

TECHNICAL GUIDANCE NOTE

MONITORING AND EVALUATION FOR IN-SERVICE TEACHER PROFESSIONAL DEVELOPMENT PROGRAMS



SUMMARY
SLIDES



COACH TOOLS
AND RESOURCES



© 2021 International Bank for Reconstruction and Development / The World Bank

1818 H Street NW, Washington, DC 20433

Telephone: 202-473-1000; Internet: www.worldbank.org

Some rights reserved.

This work is a product of the staff of The World Bank with external contributions. The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of The World Bank, its Board of Executive Directors, or the governments they represent. The World Bank does not guarantee the accuracy of the information included in this work.

Nothing herein shall constitute or be considered to be a limitation upon or waiver of the privileges and immunities of The World Bank, all of which are specifically reserved.

Rights and Permissions



This work is available under the Creative Commons Attribution 4.0 International license (CC BY 4.0) <https://creativecommons.org/licenses/by/4.0/>, with the following mandatory and binding addition:

Any and all disputes arising under this License that cannot be settled amicably shall be submitted to mediation in accordance with the WIPO Mediation Rules in effect at the time the work was published. If the request for mediation is not resolved within forty-five (45) days of the request, either You or the Licensor may, pursuant to a notice of arbitration communicated by reasonable means to the other party refer the dispute to final and binding arbitration to be conducted in accordance with UNCITRAL Arbitration Rules as then in force. The arbitral tribunal shall consist of a sole arbitrator and the language of the proceedings shall be English unless otherwise agreed. The place of arbitration shall be where the Licensor has its headquarters. The arbitral proceedings shall be conducted remotely (e.g., via telephone conference or written submissions) whenever practicable, or held at the World Bank headquarters in Washington, DC.

Attribution – Please cite the work as follows: Akmal, Maryam. 2022. “Monitoring and Evaluation for In-Service Teacher Professional Development Programs: Technical Guidance Note.” Coach Series, World Bank, Washington, DC. License: Creative Commons Attribution CC BY 4.0 IGO.

Translations – If you create a translation of this work, please add the following disclaimer along with the attribution: *This translation was not created by The World Bank and should not be considered an official World Bank translation. The World Bank shall not be liable for any content or error in this translation.*

Adaptations – If you create an adaptation of this work, please add the following disclaimer along with the attribution: *This is an adaptation of an original work by The World Bank. Views and opinions expressed in the adaptation are the sole responsibility of the author or authors of the adaptation and are not endorsed by The World Bank.*

Third-party content: The World Bank does not necessarily own each component of the content contained within the work. The World Bank therefore does not warrant that the use of any third party-owned individual component or part contained in the work will not infringe on the rights of those third parties. The risk of claims resulting from such infringement rests solely with you. If you wish to reuse a component of the work, it is your responsibility to determine whether permission is needed for that reuse and to obtain permission from the copyright owner. Examples of components can include, but are not limited to, tables, figures, or images.

All queries on rights and licenses should be addressed to Coach, The World Bank Group, 1818 H Street NW, Washington, DC 20433, USA; e-mail: coach@worldbank.org.

Cover and interior design: Karim Ezzat Khedr, Washington, DC, USA

CONTENTS

Acknowledgments.....	iv
Abbreviations.....	v
Context.....	1
Objectives.....	1
Purpose of a TPD M&E System.....	1
Designing a Results Framework for a TPD M&E System.....	2
Choosing Indicators for a TPD M&E System.....	6
Monitoring and Evaluation: Two Sides of the Same Coin.....	7
Step-by-Step Process to Design, Implement, Use, and Sustain a TPD M&E System.....	8
Context Matters.....	10
Conclusion.....	11
Appendix A. Sample Results Framework.....	12
Appendix B. Key Risks Underlying a TPD Results Framework.....	13
References.....	15

Acknowledgments

The *Monitoring and Evaluation for In-Service Teacher Professional Development Programs* guidance package was led by Maryam Akmal. The package benefits from the inputs of Jayanti Bhatia, Elaine Ding, Ann Elizabeth Flanagan, Koen Martijn Geven, Janeli Kotze, Ana Teresa del Toro Mijares, Ezequiel Molina, Adelle Pushparatnam, Manal Bakur N Quota, and Tracy Wilichowski. The team received guidance from the Technical Experts Group composed of Salman Asim, Deborah Newitter Mikesell, and Alonso Sanchez.

This package is part of a series of products by the *Coach* Team. Overall guidance for the development and preparation of the package was provided by Omar Arias, Practice Manager for the Global Knowledge and Innovation Team.

The package was designed by Karim Ezzat Khedr. Alicia Hetzner was the chief copy editor. Patrick Biribonwa and Medhanit Solomon provided administrative support.

Abbreviations

COVID-19	Coronavirus Disease 2019
CSO	Curriculum Support Officer
DfID	Department for International Development (now FCDO)
EGRS	Early Grade Reading Studies
EMIS	Educational Management Information System
ETEB	Enhancing Teacher Effectiveness in Bihar Operation
FCDO	United Kingdom Foreign, Commonwealth and Development Office (formerly DfID)
FCV	fragility, conflict, and violence
GEAK	Global Engagement and Knowledge for Education (World Bank)
GEMS	Geo-Enabling Initiative for Monitoring and Supervision (World Bank)
GPS	global positioning system
GTL	Center on Great Teachers and Leaders
IPA	Innovations for Poverty Action
IT	information technology
MLSS	Malawi Longitudinal School Survey
M&E	monitoring and evaluation
OAS	Organization of American States
PAD	Project Appraisal Document (World Bank)
PDO	program development objective
RCT	randomized controlled trial
RTI	Research Triangle Institute
SABER	Systems Approach for Better Education Results
SDI	Service Delivery Indicators
SMART	Specific, Measurable, Attributable, Realistic, Targeted
TA	technical assistance
TIPPS	Teacher Instructional Practices and Processes System
TPD	teacher professional development
USAID	United States Agency for International Development

Context

Governments and other organizations that design, implement, or manage in-service teacher professional development (TPD) programs navigate many choices and challenges to build monitoring and evaluation (M&E) systems that suit their program objectives and constraints. The primary objective of a TPD M&E system is to help guide the program toward its objectives of improved teaching practice, better quality student-teacher interactions, and, ultimately, improved student learning outcomes. An M&E system can help guide the TPD intervention toward these objectives by providing valuable data to feed into the design and implementation of the program. These data offer opportunities for course correction and could strengthen accountability relationships among stakeholders.

Objectives

This technical guidance note sets out how to navigate some of the challenges that governments and other organizations may face when designing and implementing an M&E system for an in-service TPD program. The specific design features of a TPD M&E system may vary depending on contextual factors. These factors include available resources, local technical capacity, political environment, fragile contexts, and the exact features of the TPD program—from “highly structured” to “low structured”; and school- and cluster-based versus other models.¹ This note presents high-level guidance on key factors to consider when designing an effective M&E system for a TPD program.

Before providing step-by-step guidance on how to design, implement, use, and sustain such a system, this note identifies the three key components that underpin a TPD M&E system: goals, results framework, and indicators.

This technical guidance note is part of the larger *Monitoring and Evaluation for In-Service Teacher Professional Development Programs* guidance package. The package also includes (1) a PowerPoint, which summarizes the key messages of this technical guidance note; and (2) an accompanying M&E Indicators Sheet, which contains sample indicators for a TPD M&E system.

Purpose of a TPD M&E System

A TPD M&E system can help track a program’s progress and guide it toward its objectives of improved teaching practice, better quality student-teacher interactions, and improved student learning outcomes. The M&E system uses two primary channels:

1. **Monitoring implementation fidelity.** Data from an M&E system can help ensure that the TPD program is being implemented with fidelity.² Implementation data such as the frequency of trainer observations and feedback sessions for teachers, or the number of teacher guides per teacher (to ensure that each teacher has access to a guide) can help monitor implementation fidelity and enable implementers to course-correct as necessary.

For instance, in Kenya’s Tusome program, the Curriculum Support Officers (CSOs) supporting teachers were supplied with tablets outfitted with GPS monitoring. These tablets enabled implementers to ensure that the CSOs conducted their allocated visits and made the target number of visits to a specific location (Piper and others 2018). See Spotlight 1 for more details on Tusome’s M&E system. Implementation data also can provide insights on the reasons that an intervention did or did not lead to desired outcomes. Finally, without an evaluation of implementation fidelity, it is difficult to determine whether lack of impact is due to poor implementation or to shortcomings in the program design (Carroll and others 2007).

1. For more information on highly structured and low-structured programs and on school- and cluster-based models, please consult the *Structuring Effective 1-1 Support and Structuring and Supporting School- and Cluster-Based Continuous Professional Development* guidance notes. Both notes are part of the suite of Coach Tools and Resources.

2. Implementation fidelity refers to the degree to which a program is delivered as intended. Implementation fidelity comprises five key elements: adherence to the intervention, dosage or exposure, quality of delivery, participant responsiveness, and identification of essential program features (Carroll and others 2007).

- 2. Monitoring outcomes.** Data from an M&E system also can help determine progress toward three desired outcomes³: improvements in teaching practice, quality of student-teacher interactions, and student learning. Outcome data can include both longer term results, such as change in student learning; and medium- to short-term results, such as progress (or lack thereof) in how teachers structure the lessons or engage with students. For example, under Peru's *Acompañamiento Pedagógico Multigrado* program, coaches periodically observed classroom sessions and assessed teachers on a broad range of instructional practices. These practices included lesson planning, time management, student engagement, feedback, and management of classroom environment. Coaches then discussed progress with the teachers and developed a plan for improvement. Coaches also shared monthly and quarterly reports on teachers' progress and areas for improvement with the local education authority and school principals. A randomized evaluation of the Peruvian program showed that the program improved pedagogical skills of teachers by 0.2-0.3 standard deviations (Castro, Glewwe, and Montero 2019).

Designing a Results Framework for a TPD M&E System

Underlying an M&E system that can effectively monitor implementation fidelity and outcomes is a sound results framework for the TPD program. The results framework lays out a causal pathway for how the intervention ultimately will help accomplish some of the key *program development objectives* (PDOs).⁴ These objectives can include improving teaching practice, student-teacher interaction, and/or student learning.⁵ The results framework traces the causal pathway from inputs, activities, and outputs to outcomes and impacts (figure 1). This causal pathway should be strong and coherent so that the results framework can help answer two key questions (World Bank 2020a):

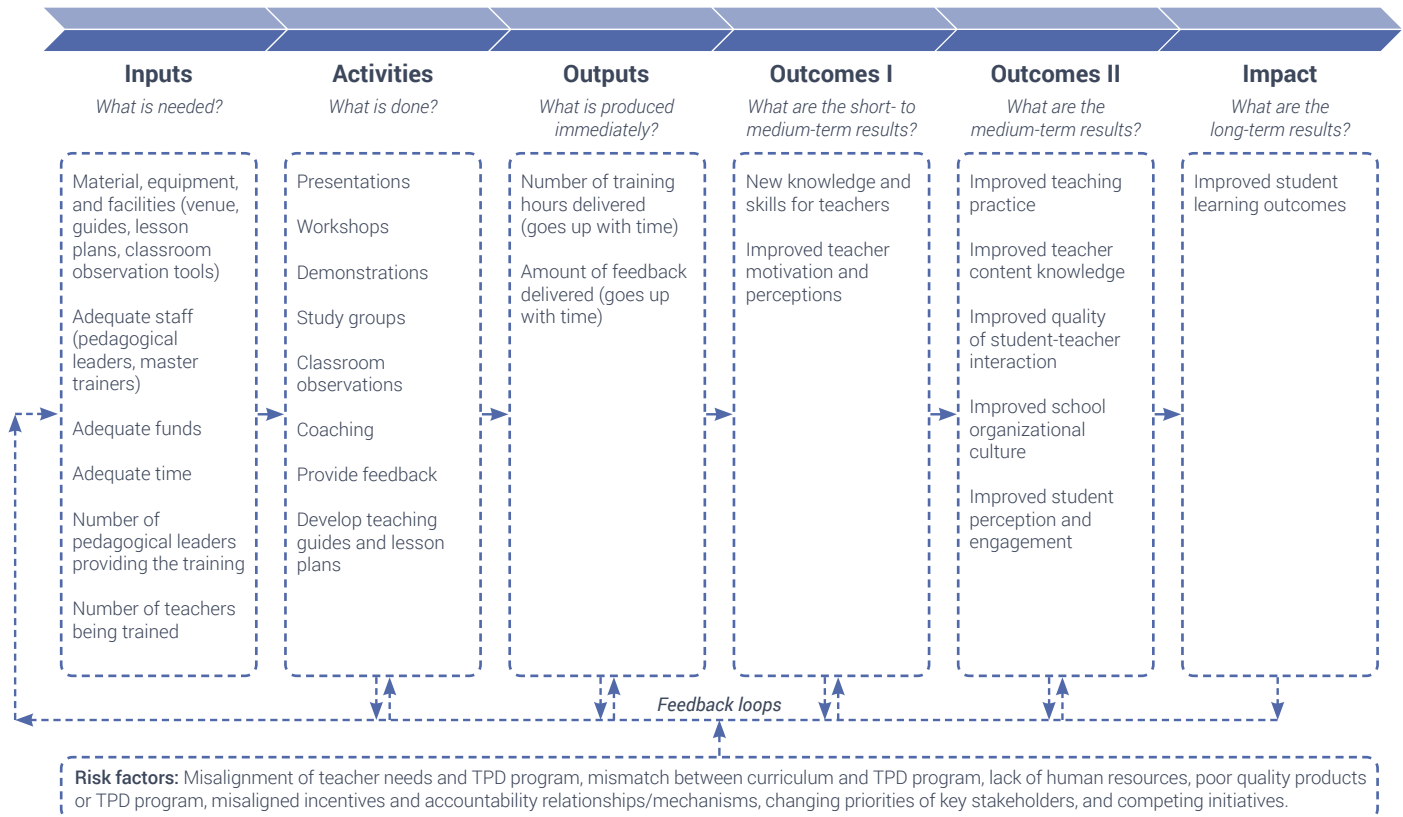
1. What are the ultimate goals of the TPD program? The goals are the *program development objectives* (PDOs): improving teaching practice, student-teacher interactions, and better student learning outcomes.
2. What are the indications that program goals are being achieved? The answer lies primarily in (a) the short- and medium-term program outcomes, such as changes in knowledge and skills of teachers and changes in teaching practice; and (b) the longer-term outcome: changes in student learning. Achieving the program goals and short- and medium-term outcomes depends on both the design of the program and its implementation fidelity. The latter relies on the inputs, activities, and outputs of the program.

3. A change in outcome can refer to a change in the target population's observed status, behavior, capacity, skill, or attitude.

4. *Program development objectives* (PDOs) define a project's overarching goals and set the foundation for how the World Bank assesses performance (Heider, Behrens, and Peng 2017).

5. The purpose of a TPD program is to improve student learning. Nevertheless, programs may choose not to benchmark the success of their programs against improvement in student learning outcomes by including the latter as a *program development objective*. Two reasons that programs may choose not to benchmark the success of their programs against student learning outcomes are that improvement in learning outcomes (a) may take a long time to materialize and (b) often is affected by a complex interaction of factors, many of which may be beyond the scope of the TPD intervention. Instead, programs may choose short- or medium-term proxy outcomes to monitor success. Examples could be change in teaching practice or change in the quality of student-teacher interaction, which may be linked to improved student learning outcomes. For example, a World Bank project in Indonesia to improve teacher performance and accountability for better learning outcomes chose improved teacher presence and service quality as the primary *program development objective*, while identifying better student learning as a longer term goal (World Bank 2019b).

Figure 1. Sample Results Framework for a TPD Program



Source:

Adapted from Haslam 2010 and Kekahio and others 2014.

Note:

See Appendix A for an example of a results framework from a World Bank program to improve pre- and in-service teacher professional development programs in the state of Bihar in India. See Appendix B for details on key risks underlying a TPD results framework.

For the M&E system to support a well-functioning TPD program, the results framework underlying the M&E system needs to identify tight feedback loops linking the steps in the causal chain leading to the program’s objectives. These objectives comprise improved teaching practice, better quality student-teacher interactions, and, ultimately, improved student learning outcomes. Tight feedback loops that iteratively direct information into decision-making processes can help agencies and implementers learn, innovate, and improve the design and implementation of TPD programs over time; and strengthen the accountability relationships among stakeholders.

A continuous data stream directed at key actors and decision-makers can be used to course-correct during the program design and implementation stages. For example, data on intermediate outcomes such as the change in teachers’ knowledge or practice can feed into the design of the training, such as the type or frequency of feedback. Similarly, implementation data such as the number of training hours received could be used to adjust critical inputs to the program, for example, assignment of adequate number of pedagogical leaders or trainers. Spotlight 1 illustrates how feedback loops were used in Kenya’s Tusome program. Ultimately, however, an M&E system is only as good as the data it collects and the extent to which these data are used for sound decision-making. Such an exercise takes time, effort, resources, and commitment; and is a continuous work in progress. See Spotlight 2 for how partnerships between the government and other actors can facilitate data collection and data use to monitor implementation fidelity and progress toward desired outcomes.

SPOTLIGHT 1. M&E Systems Can Strengthen Accountability Relationships: A Case Study of Tusome

Tusome, Kenya's national program to improve early grade education, led to a 0.6–1.0 standard deviation improvement in English and Kiswahili (Piper and others 2018). The program used government inspectors (Curriculum Support Officers, or CSOs) to support teachers. Multiple factors contributed to the program's success, such as the availability of high-quality learning materials aligned with teacher guides along with adequate training and pedagogical guidance for CSOs. In addition, Tusome had built in capacity to monitor outcomes and implementation fidelity in six important ways:

1. Tusome allowed frequent monitoring of key learning outcomes and embedded them as a key feature of the accountability relationship between CSOs and ministry officials. At the end of each classroom observation, the CSOs randomly selected and evaluated the oral reading fluency of three students, and recorded information using tablets equipped with an [offline-first software, *Tangerine Coach*](#). These data were viewable on a dashboard shared with CSOs and the Ministry of Education officials to whom the CSOs were accountable.
2. In addition to monitoring student outcomes, Tusome built in tight feedback loops (a) to monitor teachers' progress toward the desired outcomes and (b) to tailor feedback to teachers as appropriate. For example, at each visit, the CSO recorded whether the teacher employed the techniques for which s/he had received training during the lesson and provided feedback to teachers accordingly. These data then were aggregated to show how teaching practice changed over time and whether teachers in a district were meeting expectations for improved instruction.
3. CSOs received ongoing guidance and feedback from implementation firm experts and county-level education officers who had observed the CSOs' feedback sessions with teachers. For example, the Tusome technical team often observed the CSOs' one-on-one sessions with teachers and gave feedback on the quality of the CSOs' instructional support. Tusome field staff also used these data to provide daily updates to appropriate government officers on teachers' progress in delivering the expected sequence of daily lessons. Based on these data, education officers gave instructions to CSOs and head teachers to support the teachers whose performance was lagging.
4. To ensure implementation progress, Tusome employed both outcome indicators and process indicators. The CSOs' tablets, which were used for classroom observations, were outfitted with GPS monitoring, enabling implementers to ensure that CSOs conducted their allocated visits. These data were aggregated into monthly reports showing whether the CSOs had made the target number of visits to a specific location each month.
5. To strengthen accountability, Tusome linked monitoring data with incentives. For example, the data on the number of classroom visits for each school were used to determine the CSOs' travel reimbursement.
6. Monitoring data were shared widely to facilitate feedback and strengthen accountability relationships among the actors. For example, a dashboard showing the percentages of target visits at the county and national levels was used by Ministry of Education leadership to increase accountability for instructional support. The Tusome program supported the use of these data by both national education leaders and county officers.

Source: Piper and others 2018.

Underlying the results framework are risk factors or assumptions (Appendix B) that could block the pathway leading to program goals. Many of these risks will be context specific. They could include misalignment of the program with teacher needs or key components of the education system (for example, curriculum), lack of human resources or physical materials, misaligned incentives or accountability relationships and mechanisms, or political economy and financing risks. See Appendix B for details on high-level risk factors that could hamper a TPD program and potential ways to address them.

Such risks, implementation challenges, or unexpected events could prompt restructuring of the program implementation, design, or objectives, enabling World Bank country teams to adjust operations in response to new data and events. For example, 42 World Bank education projects were restructured in response to COVID-19 (Coronavirus Disease 2019) (World Bank 2020b). M&E data can help take stock of progress toward objectives and, if needed, inform project restructuring. For example, a World Bank project to enhance teacher professional development in the Indian state of Bihar underwent project restructuring. Based on data showing slower than expected implementation, results targets were adjusted, and the scope of some activities was reduced (World Bank 2019a).

SPOTLIGHT 2. Effective Partnerships to Facilitate Data Collection and Use for Monitoring Implementation Fidelity and Progress toward Outcomes

Governments may face financial constraints when designing, implementing, and using M&E systems. Even when adequate financial resources are available, technical constraints can hinder the development and use of effective M&E systems. In such contexts, collaborations with private firms, development partners, or other actors can facilitate data collection and use to monitor implementation fidelity and progress toward desired outcomes.

Governments may partner with other actors, such as universities, research agencies, or the private sector to facilitate data collection and use to monitor implementation fidelity, enabling the government to use analytical insights for operational and implementation decisions. For example, the World Bank formed a partnership with the government of Punjab, Pakistan to implement a classroom observation tool called Teach to monitor and evaluate teacher practices. The tool first was implemented by a research firm in collaboration with the World Bank and the Punjab School Education Department. The tool then was adapted by the Punjab School Education Department in consultation with the institution responsible for teacher professional development. To monitor implementation, trainings were observed by outside partners including university-based researchers. A software application for data collection was built by a private sector technical assistance (TA) team in collaboration with a specialized public sector information technology (IT) partner, enabling data to be accessible to teachers' mentors and policymakers. This multi-layered partnership along with a staggered rollout helped create opportunities early on to integrate lessons learned in program implementation and to take into account feedback from partners, teachers, and district leaders (World Bank forthcoming). Data from the tool has enabled the government to provide evidence-based, targeted, and personalized teacher professional development.^a

Similarly, the successful implementation of Tusome in Kenya (Spotlight 1) resulted from a partnership between the Kenyan Ministry of Education, and RTI (Research Triangle Institute), which was the prime implementer; as well as the United States Agency for International Development (USAID) and United Kingdom Foreign, Commonwealth and Development Office (FCDO), which funded the program. As the chief implementer, RTI provided human and technical support to monitor implementation fidelity and helped improve program implementation based on data collected. For example, RTI experts along with country-level education officers provided guidance to CSOs based on observing CSOs' feedback sessions with teachers. As seen in Tusome's case, support for monitoring implementation fidelity often goes hand in hand with program implementation support.

Partners also may help build systems for data collection, storage, maintenance, and integration, especially in contexts in which such technical skills or systems are limited. For example, the South African Department of Basic Education has partnered with a third party, DataFirst, to host and securely share de-identified administrative data. These data have provided data management benefits to the government and enabled multiple impact evaluations of various education initiatives (Rossiter 2020).

Often, however, even when data collection infrastructure is in place, other constraints limit the usability of data to monitor implementation fidelity or progress toward outcomes. Such constraints include poor integration of data sources and lack of technical capacity. For example, through a partnership among the government, World Bank, and other donors, Timor-Leste has introduced an Educational Management Information System (EMIS) infrastructure and improved data collection. However, the country still has limited capacity to use the data for analytical work to guide policy decisions (Abdul-Hamid, Saraogi, and Mintz 2017).

Beyond strengthening data collection and use to monitor implementation fidelity, partnerships also can facilitate production of new evidence of program effectiveness. For example, the government of Malawi partnered with the World Bank, Royal Norwegian Embassy, and FCDO to set up the Malawi Longitudinal School Survey (MLSS). The survey collected nationally representative data on school conditions and learning outcomes that were not being captured by the government's existing data collection systems. The survey was administered by a partnership of local and international research firms. The data helped to produce policy-relevant insights on students' learning trajectories. The data also provided evidence of the impacts of pilot interventions. This evidence then informed the design of a large, multi-faceted reform program that supports the government in scaling-up successful pilots and introduces targeted investments. Such longitudinal, nationally representative data and research evidence would be difficult for many governments to produce on their own without financial and technical support from effective partnerships.^b

Nevertheless, partnerships with external players are not always easy to implement. For example, the teacher training program under the Early Grade Reading Studies (EGRS) program in South Africa relied largely on external service providers to help monitor and implement the program. However, there was pushback from the sector on using external partners. Whether the monitoring function is performed by an external or internal actor (for example, an NGO, as opposed to a local district officer) has potential implications for the accountability relationship between the government and the actor due to the varying nature of the contractual and financial obligations between the parties. At present, even the external partners in South Africa do not have the capacity to support a national scale-up of the EGRS program.^c

Fostering such capacity may require concerted World Bank support for partnerships among governments, research institutes, and technical experts to convene the skills and resources essential for the effective design and use of M&E systems. Partnerships can be critical for implementation success in contexts with limited technical and financial resources. However, to ensure the long-term sustainability of the program, it is vital to build into partner contracts the transfer of technical assistance to build the skills and capacity of the public bureaucracy and governmental actors, including coaches and monitors.

Note:

- a. In addition to Punjab, Teach also has been implemented in Guyana and Mozambique as an M&E tool to evaluate and guide teacher professional development.
- b. Comment written on an earlier draft of this note by a reviewer involved with the Malawi project.
- c. Information provided by a researcher involved with the Early Grade Reading Studies program in South Africa.

Choosing Indicators for a TPD M&E System

A TPD M&E system can involve two main types of indicators:

1. **Process indicators to monitor implementation.** Process indicators include data that show whether activities are being implemented as planned, usually tracked by inputs, activities, and outputs (figure 1). As mentioned earlier, monitoring implementation fidelity is one key channel through which an M&E system can guide the program toward its objectives. For example, are the inputs, activities, and outputs (traditionally found in work plans) on track? These indicators are vital to track management and implementation of programs, use of resources, and delivery of services. Nevertheless, by themselves, they do not indicate whether outcomes have been achieved. For example, a TPD program could be on track to meet the target number of teacher guides produced. However, if these guides are not well designed or if teachers do not use them, these guides are unlikely to affect teaching practice.
2. **Outcome indicators to measure results.** Outcome indicators include data that show whether program inputs, activities, and outputs have made a difference in short-, medium-, and long-term outcomes (figure 1). Monitoring result outcomes (in addition to monitoring implementation fidelity) is another key channel through which an M&E system can guide the TPD program toward its goals. For example, based on information from classroom observation tools, has student-teacher interaction (an intermediate outcome) in the classroom improved due to using teacher guides (an input)?

M&E plans should include clear and measurable indicators that go beyond process indicators, such as those tracking inputs, to those measuring outcomes. Outcome indicators can provide an indication of whether the program is on track to meet intended objectives. For example, a teacher training program in Ghana tracked the number of teachers trained and number of courses run but did not use the M&E system to feed into the design of the training (World Bank 2019c), which could have benefitted from information on outcomes. Prioritizing outcome indicators can help draw the attention of policymakers and managers toward results as opposed to process-oriented tasks. Outcome indicators also may be linked to financing to incentivize progress toward results (Perakis and Savedoff 2015), such as through the World Bank's Program for Results financing instrument. For example, improvement in teaching practice based on data from classroom observation tools such as [Teach](#) or [Stallings](#), or improvement in student learning, are outcome indicators that could be linked to disbursement. Part of the payment for a TPD program could be made after pre-agreed results have been achieved and independently verified. However, with such results-based financing programs, details matter. These programs need to be designed carefully to ensure that such programs do not inadvertently lead to unintended consequences, such as corruption or misaligned incentives.

Regardless of type, it is recommended that all indicators meet the five SMART principles:

1. **Specific:** Indicators measure as closely as possible what we want to know.
2. **Measurable:** Indicators are specific and can be clearly measured.
3. **Attributable:** Indicators are linked logically and closely to the program's objectives.

4. **Realistic:** Data is obtainable at feasible cost with reasonable accuracy and frequency.
5. **Targeted:** Indicators are specific to the program's target group, for example, in the case of a TPD program, teachers, students, and trainers.

Ultimately, the decision about which indicators to choose for an M&E system depends on the goals of the TPD program (Puma and Raphael 2001). In the context of a TPD program, indicators should be chosen with the eventual goal of improving teaching practice and student-teacher interactions; and ultimately, student learning outcomes. For the M&E system to be a useful management tool, it needs to be manageable and not overloaded with too many indicators. If too many indicators are chosen, too much time will be spent managing the M&E system itself, rather than using the data to manage the TPD program (Kusek and Rist 2004). New data collection for indicators also costs time and money so indicators should be chosen judiciously. The accompanying M&E Indicators Sheet lists sample outcome and process indicators from World Bank Project Appraisal Documents (PADs) that can be used for a TPD M&E system.

It is critical that the M&E system does not devolve into a process of merely feeding checklists and data indicators to higher authorities without using these data for analysis. For example, bureaucratic norms in which the performance of local education administrators is judged largely on the speed and quantity of data collected can create perverse performance incentives that impair effort allocation (Aiyar and Bhattacharya 2016). Therefore, it is important, first, to ensure clearly defined roles and responsibilities for those who collect and use data. Second, ensure transparent processes for how data is accessed and used for decision-making and by whom. Table 4 and the accompanying M&E Indicators Sheet provide additional guidance on data collection, analysis, and use.

Monitoring and Evaluation: Two Sides of the Same Coin

Monitoring and evaluation play two different but complementary roles. Monitoring data usually gives an indication of the program's progress at any given time relative to targets and can inform ongoing course correction (Kusek and Rist 2004). Evaluation usually attempts to address issues of causality, or the reasons that targets are or are not being achieved.⁶ A monitoring system gives ongoing information about the direction, pace, and magnitude of change to see whether the program is moving in the right direction and whether implementation is happening as intended. However, monitoring data does not give the basis for causal inference, that is, why or how changes are occurring. For example, monitoring data could reveal whether coaches spend enough time providing feedback to teachers. However, one needs evaluation information to help answer whether the improvement in teaching practice is attributable to time spent by coaches on providing feedback.

SPOTLIGHT 3. Evaluating TPD Programs

Many different types of evaluations can help measure progress toward desired outcomes. Evaluations can range from randomized experiments, including nimble experiments (Karlan 2017), to quasi-experimental designs, to nonexperimental evaluations, such as process implementation evaluations or rapid appraisals, each of which produces different types and strengths of evidence. The type of evaluation conducted depends on the needs of the program and the exact research questions posed. See Kusek and Rist (2004) and Hinton (2015) for detailed discussions on the different types of evaluation methods to consider for measuring the effectiveness of a TPD program.

Monitoring data can help track movements toward desired outcomes. On a deeper level, the question of whether the changes in outcomes can be causally attributed to the TPD program (or certain features of the TPD program) would require an *impact evaluation*. For example, it is possible that teaching practice improved in schools implementing the TPD program, but how do we determine whether the results are due to the TPD program or to some other intervention or factor? If a TPD program is being piloted and offers scope for experimental variation in provision of the treatment (for example, a teacher training program), then a randomized evaluation can help determine the causal impact of the program. Ideally, such evaluations should be planned before the program is rolled out. To preserve objectivity, preferably, they also should be conducted by an independent third-party evaluator, not by the implementer.

6. Depending on the type of evaluation, data from an evaluation can be used to inform ongoing course correction.

Traditional randomized controlled trials (RCTs) often take time and money. Therefore, they should be used to answer a crucial policy question about the design and implementation of a TPD program that requires generation of new causal evidence. Such evaluations often are useful to study interventions that are testing an unproven but promising approach and produce evidence that is likely to feed into future policy decisions. RCTs can help answer questions about whether a particular TPD program is effective, test design innovations within the program, or test implementation modalities that could improve program impact or cost effectiveness (Gertler and others 2016). Additionally, governments may decide whether to adopt or scale up a particular TPD intervention based on such evidence.

Results from randomized evaluations often take time because they typically must wait for a program to mature to produce meaningful results. Thus, randomized evaluations are unlikely to feed back into the real-time processes and decisions that feed into the management of the program. At this juncture, nimble or “rapid-fire” RCTs could play a part. Rapid-fire RCTs tend to focus on short-term outcomes and, frequently, operational questions such as take-up and use. Rapid-fire RCTs rely largely on existing administrative data and tend to be faster and cheaper than traditional RCTs (Karlan 2017). Data from rapid-fire RCTs also can help decide whether to conduct and how to design a “traditional” RCT. However, in practice, nimble RCTs relying on existing administrative data systems, such as EMIS, can face problems related to the poor data quality, particularly in low-resource contexts.

Another option is nonexperimental evaluations such as process implementation evaluations. Nonexperimental evaluations can produce meaningful insights into implementation processes and operational details. The details include costs, time, staff capacity, financial and physical resources. Nonexperimental evaluations also can inform ongoing course corrections; help explore why intended results were or were not achieved; and explore unintended results. Similarly, rapid appraisals that rely on descriptive and management-focused information from key informant interviews, focus group discussions, or surveys can inform rapid decision-making and determine whether the program is on track (Kusek and Rist 2004).

Step-by-Step Process to Design, Implement, Use, and Sustain a TPD M&E System

The previous sections establish basic components underpinning the M&E system, such as the results framework and data indicators. Table 1 presents a step-by-process to build, maintain, and sustain an M&E system for a TPD program. These steps have been adapted from Kusek and Rist (2004). Each step corresponds to a key operational factor to consider at each stage of the execution of a TPD M&E system. Designing a simple yet effective M&E system often is difficult due to local contextual factors or constraints. Implementing all the steps mentioned below or following them in the sequence described may not be feasible. Table 1 is intended as a high-level guide that should be adapted to the local context and maturity of the existing M&E system.

Table 1. Process to Design, Implement, Use, and Sustain a TPD M&E System

Step	Explanation
<p>Step 1. Conduct a readiness assessment</p>	<ul style="list-style-type: none"> • Before introducing a new M&E system for a TPD program, conduct a baseline assessment to see whether any existing structures for monitoring and evaluation of TPD programs exist, their strengths and weaknesses, and whether they can be built on. • Identify the objectives of the TPD program. • Identify the capacity of local actors and agencies to perform their roles and responsibilities, ranging from collecting data to hosting, maintaining, sharing, or analyzing the data, to making decisions based on the data. Overly complicated M&E systems that lack the corresponding local capacity to use and maintain these systems are unlikely to be sustainable. • Identify whether there is a need to build partnerships with private firms, development partners, research organizations, or other actors who may help build systems for data collection, storage, maintenance, and integration; and provide analytical capacity, especially in contexts in which such technical skills or systems are limited. • Identify potential barriers to implement an M&E system, such as opposition from teacher unions or other stakeholders, and potential ways to address these barriers. • Identify incentives and demands to design and build a results-based M&E system, for example, a government mandate to improve accountability in the education system. Then identify potential ways to use such incentives to create buy-in for the M&E system.

<p>Step 2. Agree on outcomes to monitor and evaluate</p>	<ul style="list-style-type: none"> • Map out the results framework underlying the program (figure 1) by laying out the causal pathway from inputs, activities, and outputs to outcomes and impact. • Based on the results framework, identify the outcomes to monitor such as improved teaching practice, improved student-teacher interaction, and improved learning outcomes. • Through convening local meetings and focus groups, involve the government, teacher unions, CSOs, school leaders, NGOs, donors, and other education stakeholders to help build consensus and create buy-in for outcomes to monitor. At a minimum, the M&E system should have commitment from the government, head teachers, and district level officers because these local leaders play a key part in making decisions and establishing accountability based on information from the M&E system. • In practice, participatory approaches can meet challenges, including those related to time, cost, dominance by powerful stakeholders, or conflicting priorities. Moreover, organizing a participatory approach in a federal system in which different provinces or regional units have autonomous and unaligned educational systems and priorities may be difficult.
<p>Step 3. Select key performance indicators to monitor outcomes</p>	<ul style="list-style-type: none"> • Identify key performance indicators to measure improvements in outcomes linked to the effectiveness of a TPD program. • Identify key process indicators to ensure that the program is being implemented with fidelity. • Confirm that the selected indicators are specific, measurable, attributable, realistic, and targeted. • Ensure that the M&E system is not overloaded with too many indicators. • Before large-scale rollout, pilot the indicators on a small scale and make necessary changes. A pilot also can alert relevant parties to indicators for which it is impossible to collect data or for which data already exist (Kusek and Rist 2004).
<p>Step 4. Set baselines and targets, and gather data on indicators</p>	<ul style="list-style-type: none"> • Identify which data sources exist, for example, EMIS, SABER,^a SDI,^b classroom observations, or local assessments. Then determine whether these sources can be used to establish baselines for key indicators. For example, information on teacher training (in-service and pre-service) and availability of specialists or special resource centers could be obtained from the country's EMIS systems. • If new data are collected, identify tools for data collection, such as Tangerine and KoBoToolbox. • If choosing a classroom observation tool, keep in mind licensing costs, cultural relevance, and enumerator capacity required. For example, new data collection for classroom observation may require coaches and teachers to have access and training to use classroom observations tools such as Teach (Molina and others 2020), TIPPS,^c or Stallings. • Establish baselines for key indicators before the program is rolled out. • Agree on a target corresponding to each indicator. The target should be realistic, contextually appropriate, time bound, and evidence based. Targets also should specify clearly the domain in question, for example, specific student population, age cohort, or grade cohort (Azevedo and others forthcoming). Ensure that set targets do not create perverse incentives for teachers and other staff. Also ensure that teachers and staff feel supported rather than threatened by the M&E system. • Decide how frequently data will be collected. The data should be collected at regular intervals and ideally be comparable over time for trend analysis. • Develop validation processes to ensure that the data collected are accurate and reliable. • Apart from quantitative data, gather qualitative information from key stakeholders, such as teachers, coaches, and master trainers, about program implementation and results.
<p>Step 5. Monitor and report results</p>	<ul style="list-style-type: none"> • Monitor performance against established targets. • Develop a central repository or dashboard in which data are stored and can be accessed by key actors and decision-makers. • Identify plans for analysis, and dissemination of reports and data visualizations. • Ensure that results are reported in a clear and timely manner to the intended audience.

<p>Step 6. Use the findings</p>	<ul style="list-style-type: none"> • Identify key actors responsible for decision-making based on data insights, including ministry officials and implementation partners. • Identify whether the appropriate decision-making bodies have the time, capacity, and autonomy to regularly review, discuss, and act on the data. • Identify which programmatic decisions and course corrections are being made based on M&E information. For example, do results feed into how resources such as teacher guides are allocated or how training is designed? • Identify whether the program is meeting its outcome goals and use the findings to make decisions. For example, are results being used to make changes to program design or implementation or to make decisions about budgetary allocations and scale-up? • Identify whether the program is being implemented with fidelity. For example, are trainers conducting their target number of visits to teachers? Does every teacher have access to a guide or other inputs?
<p>Step 7. Sustain the M&E system</p>	<ul style="list-style-type: none"> • Lay out clear roles and responsibilities, specifically, for who is responsible for collecting, hosting, maintaining, sharing, analyzing, and using the data. • Ensure that teachers, school leaders, pedagogical leaders, and other actors who are collecting or providing data are given appropriate time and resources to perform their tasks. • Ensure that the M&E system is producing credible, valid, timely, and reliable information. • Identify (a) whether there is technical and financial capacity to sustain the system and identify (b) actions to enhance the capacity and performance of the agencies involved. • Identify whether stakeholders' incentives are sufficiently aligned to help sustain the M&E system. • Identify whether there is a strong political champion for the M&E system. • Identify and establish a process to evaluate the M&E system.

Source:

Adapted from Kusek and Rist 2004.

Note:

- a. See the accompanying M&E Indicators Sheet for additional guidance on data collection, analysis, and use.
- b. SABER = Systems Approach for Better Education Results; SDI = Service Delivery Indicators; TIPPS = Teacher Instructional Practices and Processes System.

Context Matters

Country-level constraints, for example, political and economic realities, can impact how M&E systems are designed and used. The politics of M&E often are more difficult than getting the technical pieces right. M&E systems provide policy actors with crucial information about progress of the TPD program so can help build support for the program. Conversely, M&E results can pose political challenges because not all stakeholders may be pleased about disseminating results-based information and increasing transparency and accountability. For M&E systems to be sustainable, they may need political support from a powerful champion (table 1; Kusek and Rist 2004). Obtaining political support may be particularly challenging in highly authoritarian regimes.

Similarly, in low-resource settings, it may be more feasible to first focus on establishing a basic M&E system foundation by using simple indicators. Countries need resources to build data systems, hire technical capacity, and implement mechanisms to monitor progress. At a minimum, an effective M&E system needs the capacity to construct, collect, analyze, and report basic data in relation to baselines for key indicators and other benchmarks (Kusek and Rist 2004). Officials need to get training in data collection, methods, and analysis. Obtaining such training can be particularly challenging in low-resource settings that have limited technical capacity. Donor-funded

assistance may be an essential first step to set up a basic system. The next step would be constant capacity building of government or other local staff to ensure sustainability and prevent an excessive long-term dependence on donor funding and support. Spotlight 4 details the challenges that may affect the development of a TPD M&E in fragility-, conflict-, and violence-affected (FCV) contexts.

SPOTLIGHT 4. Developing M&E Systems in Fragile and Conflict-Affected Areas

Situations in FCV environments can change quickly. Therefore, it is important that both the M&E system and the TPD program are flexible and can adapt rapidly based on the changing situation. This flexibility requires identifying and monitoring the key risks specific to the local context and appropriately adapting the TPD program and M&E system. The first step is to understand the local conflict dynamics, for example, the key actors involved, conflict triggers, and conflict resolution mechanisms. It is also important to ensure that the incentives or accountability relationships created by the M&E system or the TPD program do not exacerbate tensions among stakeholders and worsen the conflict or fragility (Corlazolli and White 2013).

FCV countries are likely to face financial and technical capacity challenges. Consequently, they may have limited access to the physical infrastructure required to implement a TPD program, such as tablets or data collection instruments. FCV contexts also may have constrained physical access on the ground that impedes in-person data collection and monitoring, as experienced during the COVID-19 pandemic in non-FCV countries as well. M&E systems in such contexts may benefit from remote data collection and monitoring through open-source data collection tools particularly suited to FCV settings. Such tools include those facilitated by the World Bank's [Geo-Enabling Initiative for Monitoring and Supervision \(GEMS\)](#) program.

Due to financial and technical capacity limitations, FCV contexts may choose to collect very few basic indicators, at least in the beginning when setting up the M&E system. For example, Lebanon has established a school information system (a component of EMIS) that includes basic information on schools, resources, enrollments, and basic learning data. The country's move from a paper-based to a web-based school census has reduced data entry issues and simplified data verification. Nevertheless, challenges persist. There is limited capacity to use the data collected for M&E decisions, and the data is de-linked from the education policy management and decision-making processes of schools and the ministry. Furthermore, there is heavy dependence on donors and external vendors, raising issues about the financial sustainability of the system (World Bank 2018). As its data system matures, Lebanon could focus on building local technical capacity to reduce reliance on external vendors. The country also could strengthen feedback loops between data providers (teachers, students, parents) and data users (policymakers, principals) to increasingly use the data for M&E decisions (World Bank 2018).

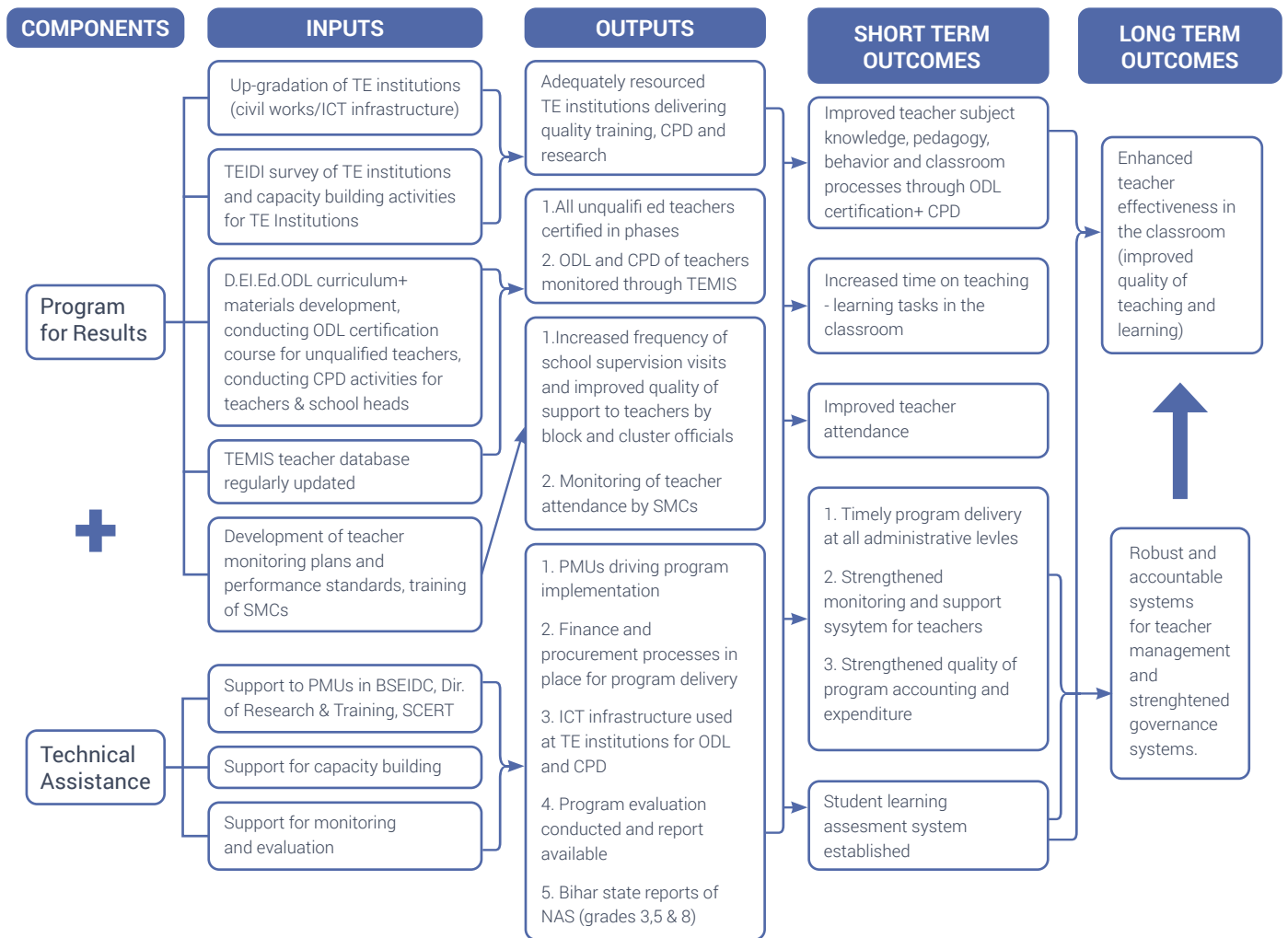
Conclusion

An effective monitoring and evaluation system can help ensure the progress of a teacher professional development program in reaching its goals, such as improved teaching practice, better quality student-teacher interactions, and improved student learning outcomes. A TPD M&E system can provide valuable data to feed into program implementation and design, strengthen accountability relationships among the stakeholders, and enable governments to make evidence-based decisions about program expansion and budgetary allocations.

However, an M&E system is only as good as the data collected and the capacity and willingness of the governments and key actors to make evidence-based decisions. Designing, implementing, using, and sustaining an effective M&E system require more than carefully selecting outcome indicators, establishing data systems, and getting the technical details right. Also required are tight regular feedback loops that link data to decision-making processes. These data can help governments and other actors to improve the design and implementation of TPD programs over time by offering opportunities for course correction. Maintaining effective feedback loops requires adequate local capacity for data collection, analysis, and use, along with access to sufficient technical, physical, and financial resources. Equally important, well-functioning M&E systems require strong political support and buy-in from key stakeholders who are willing to use the system to make evidence-based decisions.

Appendix A. Sample Results Framework

Figure A1. Results Framework for a World Bank Teacher Professional Development Program in India



Source:

World Bank 2015, figure 3.

Note:

All abbreviations in figure A1 are defined in World Bank 2015, <http://documents1.worldbank.org/curated/en/184631468000251240/pdf/92972-PAD-P132665-IDA-R2015-0096-1-Box391421B-OUO-9.pdf>.

Appendix B. Key Risks Underlying a TPD Results Framework

Underlying the results framework are risk factors that could disrupt the causal chain leading to program objectives. Eight of these key risk factors and ways to address them follow.

1. Misalignment with Needs

The professional learning activities and materials should directly address the needs and deficiencies of teachers and coaches. Information from diagnostic tools, such as Teach, can help identify pedagogical gaps and needs, which can inform the program's design. Additionally, a participatory approach that involves stakeholders who, early on, are involved in the design of the TPD system can help ensure alignment between the program and the needs of key actors and help generate buy-in for the TPD program and the M&E system.

2. Misalignment with Curriculum and Assessments

The content of TPD program should be aligned with curriculum and assessment requirements and with other key components of the education system. For example, if teacher TPD programs emphasize building critical thinking skills in students, but the curriculum and assessments prioritize rote memorization, the results framework could break down.

3. Incoherent Accountability Relationships

Similarly, accountability relationships among teachers, coaches, inspectors, school leaders, students, and parents should be coherent. For example, if coaches expect teachers to attend TPD trainings, but school leaders want teachers to spend time on competing initiatives, the accountability relationships among school leaders, teachers, and coaches become "incoherent" (Pritchett 2015). Governments should ensure that all actors (for example, principals, coaches, inspectors, teachers) are given a consistent message of what is required from both the pedagogical leaders and the teachers.

4. Lack of Human Resources

There should be sufficient availability of expert pedagogical leaders and coaches. They should receive good quality training in how to observe and provide feedback. Ideally, they should have in-depth knowledge of content pedagogy, possess content- and grade-specific teaching experience, and have experience modeling teaching practices for teachers (Wilichowski and Popova 2021). Similarly, poor classroom conditions due to limited human resources, for example, high pupil-teacher ratios, which are prevalent in many classrooms around the world, can hinder the most effective and well-trained teachers from improving student learning outcomes.

5. Poor Availability of Materials

The professional learning activities and materials for teacher training should be available and accessible to teachers and coaches. Such material should be well designed, deliver high-quality content, and engage teachers. For example, ideally each teacher may need access to his/her own guide or lesson plan, and each pedagogical leader may need his/her own tablet to record observations. Similarly, lack of adequate physical infrastructure in the classroom, such as books, electricity, or writing materials, can hinder even well-trained and otherwise effective teachers from improving student learning outcomes.

6. Lack of Incentives

Teacher motivation is linked closely to teacher practice, psychological fulfillment, and well-being (Han and Yin 2016), and consequently, is an important determinant of student learning outcomes (Bennell and Akyeampong

2007; Guajardo 2011; Richardson 2014). Incentives may be used to improve teacher motivation, attendance, and engagement. Prior evidence suggests that effective TPD programs link participation to incentives, such as promotion or salary implications (Popova and others 2018). Incentives for teachers, leaders, and trainers of leaders also can be nonpecuniary recognition (Cruz and Loureiro 2020). For example, in the state of Ceará in Brazil, coordinators who participated in at least 75 percent of the activities and achieved a threshold score on the quality of their feedback to teachers received a certification (Bruns, Costa, and Cunha 2018).

7. Poor Accountability Mechanisms

The accountability mechanisms to confirm that master trainers, coaches, and teachers are performing their duties under the TPD program should work well. In South Africa, Cilliers and others (2020) found that on-site coaching was more effective than virtual coaching. Part of the reason may have been that on-site coaching allowed coaches to directly observe teachers and hold them accountable to complete curricula and attempt new teaching techniques. Other accountability mechanisms, such as making de-identified performance metrics public, may be used to encourage both teachers and coaches to improve classroom practice. For example, again in Ceará, at the beginning of each year, some schools received a two-page feedback bulletin that compared the school's performance against a Stallings benchmark and broke out the results for individual classrooms, including teacher time spent on different teaching methods. Publishing the comparisons sparked substantial discussion in the schools (Bruns, Costa, and Cunha 2018).

8. Political Economy and Financing Risks

Changing priorities of governments and other stakeholders can affect political support and financing available for the TPD program. Successfully sustaining TPD programs may require stakeholder buy-in from a variety of actors, including teacher unions, CSOs, NGOs, donors, school leaders, and other actors in the local education ecosystem. For example, in Chile, attempts by the government to introduce monitoring tools to assess progress in learning has faced criticism in the past over concerns around homogenization of teacher practices (Yoshikawa and others 2015). Having a strong champion in the government may help overcome such challenges.

References

- Abdul-Hamid, Husein, Namrata Saraogi, and Sarah Mintz. 2017. "Lessons Learned from World Bank Education Management Information System Operations: Portfolio Review, 1998–2014." *World Bank Studies*. World Bank, Washington, DC. <https://doi.org/10.1596/978-1-4648-1056-5>.
- Aiyar, Yamini, and Shrayana Bhattacharya. 2016. "The Post Office Paradox A Case Study of the Block Level Education Bureaucracy." *Economic and Political Weekly* 51(11). <https://www.epw.in/journal/2016/11/special-articles/post-office-paradox.html>.
- Azevedo, João Pedro, Michael F. Crawford, Marta Carnelli, and Diana Goldemberg. Forthcoming. "Setting Targets for Progress in Reducing Learning Poverty and Improving Developmental Reading Subskills: Toward an Evidence-Based Strategy to Accelerate Learning." *Education Global Practice (GEAK)*, World Bank, Washington, DC.
- Bennell, Paul, and Kwame Akyeampong. 2007. "Teacher Motivation in Sub-Saharan Africa and South Asia." Department for International Development (DfID), London, UK. <https://assets.publishing.service.gov.uk/media/57a08be640f0b652dd000f9a/ResearchingtheIssuesNo71.pdf>.
- Bruns, Barbara, Leandro Costa, and Nina Cunha. 2018. "Through the Looking Glass: Can Classroom Observation and Coaching Improve Teacher Performance in Brazil?" *Economics of Education Review* 64: 214-50. <https://doi.org/10.1016/j.econedurev.2018.03.003>.
- Carroll, Christopher, Malcolm Patterson, Stephen Wood, Andrew Booth, Jo Rick, and Shashi Balain. 2007. "A Conceptual Framework for Implementation Fidelity." *Implementation Science* 2 (1): 40. <https://doi.org/10.1186/1748-5908-2-40>.
- Castro, Juan F., Paul Glewwe, and Ricardo Montero. 2019. "Work with What You've Got: Improving Teachers' Pedagogical Skills at Scale in Rural Peru." Working Paper 158. Peruvian Economic Association, Lima, Peru. <https://ideas.repec.org/p/apc/wpaper/158.html>.
- Cilliers, Jacobus, Brahm Fleisch, Cas Prinsloo, and Stephen Taylor. 2019. "How to Improve Teaching Practice? Experimental Comparison of Centralized Training and In-Classroom Coaching." *The Journal of Human Resources* 0618-9538R1. <https://doi.org/10.3368/jhr.55.3.0618-9538R1>.
- Cilliers, Jacobus, Brahm Fleisch, Janeli Kotze, Nompumelelo Mohohlwane, Stephen Taylor, and Tshegofatso Thulare. 2020. "Can Virtual Replace In-Person Coaching? Experimental Evidence on Teacher Professional Development and Student Learning in South Africa." RISE Working Paper 20/050. RISE Programme, Oxford, UK. https://riseprogramme.org/sites/default/files/2021-01/RISE_WP-050_Cilliers_etal_2021_update_0.pdf.
- Corlazolli, Vanessa and Jonathan White. 2013. "Back to Basics: A Compilation of Best Practices in Design, Monitoring & Evaluation in Fragile and Conflict-Affected Environments." Department for International Development (DfID), London, UK. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/304632/Back-to-Basics.pdf.
- Cruz, Louisee, and Andre Loureiro. 2020. "Achieving World-Class Education in Adverse Socioeconomic Conditions: The Case of Sobral in Brazil." Working Paper. World Bank, Washington, DC. <https://openknowledge.worldbank.org/handle/10986/34150>.
- Gertler, Paul, Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M.J. Vermeersch. 2016. "Impact Evaluation in Practice." World Bank, Washington, DC. <https://openknowledge.worldbank.org/handle/10986/25030>.
- Guajardo, Jarret. 2011. "Teacher Motivation: Theoretical Framework, Situation Analysis of Save the Children Country Offices and Recommended Strategies." Save The Children, Washington, DC. https://teachertaskforce.org/sites/default/files/migrate_default_content_files/savethechildren_1.pdf.
- Han, Jiying, and Hongbiao Yin. 2016. "Teacher Motivation: Definition, Research Development and Implications for

- Teachers." *Cogent Education* 3 (1). <https://www.tandfonline.com/doi/full/10.1080/2331186X.2016.1217819>.
- Haslam, M. Bruce. 2010. "Maryland Teacher Professional Development Evaluation Guide." Center on Great Teachers and Leaders (GTL Center), Arlington, VA. <https://gtlcenter.org/sites/default/files/docs/MarylandPDEvaluationGuide.pdf>.
- Heider, Caroline, Joy Behrens, and Xiaoxiao Peng. 2017. "Is the World Bank Achieving its Development Objectives?" World Bank, Washington, DC. <https://ieg.worldbankgroup.org/blog/world-bank-group-achieving-its-development-objectives>.
- Hinton, Rachel. 2015. "Assessing the Strength of Evidence in the Education Sector: BE2." USAID (United States Agency for International Development), Washington, DC. https://www.usaid.gov/sites/default/files/documents/1865/BE2_Guidance_Note_ASE.pdf.
- Karlan, Dean. 2017. "Nimble RCTs: A Powerful Methodology in the Program Design Toolbox." Innovations for Poverty Action (IPA), New Haven, CT. <https://pubdocs.worldbank.org/en/626921495727495321/Nimble-RCTs-WorldBankMay2017-v4.pdf>.
- Kekahio, Wendy, Brian Lawton, Louis Cicchinelli, and Paul R. Brandon. 2014. "Logic Models: A Tool for Effective Program Planning, Collaboration, and Monitoring." US Department of Education, Washington, DC. https://ies.ed.gov/ncee/edlabs/regions/pacific/pdf/REL_2014025.pdf.
- Kipp, Scott, Sarah Pouzevara, and Benjamin Piper. 2018. "The National Rollout of Coaching with Tangerine in Kenya." RTI International (Research Triangle Institute), Washington, DC. <https://shared.rti.org/content/national-rollout-coaching-tangerine-kenya>.
- Kusek, Jody, and Ray Rist. 2004. "A Handbook for Development Practitioners: Ten Steps to a Results-Based Monitoring and Evaluation System." World Bank, Washington, DC. <https://openknowledge.worldbank.org/bitstream/handle/10986/14926/296720PAPER0100steps.txt?sequence=2&isAllowed=y>.
- Majerowicz, Stephanie, and Ricardo Montero. 2018. "Can Teaching Be Taught? Experimental Evidence from a Teacher Coaching Program in Peru (Job Market Paper)." Working Paper. Harvard University, Cambridge, MA. <https://scholar.harvard.edu/smajerowicz/publications/job-market-paper-can-teaching-be-taught-experimental-evidence-teacher>.
- Molina, Ezequiel, Syeda Farwa Fatima, Andrew Dean Ho, Carolina Melo, Tracy Marie Wilichowski, and Adelle Pushparatnam. 2020. "Measuring the Quality of Teaching Practices in Primary Schools: Assessing the Validity of the Teach Observation Tool in Punjab, Pakistan." *Teaching and Teacher Education* 96. <https://doi.org/10.1016/j.tate.2020.103171>.
- Perakis, Rita, and William Savedoff. 2015. "Does Results-Based Aid Change Anything? Pecuniary Interests, Attention, Accountability and Discretion in Four Case Studies." Working Paper 052. Center for Global Development, Washington, DC. <https://www.cgdev.org/sites/default/files/CGD-Policy-Paper-52-Perakis-Savedoff-Does-Results-Based-Aid-Change-Anything.pdf>.
- Piper, Benjamin, Joseph Destefano, Esther M. Kinyanjui, and Salome Ong'ele. 2018. "Scaling up Successfully: Lessons from Kenya's Tusome National Literacy Program." *Journal of Educational Change* 19 (3): 293–321. <https://doi.org/10.1007/s10833-018-9325-4>.
- Popova, Anna, David K. Evans, Mary E. Breeding, and Violeta Arancibia. 2018. "Teacher Professional Development around the World: The Gap between Evidence and Practice." Policy Research Working Paper 8572. Washington, DC: World Bank. <https://openknowledge.worldbank.org/handle/10986/30324> License: CC BY 3.0 IGO.
- Pritchett, Lant. 2015. "Creating Education Systems Coherent for Learning Outcomes." Working Paper. RISE Programme, Oxford, UK. https://riseprogramme.org/sites/default/files/inline-files/RISE_WP-005_Pritchett_2.pdf.

- Puma, Michael, and Jacqueline Raphael. 2001. "Evaluating Standards-Based Professional Development for Teachers: A Handbook for Practitioners." The Urban Institute, Washington, DC. <http://webarchive.urban.org/Uploadedpdf/410432.pdf>.
- Richardson, Emily. 2014. "Teacher Motivation in Low-Income Contexts: An Actionable Framework for Intervention. Teacher Motivation and Strategies." OAS (Organization of American States), Washington, DC. <https://www.oas.org/cotep/GetAttach.aspx?lang=en&cld=593&aid=876>.
- Rossiter, Jack. 2020. "Link It, Open It, Use It: Changing How Education Data Are Used to Generate Ideas." Note. Center for Global Development, Washington, DC. <https://www.cgdev.org/publication/link-it-open-it-use-it-changing-how-education-data-are-used-generate-ideas>.
- Wilichowski, Tracy, and Anna Popova. 2021. "Structuring Effective 1-1 Support: Technical Guidance Note." Coach Series, World Bank, Washington, DC. License: Creative Commons Attribution CC BY 4.0 IGO.
- World Bank. 2015. "Enhancing Teacher Effectiveness in Bihar Operation." Project Appraisal Document (PAD), India. World Bank, Washington, DC. <http://documents1.worldbank.org/curated/en/184631468000251240/pdf/92972-PAD-P132665-IDA-R2015-0096-1-Box391421B-OUO-9.pdf>.
- World Bank. 2018. "Data Collection and Management for Improved Institutional Development." Education Global Practice (GEAK), World Bank, Washington, DC.
- World Bank. 2019a. "Enhancing Teacher Effectiveness in Bihar (ETEB) Operation." World Bank, Washington, DC. <https://www.worldbank.org/en/country/india/brief/enhancing-teacher-effectiveness-bihar>.
- World Bank. 2019b. Implementation Completion and Results Report for Indonesia. World Bank, Washington, DC. <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/169921570551716879/implementation-completion-and-results-report-icr-document-indonesia-improving-teacher-performance-and-accountability-kiat-guru-p159191>.
- World Bank. 2019c. "Selected Drivers of Education Quality: Pre- and In-Service Teacher Training." World Bank, Washington, DC. https://ieg.worldbankgroup.org/sites/default/files/Data/Evaluation/files/Drivers_of_Education_Quality.pdf.
- World Bank. 2020a. "Developing a Strong DPF Results Framework." World Bank, Washington, DC. <https://olcsb.worldbank.org/content/developing-strong-dpf-results-framework>.
- World Bank. 2020b. "The World Bank's Education Response to Covid-19." World Bank, Washington, DC. <https://pubdocs.worldbank.org/en/487971608326640355/External-WB-EDU-Response-to-COVID-Dec15FINAL.pdf>.
- World Bank. Forthcoming. "Teach in Action: Three Case Studies of Teach Implementation to Date." Education Global Practice (GEAK), World Bank, Washington, DC.
- Yoshikawa, Hirokazu, Diana Leyva, Catherine E. Snow, Ernesto Treviño, M. Clara Barata, Christina Weiland, Celia J. Gomez, and Lorenzo Moreno. 2015. "Experimental Impacts of a Teacher Professional Development Program in Chile on Preschool Classroom Quality and Child Outcomes." *Development Psychology* 51 (3): 309-22. <http://doi.org/10.1037/a0038785>.

Access Coach Tools
and Resources 