ipa Innovations for Poverty Action | PLAN INTERNATIONAL | The LEGO Foundation | VSO | vvob education for development

Play Our Part

Photo credits © Sarah Farhat / World Bank

Class observation:

A Play Our Part CoP guide on using class observation to improve a program

Advancing Learning through Play in Early Childhood Education

Who are we:

The **LEGO Foundation**-funded **Play Our Part (PoP)** initiative is a community of practice made up of three early childhood education (ECE) implementers - **VVOB - Education for Development**, **Plan International**, and **Voluntary Service Overseas (VSO)** - working to strengthen Learning through Play (LtP) instruction and holistic learning in schools and centers. The initiative uses a multi-level approach, engaging schools, government institutions, and communities.

Engage in our Resource Package:

Innovations for Poverty Action (IPA), as PoP's Learning and Design Partner, authored this research package to consolidate implementation findings into credible recommendations and guidance for those implementing—or considering implementing—ECE and LtP programs. PoP implementing partners contributed findings and insights from their programs, with support from the LEGO Foundation. We invite you to explore these resources and join us in bringing the LtP vision to life.

Table of contents

3 What is class observation? Why is it important for teacher capacity building and improving programs?

5 Aligning class observations with learning purpose

7 How can you better design class observation data for program improvement?

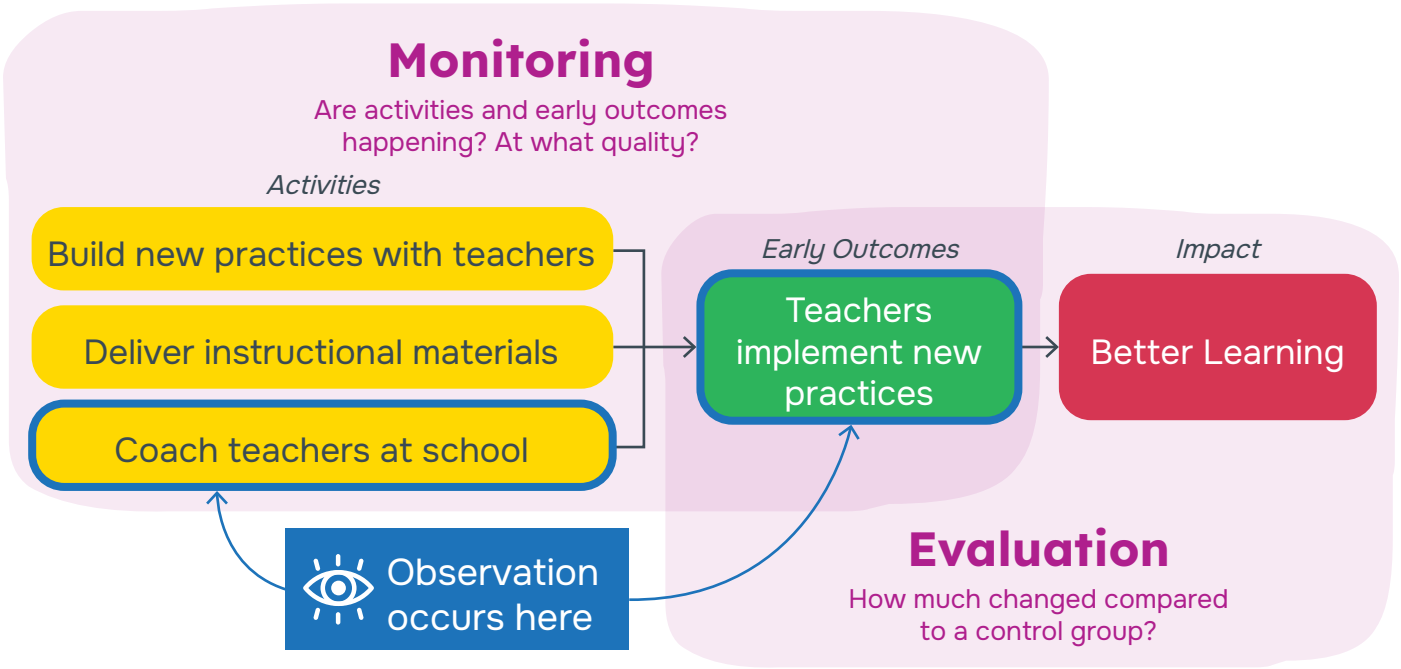
8 Ensure consistent and accurate data quality

10 How can you better use class observation data for your program?

What is class observation? Why is it important for teacher capacity building and improving programs?

Class observation is a tool for viewing a teacher's instructional methods in action by observing them teaching real lessons to real students. They are often used to assess the overall classroom environment, guide improvement of instructional or physical characteristics, and are complemented by an observation guide to focus attention to critical instructional behaviors.

Figure 1. Classroom monitoring is crucial in a Teacher Professional Development (TPD) program, because the pathway to improved student learning requires teachers using new practices at a sufficient quality.



If monitored throughout a program, observations can be used to track key instructional practices: 1) whether they are being applied at all, and 2) at what quality. Answers to these two questions are the most informative guideposts for deciding which program elements need to be adapted in order to better impact learning.

When combined with qualitative data and group discussions, class observation data can guide the strategy for *how* to adapt a program. Analyzing this data with your team can help program design and teacher capacity building by showing:

1. **Why practices are working well:** frequent and high-quality teaching practices can help identify which training methods may be more effective or reveal mechanisms for behavior change that are more likely to influence teachers.

2. **Where support or more buy-in is needed:** Program and teacher accounts can indicate if inconsistently observed practices require improved training approaches, or stronger efforts to demonstrate the value and practicality of new methods in order to increase buy-in.

3. **When to rework program concepts:** New methods may not address teachers' lived realities, such as large classes, little space or materials, or a high need for inclusivity. In such cases, programs should consider significant modifications to what teachers are asked to do.

Class observations also give key information about whether your program is ready to be more rigorously evaluated. Rather than relying on potentially inaccurate self-reports from teachers, class observations help verify whether the minimum instructional standards are being met and provide insights about overall quality. Regular collection will show when key teaching practices essential for learning are being sustained across teachers, suggesting readiness for impact measurement.

Aligning class observations with learning purpose

Class observations can be valuable at various stages of a program's development and should be designed in a way that serves those purposes. Such insights range from identifying early indications that teachers are adopting a practice to whether specific instructional practices are associated with stronger student learning.

Consider class observation purposed to serve three different programmatic purposes:

Observation for teacher improvement	Observation for program improvement	Observation for program evaluation
A *teacher professional development activity* associated with **coaching** in which a coach uses an observation to identify strengths, weaknesses, and tailor support to *build the teacher's skills* (See the yellow box activities in Figure 1). Frequency: based on teacher need and usually ongoing to support professional growth.	A Monitoring, Evaluation, and Learning (MEL) *activity*, **which along with iteration**[1], makes the program more effective. A sample of participating teachers are observed in a consistent way to see if a predetermined set of practices are happening during instructional time, and at what level of quality. Frequency: multiple learning cycles per year to ensure consistent implementation and regular improvement.	In many ways similar to program improvement, but with an added **control group** against which the *amount of change* can be compared, in order to quantify the impact on teaching quality or relation between classroom practices and student learning. Frequency: at start and by end of implementation strategically to determine impact of the program.

If you are conducting class observation for program improvement be sure to:

- Observe a representative sample of your teachers
- Consistently and reliably rate their use of key practices
- Use the data to determine if the desired implementation is working and how to make it even better
- Conduct as many rounds of observation as feasible and affordable, iterating at each cycle (prioritize the speed of turnaround with 'just-enough' data to support evidence-based decision making)

[1] Program iteration is the process of refining a program through repeated cycles of implementation, feedback, and adjustment to enhance its effectiveness.

If you do coaching, avoid collecting monitoring observations at the same time.

It is tempting to combine coaching activities with program improvement because it seems more efficient and cost-effective to do both at once. There are two threats to doing so:

- **Less effective coaching:** The time required to train coaches to meet the monitoring observation protocols and maintain consistency takes away from the time they can spend with teachers. Additionally, sharing with teachers that observations will be used for monitoring may make them less open to learning, feel threatened, or even question their relationship with the coach. Be especially careful where coaches hold public or governmental roles.
- **Poor data quality:** Coaches' desire to individualize ratings or give attention to teacher development could bias the way coaches score teachers.

Ideally, an independent observation team such as the program team or external enumerators would arrange a window of time for observing the class, and visit randomly within that period to **ensure the most authentic and consistent data.**

When ready to evaluate how much you've changed teaching or learning, use a control group.

Class observation for the purpose of program improvement tells you if a practice is happening and at what quality, *not* how much a practice's change can be attributed to your program. Additionally, observation for program improvement can let you know when you are ready to measure impact.

However, in order to attribute the changes you observe to your program, **a control group should be established from the outset** against which to compare program classroom observations. A true impact evaluation with a control is a worthwhile investment once classroom observation data demonstrates that the teacher professional development approaches intended to be influenced are consistently reflected in the observations of teachers.

Example: Why is a control group necessary to attribute classroom changes to a program?

The Twigire Mumikino Rwanda (TMR) program aimed to boost playful instruction and learning in Rwanda. Strong improvements from the beginning of the year were measured in the way teachers built real-life connections into play activities. However, when a control group was compared to play-innovative classes, it was clear that both groups of teachers had **improved equally**. Without a more rigorous research design, TMR would have likely overlooked the possibility that there was another explanation besides the effect of their program, such as teachers having better knowledge of the student's home-life at the end of the school year.

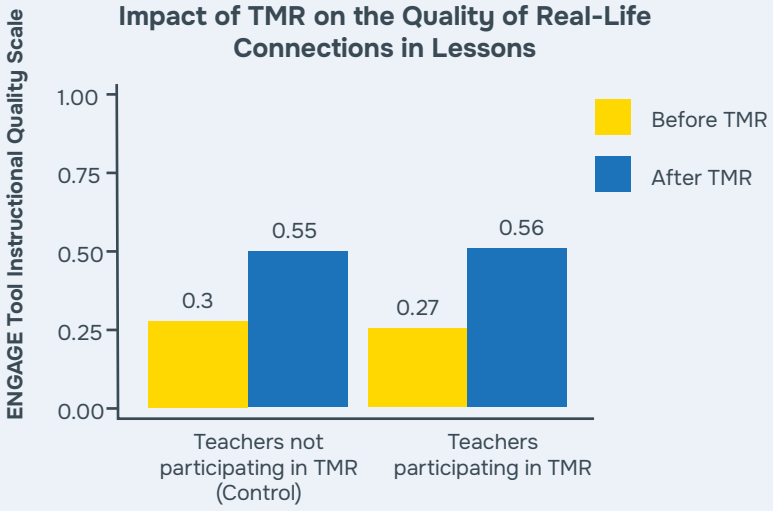


Figure 2. TMR data illustrates why a control comparison is important for measuring program effect.

How can you better design class observation data for program improvement?

Make sure the observation tool captures teaching practices and environmental factors that are key to the way your program functions

Your ability to adapt your pedagogical approach depends heavily on **whether the practices measured are within your control.** Reflect with education technical specialists and the program team on the most important practices that are being asked of teachers in their classrooms, using your teacher professional development program content as a guide for what you can reasonably expect teachers to do as part of your program. Make sure your team can describe how they might be consistently defined, observed during instruction, and measured. User friendly and concise observation tools will help your team consistently rate teachers.

Example: Using the ENGAGE Tool for Effective Teacher Monitoring

Plan International's *Playful Futures* team in Uganda chose the ENGAGE tool (previously named the PLAY tool) for monitoring whether teachers were applying behaviors from training engagements. The technical team mapped the "5 Tips for Play" from its teacher guidebook to specific ENGAGE tool indicators, narrowing it down to 6 core practices. They constructed and added 2 additional practices to measure equal gender responsive and inclusive teaching. Then the team reviewed each practice definition, simplified the criteria for how it could be consistently counted, and modified the example descriptions to meet what would be expected in a typical Ugandan pre-primary classroom. Using the tool, Playful Futures conducted multiple rounds of observations across all participating pre-primary centers, acquiring an accurate gauge of the relative frequency of practices, and those which seemed to be improving (although a counterfactual was needed to ascertain whether change was attributable to the program).

We recommend seeking out observation tools that have already been developed and validated in similar contexts to your own. Please consider these tools which were used by PoP programs:

- ENGAGE tool (formerly known as PLAY)
- MELE class observation tool
- TIPPS

Include information for collection which cannot be easily viewed during an observation of an instructional lesson. For example, it may be useful for the observation team to look at the teacher's lesson plan and rate a few simple aspects. Ignore unrelated elements within the classroom that your program has no intention of changing; there will be little purpose in collecting such data, saving you time and effort.

Finally, be sure to gather environmental data that highlights barriers to desired changes. For example, PoP programs have identified the following as major constraints to teachers' ability to implement play:

- Large class sizes or mixed aged groups and abilities
- Not having enough play materials for activities
- Teacher perception that parents or the head teacher do not support playful learning

Including these environmental factors in the class observation tool will allow you to investigate to what extent those factors constrain teachers (or not) to implement the key classroom practices.

Ensure consistent and accurate data quality

The intention of monitoring implementation is not to respond directly to individual teachers' needs, but to give an honest picture of what's happening in the field so you can adjust teacher support accordingly. This implies observations should aim to rate all teachers in the same way and maintain clear criteria for each rating.

The first way you can control this is by **selecting and training a strong observation team**:

- Train more observers than needed, and choose those who show the best accuracy.
- Select observers with background in the education field (but be aware that this doesn't ensure they are strong raters, nor that they agree on key concepts).
- At training, prioritize exposing your team to the diverse ways an instructional practice can be used, making use of real classroom examples (including video) and discussions on applying the tool. After watching the video, facilitate conversations on correct ratings and the reasoning behind them.
- Keep contact with your team during data collection and conduct ongoing calibration of how to use the tool; they will undoubtedly encounter novel examples in the field, and these can be used to further train the team and develop consistency.
- Conducting class observations will produce a highly skilled observation team, so consider selecting those that may have extended roles working in the field beyond the observation itself.

Example: Improving Observation Consistency Through Training and Practice

IPA trained a team of Rwandan observers with a wide range of exposure to education and teaching. Despite this diversity, the team achieved a high level of consistency by the end of the observation. The team reported the greatest value in watching video footage from real classes, being able to hear both peers' and experts' reasoning about how to categorize different class behaviors, and participation in a range of real play activities (which were then discussed as they relate to the tool's indicators).

Set protocols that maintain a normal classroom setting and maximize consistency

- Ask teachers to prepare as they normally would, and clearly explain that the data will not be used to evaluate their individual performance.
- Consider providing teachers with a general timeframe for the visit but arriving on a random day during that period at the usual lesson time. This helps prevent teachers from preparing a lesson solely for the observation.
- Avoid interactions between observers, teachers, and children during the observation.
- Standardize where the observer sits, their movement (we recommend none), and what they look at during the observation.
- If resources allow, have two observers rate the same lesson independently, and compare their ratings afterwards, discussing where differences are found in scores.

Photo credits © Loïc Watine/IPA, J-PAL

Observe only as many teachers as you can while still being able to analyze and use the data in a timely manner (eg. quarterly)

Trying to observe too many teachers may reduce your data quality, leave fewer resources to do teacher training, and increase the length of each learning cycle. For this reason, it is usually most appropriate to *sample only a portion of your teachers*. While the quantity should depend on your program size, MEL capacity, and geographic dispersion, consider the following guides:

- Estimate how many teachers need to be viewed to give a representative picture of the program as a whole, so that adding 1 more teacher is unlikely to give new information or hold a big surprise; this may depend on whether there are important groups within your teaching population –each of which needs to be observed
- Aim for completing multiple rounds of observation in the calendar year (quarterly, triannually, biannually, etc.)
- Consider how many observers you can train and how long it will take them to travel among the sampled teachers; stretching this team will result in worse data
- Factor in how long it will take the MEL team to collate, clean, analyze data

How can you better use class observation data for your program?

Have an active plan for how the data will be used.

Data on the shelf is a waste of resources and an irresponsible use of time for you and your teachers. Planning in advance is the easiest way to ensure that you will use class observation data.

Figure 3: Consider planning your class observations based on how you will complete a full learning cycle: design, collect, analyze, and iterate.

1. Design observation tools

Prioritize actionable data from classroom observations to inform program improvements and iteration

Key tip: Set Learning Ojectives
Target key questions that which can be credibly answered by classroom observations and will inform future actions. For Example:

1. How many teachers were observed doing a desired classroom practice?
2. What average level of quality was observed across all classrooms?
3. What combinations of practices tended to occur together?

2. Collect

Collect observation data regularly, which factoring in MEL capacity and geographic reach to ensure cost personel feasibility.

Key tip: Plan ahead and be accountable
Set clear expectations with your team for *when* and *how* data will be collected, submitted, and processed. Be accountable to that schedule.

3. Analyze and Share

Summarize findings clearly and make them accessible to enhance team engagement and support data-informed decision-making.

Key tip: Engage your team
Schedule dedicated time with a range of stakeholders such as program leads, education technical specialists, and field coordinators. Analyze the data, unpack its implications, and identify key actions for enhancing teacher professional development (TPD) or informing other program design considerations.

4. Iterate

Refine TDP approaches to address key gaps and align data collection with evolving needs.

Key tip: Adapt your program
Communicate an action plan to your team. Focus on improving controllable elements within your program, such as how you do teacher professional development (TDP) or support instructional practices. Update your learning priorities and restart the cycle.

Incorporate different data points and experiences to make meaning of the data

A combination of different kinds of data can be used to supplement your class observations and lead to better informed decisions. Consider collecting qualitative data that describes teachers' experiences while participating in your program, challenges, solutions, and reactions to the new practices and have a plan for how positive or negative findings may lead to different iterations:

Positive signs in the data

- Consistent demonstration of a teaching practice across a wide selection of teachers
- Specific teachers demonstrating high quality delivery of certain practices.

Implications: Something seems to be right. Use qualitative data to...

- Identify which 'mechanisms' or enablers are leading to success
- Explore how 'mechanisms' for success be transferred to strengthen other practices or teachers

Red flags in the data

- Practices with consistently low scores
- Where teachers express:
  - Difficulty understanding key concepts
  - Infeasibility of a practice
  - Resistance to doing a practice

Implications: Change is needed. Use qualitative data to...

- Identify the barriers to skill improvements or adoption of practices.
- Look for solutions that supplement current programmatic strategies

Example: Making data accessible for collaborative decision making

Plan International's *Playful Futures* program collected a combination of class observations, teacher surveys, mentor feedback, and parent engagement data in a Power BI dashboard.

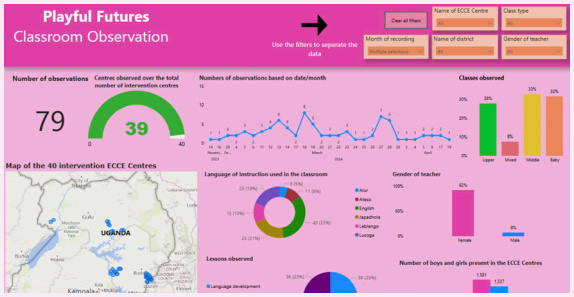


Figure 4: Playful Future's visually engaging dashboard.

The MEL team ensured quantifiable data was presented in easy-to-understand dashboards and arranged next to related qualitative comments listed in scroll down boxes. For example the frequency of 9 key teaching practices were displayed in simple bar charts with rankings and comments from mentors about practices showing the most improvement. Access was made available to team members, who were prompted to play with the findings during team meetings and discuss how the program was working. Playful Futures says this changed its internal culture by leading to more access to real-time data, collaboration across the team and data-informed decision-making.

Display the data for your team, making sure to invite both decision makers and those delivering the project. Engage the group to comment on how their experiences on the project **confirm or contradict** the data, **explain** the results, and **propose adaptations** to controllable program elements.

The data can also be used to engage external stakeholders to communicate how the project is going. While this is often meant as a means of accountability to your stakeholders, you can also use such opportunities to engage stakeholders in the same questions asked internally, giving them a voice to help decision making around program elements that affect them.

ipa Innovations for Poverty Action

PLAN INTERNATIONAL

The LEGO Foundation

VSO

vvob education for development