LEBANON - Coaching teachers on a new student-centered teaching approach:

Results of an impact evaluation

collaboration of the World Bank and
the Ministry of Education and Higher Education (MEHE)
Department of Scholastic and Pedagogical Guidance (DOPS)

March, 2024

Table of Contents

Acronyms

CERD	Center for Educational Research and Development
CLASS	Classroom Assessment Scoring System
DGE	General Directorate of Education *(french: Direction Générale de l 'Education)*
DFID	Department for International Development *(changed to FCDO)*
DOPS	Department of Scholastic and Pedagogical Guidance
FCDO	Foreign, Commonwealth and Development Office (formerly DFID)
MEHE	Ministry of Education and Higher Education
PRIMR	Primary Math and Reading program
RACE	Reaching All Children with Education
USAID	United States Agency for International Development

Acknowledgments

This work was a collaboration of the Ministry of Education and Higher Education (MEHE) Directorate of Guidance and Counseling (DOPS), the World Bank Middle East and North Africa Education Team, and New York University's Global TIES for Children Institute. The literature review and teaching model were developed jointly by the three parties, while the intervention and evaluation were designed and executed by the World and MEHE.

The **MEHE** team was led by Hilda Khoury, Director of the Office of Guidance and Counseling at MEHE with guidance from Fadi Yarak, Director General of Education at MEHE. MEHE's core project team of coaches contributed to the co-construction of the model, implementation, and evaluation support.

The **World Bank** team was led by Ms. Kaliope Azzi-Huck, Sr Operations Specialist; Mr. Alonso Sanchez, Senior Economist; and Mr. Mohammed Yassine, Education Specialist, and was comprised of Marwan Alawieh, Nadine El Franji, Wael Moussa, Cynthia Yammine, and Aarya Shinde.

The **New York University** team was led by Lindsay Brown, Senior Research Scientist with the Global TIES for Children Institute and was comprised of Ms. Roxane Caires, Laura Killips, Douha Jaballah, and Ha Yeon Kim.

This report was co-authored by Kaliope Azzi-Huck, Wael Moussa, Nadine Franji, and Aarya Shinde. Technical and editorial inputs were received from Wendkuni Ouadraogo and Emma Etori. Inputs were also received from Dr. Hilda Khouri and the MEHE Core Project Team.

Technical and financial support were received from the following World Bank -administered Trutfunds:

REACH Trust fund;
SIEF Trust fund;
Forced Displacement Trust fund

Data collection at baseline and endline was managed by InfoPro Research Sal.

An appreciation is extended to the coaches, teachers, and principals who participated in the pilot phase of the coaching model. Their dedication, enthusiasm, and willingness to embrace change have been crucial in the successful implementation of the pilot and the generation of valuable data for the evaluation.

I. Executive Summary

i. This report summarizes the effects of introducing a new teaching method to teachers of key subjects at elementary and middle school levels using an on-site, recurrent coaching approach. Specifically, it documents changes in teacher behaviors as well as teacher and coach perspectives after the model was implemented, with the aim of providing policy recommendations for further scale up where relevant.

ii. In 2021, Lebanon completed the pilot of a new coaching model that introduced a student-centered approach to teaching among teachers of grades 4 (elementary level) and grade 8 (middle school level). The design in this case focused on co-construction of a new coaching model to ensure adaptability to local context, a drive for on-site and ongoing support, and introduction of new elements in gradual manner that allowed for real-time practice, observation, and feedback. The model was selected by adapting similar prior experiences of Brazil, Kenya, Peru, South Africa, and the United States. The model focused on on-site and ongoing support to teachers using pedagogical coaches who visited each targeted teacher 6 times during the intervention.

iii. The evaluation study used a randomized controlled trial approach, with coaches randomly assigned to either a treatment group that received training on the new coaching model or a control group that continued with business-as-usual practices.

iv. The results showed that the coaching model had a positive impact on teacher practices and outcomes, with teachers in the treatment group reporting higher levels of self-efficacy, job satisfaction, and perceived usefulness of coaching visits compared to the control group. The coaching model also led to improvements in classroom practices, with teachers in the treatment group demonstrating greater use of student-centered teaching approaches, student engagement, and group work. The coaching model emphasis on deliberate practice, reflection, and feedback, helped teachers develop new skills and improve their instructional practices. Teachers who received coaching reported increased satisfaction with their coaching experience, perceived improvements in their teaching practices, and positive impacts on student engagement. Additionally, the analysis indicates a significant impact of the coaching model on promoting whole-group discussions in lessons.

v. However, the coaching model faced challenges in implementation, including resistance from some teachers, disruptions due to the COVID-19 pandemic and political unrest, and limited time for teacher-coach interactions. The prolonged school closures and delays in implementation prevent the evaluation of student learning outcomes as those students who were assessed at baseline had moved on before the project reached endline stage.

vi. Overall, the results suggest that a well-designed coaching model can support teachers in enhancing their performance and improving student outcomes. The effect sizes in various dimensions of classroom practice further support the notion that the coaching model is effectively

translating into tangible changes in teaching approaches and classroom dynamics, as intended by the theoretical framework underlying the intervention.

vii. The report recommends a shared vision and unified framework for coaching, coordination of coaching and professional development initiatives, increased time for teacher-coach interactions, and support for teachers in inclusive classrooms and schools. Lebanon's coaching pilot has demonstrated effectiveness in improving teaching practices and student learning outcomes. Based on these findings, key recommendations include a shift towards a minimum number of coaching visits per coach to teachers, along with a support framework. This would ensure regular and consistent support for teachers and help improve teaching practices and student outcomes. The findings from these coaching programs can inform future decision-making and planning for coaching programs in education systems worldwide.

II. Background

1. The education system in Lebanon covers a combination of public and private schools, overseen by the Ministry of Education and Higher Education. It is divided into three levels: primary, intermediate (middle school), and secondary education. Primary education is compulsory and lasts for six years, while intermediate education lasts for three years. Secondary education is divided into two cycles, with the first cycle lasting for three years and the second cycle lasting for two years. At the end of secondary education, students take the Lebanese Baccalaureate exam to determine their eligibility and track for higher education. A technical-vocational track is also offered in private and public schools and runs parallel to secondary education.

2. Despite grave achievements and a well-earned reputation for rigor, Lebanon's education system faces several challenges. There are disparities between public and private schools, with private schools generally offering higher quality education but being expensive and inaccessible to all students. The system also grapples with overcrowded classrooms, limited resources, and outdated curriculum. Exogenous factors including political instability, economic crises, and COVID-19 pandemic closures have further strained delivery and accessibility of education services. Efforts are being made to address these challenges, including reforms and initiatives to improve access, quality, and teacher training.

3. The 1926 Lebanon constitution set the right to education for everyone regardless of their culture and religion. Lebanon also committed itself to providing free education at the Universal Declaration of Human Rights adopted in 1948. This right was further reiterated in the 1989 Taif agreement after the civil war which also provided for reforms and the strengthening of the education sector. Since then, further reforms have ensued including the 1994 Education Recovery Plan, the 1997 curriculum reform, the 2003 Education for All Plan of Action (EFA), and the 2010 Education Sector Development Plan (ESDP). Addressing historical challenges like the Syrian refugee influx, the Reaching All Children with Education (RACE) comprehensive sector program was launched in 2014 and elaborated on in 2016. Within that program, the government of Lebanon highlighted the priority of improving teaching and support to teachers across the country's public schools.

Table 1: Chronology of educational reforms in Lebanon

DATES	REFORM ATTEMPTS	HISTORICAL MILESTONE
1926	The Constitution	French Mandate
1943–75	Post-independence reforms establishing the DGE[1] and the CERD[2]	Post-independence
1989	Provision of the Taif Agreement	
1994	Plan for Education Reform in Lebanon	Post–civil war reconstruction efforts
1997	Curriculum reform	

[1] DGE= Direction generale de l'education

[2] CERD = Center for Education Research and Development

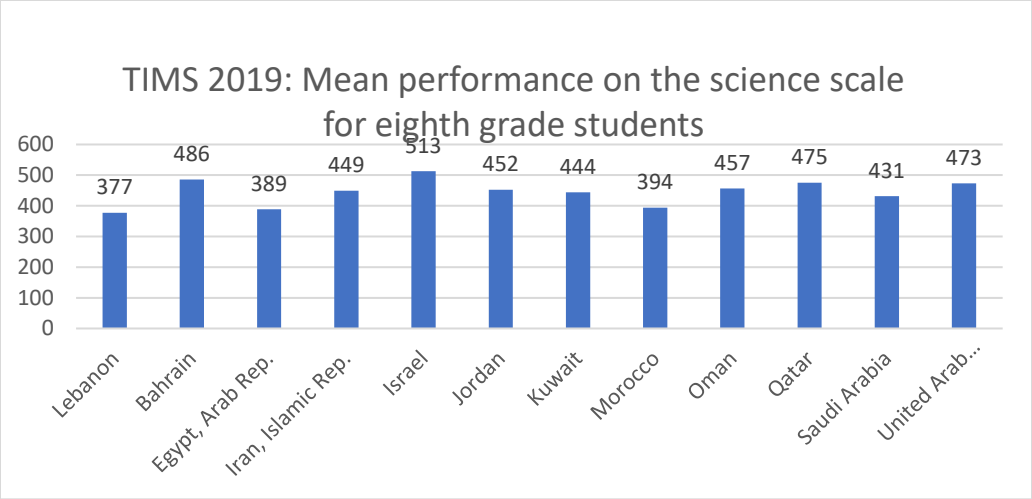
2003	Education for All (EFA) Plan of Action	Contemporary sector development reforms
2010	Education Sector Development Plan (ESDP)	
2014	Reaching all Children with Education (RACE)	Response to the influx of Syrians into Lebanon

Source: Lebanon Political Economy Study (World Bank Group)

4. Over recent decades, significant efforts have advanced access and quality of education in Lebanon, despite social and political challenges. The country thus boasts of historical rigor in education excellence, with a large proportion of its population having advanced degrees from domestic and international institutions. However, the success of that small segment does not mirror nationally aggregated results, as highlighted by Lebanon's *Human Capital Index,* which stood at 0.54 in 2022, i.e., a person born today will be only 54 percent as productive as they could have been if they were to receive complete education and total health. Similarly, by World Bank estimates, the Learning Adjusted Years of Schooling (LAYS) is 6.3, meaning that twelve years of schooling is estimated to produce 6.3 years of actual learning. In the latest international assessments (Program for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMMS)) Lebanese students performed below OECD and peer countries in the Middle East and North Africa (MENA) region (Figure 1)[3].

Figure 1: PISA 2018 mean score in reading and TIMS 2019 mean score in science

[3] Schools, Students, and Teachers in Lebanon: Research for Results Program 2021. Washington, D.C.: World Bank Group



5. Lebanon's education has been driven mainly by multi-dimensional political, social, and economic challenges with a compounding effect. Political instability intensified in 2019 and was followed by the collapse of the Lebanese economy and financial markets, and subsequently followed by the COVID-19 pandemic, which saw over 1.3 million students kept out of schools for lengthy periods. Then the Beirut Port blast of August 2020 and the infrastructural and systemic damages crippled a society already suffering from decades-long domestic socio-economic fragility[4]. These circumstances have put social services like education in a constant state of emergency or mitigation-of-losses mode. In addition to exogenous issues, endogenous factors have also been significant and include an overstretched education system that has increasingly struggled to maintain quality, relevance, and effectiveness (as measured by LAYS) and a disconnect between education and labor market demands. It is also worth noting that the 2016 SAHWA Youth survey revealed that 92% of youth felt that their education does not match the needs of the labor market.

6. **An Education Strategy:** In 2021, the Ministry of Education and Higher Education formulated Lebanon's Education Reform Plan to enhance the accessibility and quality of education in the country. This comprehensive strategy identified seven key pillars that needed to be addressed. Figure 2 provides the highlights of each pillar- strategy and curriculum, programs, governance, accountability, learning assessments, teachers, financing, and schools (including infrastructure and administration).

7. Out of these pillars, the Government recognized the significance of supporting teachers who at the time numbered about thirty-six thousand as a teaching workforce. Thus, MEHE identified them as a priority (*Figure 2*), developing with specific focus on school-based professional development, coaching, and structured pedagogy. This emphasis on teachers aimed to provide them with the necessary support and resources to enhance their professional growth and improve the overall quality of education. This approach was aligned with the growing evidence about the need to focus on the teachers as a critical input to better learning outputs: e.g. the Global Education Evidence Advisor Panel's 2023 Report on "Cost-effective Approaches to Improve Global

[4] Lebanon- Almost Three-quarters of the Population Living in Poverty (ESCWA, September 2021)

Learning,"[5] supporting teachers with structured pedagogy was cited as among the most valuable investments and thus a "great buy" that is backed by robust evidence. The report noted that one effective approach to support teachers with structured pedagogy is through skills-based ongoing teacher training, coupled with teacher mentoring. These two components, when carefully coordinated, can mutually reinforce each other, resulting in improved teaching practices and ultimately enhancing global learning outcomes. Teachers are the most important in-school factor contributing to student academic and social-emotional learning[i]. Going from a low-performing teacher to a high-performing one increases student learning by multiple years over business-as-usual schooling.[ii] Effective teachers also have a substantial impact on the long-term learning outcomes as well as the general well-being of students.

Figure 2- Lebanon Education Reform Plan-
Government Key Priorities for Education Sector Reform (2021)

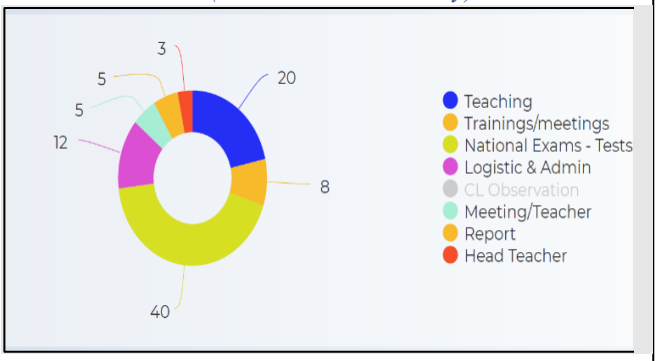
8. **Supporting teachers in the Lebanese Context:** Traditionally, the role of a coach in the education organizational structure is to support and advise on teaching. Thus, instructional coaches "work with teachers to help them incorporate research-based instructional practices into their teaching so that students will learn more effectively" (Knight, 2017). They do this by addressing instruction, classroom management, planning, and reflection. But not all education systems have personnel dedicated to coaching. Some administrations designate this responsibility to principals, supervisors, team leads, or others.

[5] 2023 Cost-effective Approaches to Improve Global Learning - What does Recent Evidence Tell Us are "Smart Buys" for Improving Learning in Low- and Middle-income Countries? (English). Washington, D.C.: World Bank Group.

9. In the Lebanese public education system, two primary entities are responsible for teacher professional development: The Department of Scholastic and Pedagogical Guidance (DOPS) within the Ministry of Education and Higher Education (MEHE) and the Center for Educational Research and Development (CERD).

10. The responsibilities of coaching in Lebanon fall within the work program of DOPS, but not surprisingly, they are not exclusive, as coaches in Lebanon are also themselves teachers, administrators, and others. During National Examinations, coaches are called upon to proctor and grade. As part of the baseline data collected, information was gathered from coaches regarding their primary and secondary roles. Figure 3 below shows the various duties taken on by a coach in Lebanon which shows preparing for and proctoring national examinations as their most consuming task.

Figure 3: Percentage of time a coach spends on various activities (2019 baseline survey)



Source: Lebanon Coaching Model Baseline Survey Report - 2021

Who are DOPS

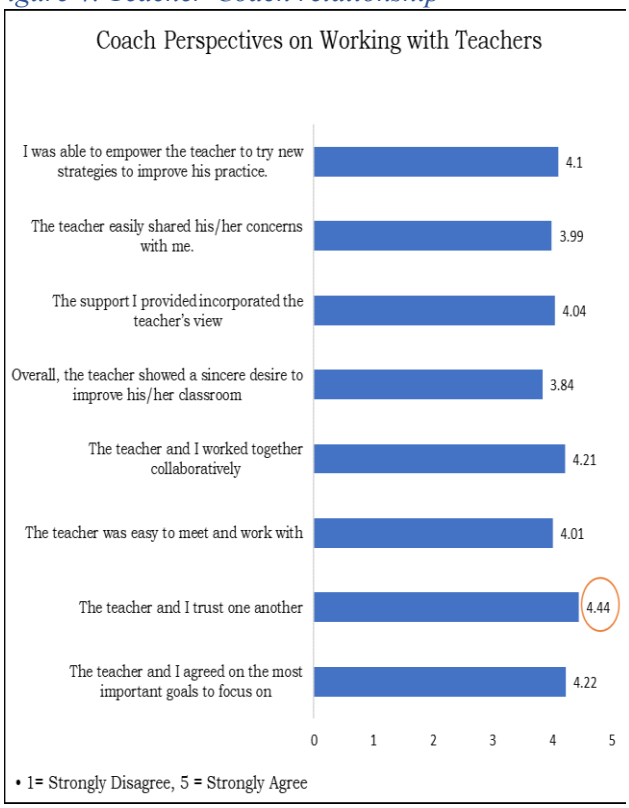
DOPS refers to the Department of Scholastic and Pedagogical Guidance in the Ministry of Education and Higher Education in Lebanon. The team led by a Minister-appointed director, is tasked with providing follow-up, counselling, modelling, and mentoring to teachers. They provide tailored coaching to Lebanese public-school classrooms teachers on pedagogy, health and psycho-social issues, and child protection. Coaching process is managed and monitored through its in-house developed coaching system.
Educational meetings, that allow subjects coordinators and relevant coaches and units to meet on regular basis, aim towards:

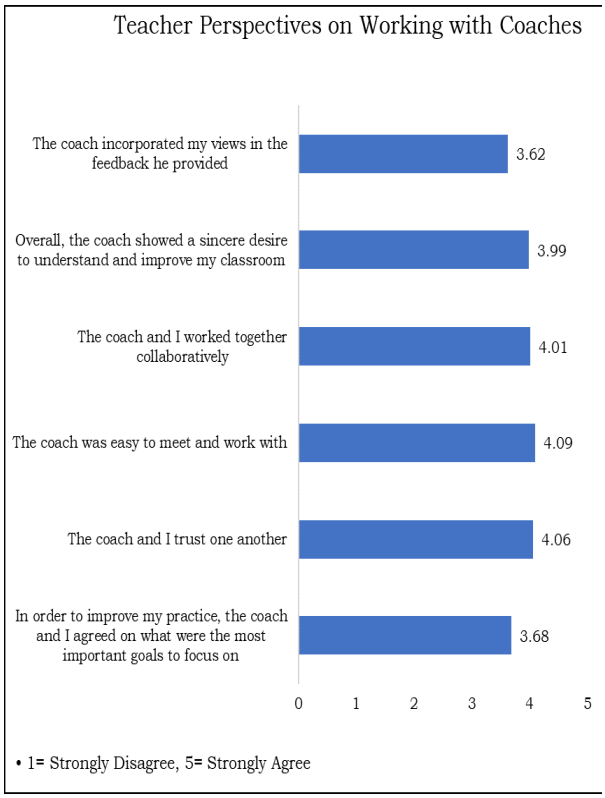
- analyzing teachers' practices and providing advice/guidance to enhance teachers' performance.
- correcting erroneous educational practices that are usually observed during the coaching visits.
- supporting teachers in overcoming challenges that they might be facing.
- reviewing examination and assessment tools that are being used to evaluate student's performance.

supporting teachers in reflecting on their professional practices.

11. The information also showed that the responsibilities expected of a coach meant that their interactions with teachers, while expected to be regular and structured, are inconsistent and often deprioritized due to other commitments. Coaches and teachers were asked to articulate their views of contribution to their work program. Figure 4 summarizes the responses, highlighting significant disconnects related to each party's value-add to the practice. For example, 86% of coaches said they thought they were incorporating teachers' views when giving feedback, whereas only 68% of teachers felt that their views had been considered. This indicates that teachers may sometimes disagree with coaches about important areas to focus on and that the former may feel they are not as heard by the latter as they would like. It is, therefore, vital to ensure both parties agree on the objective and value of coach-teacher interactions.

Figure 4: Teacher-Coach relationship





12. Key differences between coaches and teachers' perceptions regarding goal agreement and incorporating teachers' views into coaches' feedback shown a lack of trust or understanding among teachers of the significance of their (coaches') contribution. Conversely, it can be construed that teachers may have felt that they did not benefit significantly from the coaching process.

III. Coaching Model Design

Supporting Student-Centered Teaching

13. The **Lebanon Education Reform Plan**, ***actions for Building Forward Better (2021)*** identified several factors contributing to low returns in Lebanon's decreasing learning outcomes. These included poor learning environments, quality of teaching, rigid systems, and outdated curriculum[6]. The Reform Plan, spanning five years, highlights seven (7) key education priorities, seen as requirements to improve the quality of education. Prominent among these priorities is Teachers, specifically improving the use of the existing teaching workforce.

14. The Government of Lebanon recognizes the importance of school-based professional development, coaching, and structured pedagogy as strategies for improving teaching practices. In alignment with this priority, the Department of DOPS (Directorate of Professional Development

[6] Foundation for Building Forward Better: An Education Reform Path for Lebanon

and Educational Research) at the Ministry of Education and Higher Education (MEHE) has developed a comprehensive coaching system for teachers in Lebanon. This initiative aims to shift teachers' instructional practices away from traditional rote learning methods towards more deliberative approaches that foster engaging and effective learning experiences, specifically:

Moving from….	To……
TRADITIONAL TEACHING PRACTICES • Generalized • Unfocused and Repetitive • Unconstructed ⇨ Rote learning and slower growth	**DELIBERATIVE TEACHING PRACTICES** Targeted Focused and purposeful Systematic ⇨ Engaged learning and faster Growth

Co-Construction of a New Model of Teaching and Coaching

15. Executing the priorities set forth in the Support to Reaching All Children with Education (S2R2) program (financed by the World Bank and other partners), the Government of Lebanon (GoL) aimed to enhance the quality of education services (pillar 2) and strengthen the capacity of the education system to deliver high-quality education (pillar 3). As part of this effort, the GoL committed to implementing a new coaching model that aligns with the Teacher Competency Framework adopted by the Ministry of Education and Higher Education (MEHE) in 2019. The focus of the coaching model is on the use of teaching tools that emphasize Facilitation moves and Deliberate Practice.

16. In collaboration with the World Bank and New York University, the Ministry of Education and Higher Education established a teaching model prioritizing student-centered learning. This model was developed after conducting a comprehensive literature review of coaching models implemented in various regions, including the Americas, South Asia, and East Asia. The collaborative efforts of MEHE/DOPS, NYU, and the World Bank resulted in identifying a set of high-leverage competencies. The group also worked on defining "core practices" and applied these concepts to the Lebanese Teacher Competency Framework (refer to Figure 5) to ensure a targeted focus on areas where teachers can benefit the most from support provided by the Directorate of Professional Development and Educational Research (DOPS).

17. Research studies have consistently shown that teachers' instruction can be positively influenced by professional development focused on core practices (Cohen et al., 2016). Additionally, studies have indicated that teachers who effectively implement core practices in their teaching demonstrate improved student learning outcomes as measured by rigorous assessments (Grossman et al., 2014).

18. Research indicates that teachers can enhance their skills through continuous support and training. Successful teacher professional development programs exhibit certain shared characteristics, such as face-to-face interaction, subject-specific focus, professional incentives, teacher collaboration, and follow-up visits. These program elements significantly enhance

effectiveness, facilitating ongoing growth and improvement in teaching practices (Béteille & Evans, 2021). Building upon this research, a workshop conducted in 2019 by the Directorate of Professional Development and Educational Research (DOPS) and its partners identified three specific focal competencies to prioritize in the development of a coaching model:

Figure 5: Focal competencies and affiliated elements from MEHE's Teacher Competency Framework

T.SPP.METH: Competency	T.SPP.METH1 - Competency Components	T.SPP.METH1 Indicators* *These indicators are given as examples, but the list is not restrictive.
Employs a variety of teaching and learning methods based on the theories and concepts of learning and teaching appropriate to the learning styles of the various learners and on his/her training acquisitions	Links between teaching and learning methods and basic pedagogical theories, taking into account the type of the educational material and the characteristics of the learners' age group	a. Uses diverse strategies, situations and techniques that inspire learners' motivation and suit their interest b. Encourages communication among learners c. Applies reinforcement in all its forms d. Uses problem-situations that urge learners to mainstream their knowledge into similar or new situations e. Uses activities ranging from recalling, understanding, applying, analyzing summarizing, evaluating, criticizing and innovating f. Provides learners with the appropriate means and documents for the activity g. Assigns homework in quantity and quality to suit the learners' age and abilities

Source: The Competency Frameworks - Supporting Quality Teaching in Lebanon

Note: *T(Profession) = Teacher, SPP(Domain) = Specialized Professional Practice, T.SPP.METH (Competency) = Employs a variety of teaching and learning methods based on the theories and concepts of learning and teaching appropriate to the learning styles of the various learners and on his/her training acquisitions*

19. Several critical challenges were identified at the institutional level that needed to be addressed to successfully implement the new coaching model and any changes to the coaching practice in Lebanon. These challenges encompassed areas that required reform and improvement. Some of the key challenges identified include:

Challenge or Area for Improvement	Recommendation
• No shared vision of what coaches should accomplish with teachers or how	⇨ Align and unify the coaching system around a common framework.
• Uncoordinated coaching and professional development initiatives	⇨ Focus and align coaches and teachers on cross-curricular core practices.
• Insufficient time spent in teacher-coach interactions	⇨ Recommendation #5: Increase

20. The above recommendations were further developed into specific components or focus areas of the coaching package. To ensure feasibility, alignment, and contextual fit, the World Bank, NYU-TIES, and MEHE's DOPS co-constructed specific activities and processes within each recommendation.

21. Furthermore, by drawing upon current research, classroom experience, existing observation protocols, and Lebanese classroom videos, the team created a rubric for each focal competency, showcasing the range of quality levels associated with each competency. This spectrum was devised to assess the degree to which an instructor incorporates the methodology and subsequently transformed it into a classroom observation tool. Figure 6 below presents the classroom observation rubric that was developed to evaluate teachers' adoption of the new approach to teaching. It focuses on Instructional Dialogue.

Model Modifications – Adapting for in-person and virtual teacher support

Due to the challenges posed by the COVID-19 pandemic and school closures, the initial coaching model, consisting of 14 elements, became unfeasible to implement. With prolonged closures and uncertainties surrounding the return to in-person learning, the team had to adapt the model to accommodate both in-person and virtual settings. Consequently, a simplified model was developed, focusing on the elements that could be effectively transferred across both modalities. The modified coaching model comprised the following elements:

a. Instructional Dialogue (teaching through exchange and reflection)

22. Instructional dialogue is a vital component of student-centered teaching, as it facilitates active learning and promotes deep understanding. Through meaningful conversations and exchanges, instructional dialogue fosters collaborative problem-solving, critical thinking, and knowledge construction. It empowers students to express their thoughts, reflect on their understanding, and participate in peer-to-peer discussions. Learners experience enhanced comprehension and long-term learning outcomes by actively constructing their knowledge.

23. The levels of engagement in instructional dialogue can vary, ranging from minimal interactions or absence of dialogue to fully engaged, conversational situations where information is delivered.

The DOPs team has developed a classroom observation tool that captures and assesses these levels of instructional dialogue. The figure below illustrates the various levels of instructional dialogue as measured by the observation tool.

Figure 6: Classroom Observation Tool (Instructional Dialogue)

	1	2	3	4
Level of Questioning	Students are never or rarely asked higher-order questions that are thought-provoking, require student reasoning, use of his/his own ideas, etc.	Students are occasionally asked higher-order questions that are thought-provoking, require student reasoning, use of his/her own ideas, etc.	Students are frequently asked higher-order questions that are thought-provoking, require student reasoning, use of his/her own ideas, etc.	Students are consistently asked higher-order questions that are thought-provoking, require student reasoning, use of his/her own ideas, etc.
Response to Student Answers	Teacher and/or students rarely takes students' response into consideration; answers may be automatic ("okay", "good")	Teacher and/or students responds briefly to students' answers; s/he may re-voice students' answers verbatim or in their own words	Teachers and/or students sometimes pick up on the students' responses; students have the opportunity to explain, clarify, criticize, and/or evaluate the ideas of the content	Teacher and/or students frequently pick up on the student responses; students have the opportunity to explain, clarify, criticize, and/or evaluate the ideas of the content
Student-to-Student Dialogue	Teacher does not provide opportunities for students to engage in dialogue with each other	Teacher rarely provides opportunities for students to engage in dialogue with each other	Teacher sometimes provides opportunities for students to engage in dialogue with each other	Teacher frequently gives students the opportunity to engage in dialogue with each other

b. Facilitation Moves (to encourage greater student uptake)

24. Facilitation Moves for Student Uptake consists of three key steps: Ask, Press, and Revoice. These steps serve as guiding principles for teachers to facilitate student engagement and participation in the learning process effectively.

Ask

Teacher or student launches or redirects discussion with a generative question allowing for students to develop and share ideas and reasoning about content, concepts, and material under study. Some questions are planned by both teacher and student, and some are spontaneous.

Questions should be authentic and will vary based on text at hand

In general, Generative Questions:

- Require students to use their own ideas to answer

- Are open-ended questions
- Require some level of inference or analysis on the part of the student (e.g., not recall questions where students can easily find the answer in a text/figure, etc.)

Press

Teacher or student responds to a comment in order to (1) press for elaboration or clarification or (2) press for evidence (substantiation)

Examples include:

- **"Say more about what you mean by ___."**
- **"Can you clarify what you mean by ___?"**
- **"What evidence do you have for ___?"**
- **"What in the text makes you think ___?"**

Revoice

Teacher or student strategically restates another person's claim (and then checks in with the speaker to confirm) in order to:

3. Verify/check their understanding of someone's claim
4. Summarize what has been said so far
5. Introduce or encourage use of academic language

Examples include:

- "I'm hearing _____ say that _____."
- So let me see if I understand. Your claim is _____."

c. Deliberate Practice (shifting to improving skills and knowledge)

25. Changing teacher practice can be a challenging endeavor, and one of the barriers to change is the influence of habits. For many teachers, sticking to familiar teaching methods is more convenient and comfortable than trying (and potentially failing at) new approaches. The sense of security that comes from familiarity can make it difficult to disrupt established instructional patterns. Fear of the unknown can also contribute to resistance to change (Fullan, 2001; Greenberg & Baron, 2000).

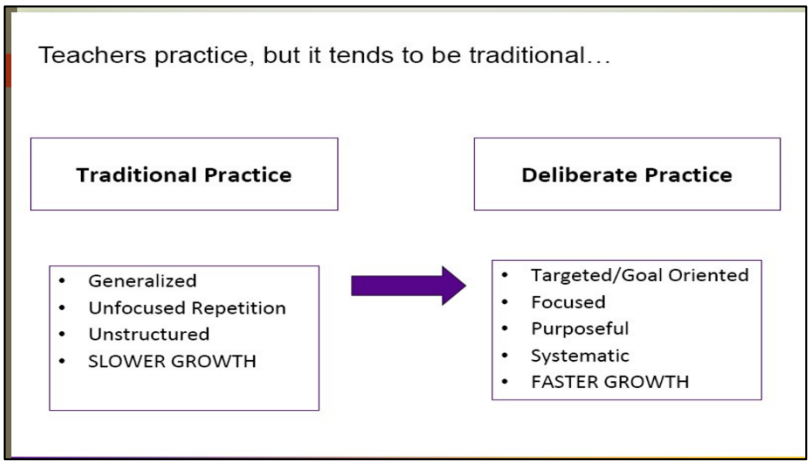
What can coaches do about this?

26. Creating a safe and supportive environment for teachers to develop new habits and skills is crucial. While there are opportunities for teachers to practice and refine their abilities, providing a safe space becomes paramount. This safe space allows teachers to experiment, take risks, and learn from their experiences without the fear of judgment or negative consequences

What is Deliberate Practice?

27. Research suggests that expertise is not innate but is instead developed through a process known as "deliberate practice" (Ericsson, Krampe & Tesch-Romer, 1993). Deliberate practice differs

from regular practice in its intentional and focused nature. Rather than simply going through the motions, deliberate practice involves targeted efforts to improve specific skills or areas of knowledge. It requires individuals to engage in purposeful practice, receive feedback, and continuously adjust to refine their abilities. Individuals can develop and enhance their expertise in a particular domain by consistently engaging in deliberate practice.



What does Deliberate (vs. Traditional) Practice look like?

28. Imagine that a group of children meet several times a week to play football. Over time, they will likely get better at the game because they will have spent time playing- they will have practiced. But this is an example of general or traditional practice; the children aren't necessarily thinking purposefully about how to get better at specific skills.

29. On child, however, is determined to get better at shooting the ball. At first, she focuses only on trying to kick the ball with her shoelaces while standing still. Once she has mastered that skill, she dribbles the ball before shooting. Then, she adds an obstacle to act as a defender. She practices and masters each of these skills individually before adding the next move. This is deliberate practice. The child's specific goal is to shoot with greater accuracy. Her practice is purposeful because it is designed to help her achieve this goal. It is systematic because she breaks down the skill into more manageable parts, practices each part until she improves, and then adds the next step.

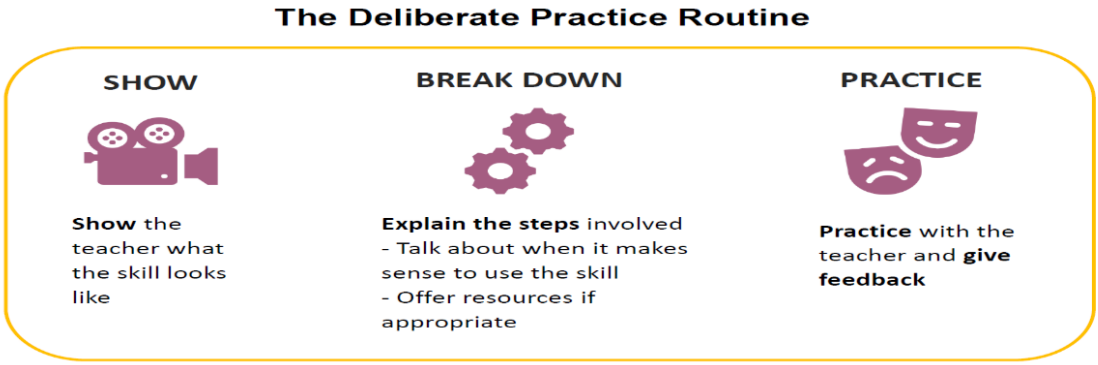
30. Deliberate practice isn't just for children learning new skills; many professionals also rely on it. Doctors first practice on cadavers and pilots use flight simulators, for example, because it gives them the chance to improve their skills without risking others' lives.

What does Deliberate Practice Look Like in Teacher Coaching?

31. Deliberate practice for teacher coaching consists of three key components: show, break down, and practice. In the ***Show*** phase, the coach demonstrates or models the new skill for the teacher, allowing them to observe it in action. In the ***Break Down*** phase, the coach explains the steps and considerations to acquire the new skill effectively. This includes setting clear goals, discussing the

appropriate context for its application, and providing any relevant resources. Finally, in the **Practice** phase, the teacher actively engages in implementing the new skill while receiving feedback and guidance from the coach. This iterative process of trying, receiving feedback, and trying again allows the teacher to refine and enhance their proficiency in the skill being practiced.

- (**Show**) focus the coaching conversation around specific teaching skills;
- (**Break Down**) Breaking Down skills to help teachers better understand the techniques associated with implementing a new skill;
- (**Practice**) Giving teachers a safe environment where they can try out the skill and receive feedback



32. These components of a coaching program, including reflection and discussion, are indeed crucial. While it may initially feel time-consuming, the emphasis on deliberate practice may reduce the time allocated for reflection and discussion. As coaches become more experienced, they become more efficient in facilitating each component, allowing teachers to consistently engage in reflection and deliberate practice within a manageable timeframe. The goal is to create a balanced coaching approach that maximizes the benefits of both reflection and deliberate practice for continuous professional growth.

Why use Deliberate Practice in Teacher Coaching?

33. Incorporating deliberate practice into teacher coaching can have several benefits for teachers. Firstly, it provides a focused approach by centering the coaching conversation around specific teaching skills, allowing for targeted improvement. Secondly, the process of showing, modeling, and breaking down skills helps teachers clearly understand how to implement new skills effectively. It provides them with tangible examples and practical guidance. Thirdly, deliberate practice creates a safe and supportive environment for teachers to learn new skills and receive constructive feedback. This safe space allows them to take risks without fearing negative consequences on classroom management or student achievement.

34. Given the multitude of responsibilities that teachers face, attempting a new skill without proper practice and support can be challenging and may have wide-ranging impacts. However, through multiple opportunities for deliberate practice in a safe environment, teachers can enhance their strategic thinking, accelerate their skill development, and cultivate new habits more efficiently.

This approach enables teachers to improve their instructional practices and ultimately positively impact student learning outcomes.

IV. Intervention

Target group and timeline

35. After reviewing the coaching model, the Ministry formed technical teams to develop an implementation plan for rolling out the model within the current school year. In addition to piloting the rollout, the team also set out to evaluate effectiveness.

36. The pilot phase involved 111 coaches who were assigned to work with teachers in English, French, and Biology subjects at grade levels 4-8. Of these coaches, 56 received comprehensive training on the entire model and its expected timeline for implementation. Various entities, including the World Bank, NYU, and the DOPs Core Project Team, conducted the training. It utilized presentations, simulation exercises, and videos to illustrate specific activities aligned with the new coaching model.

37. The dissemination of the coaching model during the pilot phase followed a multi-tier approach. The first type of training targeted the Core Project Team, which consisted of regional coach coordinators and other administrators in DOPS. This training aimed to ensure that the core team involved in the model's design thoroughly understood all its elements and implementation procedures. Additionally, this team played a role in identifying content for training videos to be used in subsequent model trainings.

38. The second type of training focused on coaches randomly selected to participate in the pilot. These coaches received training in the fall of 2019 and again in the fall of 2021 after schools reopened for the 2021-22 school year. The objective was to familiarize them with any modifications made to the coaching model. Before the training sessions commenced, the coaches watched a 7-minute video and scored it using the DOPS observation protocol, providing baseline data for the Impact Evaluation (IE) analysis.

39. It's worth noting that the coaching model also introduced peer learning activities for coaches and teachers, as well as fostering greater interaction between teachers and coaches. However, these activities could not be piloted or tested as planned due to prolonged school closures and teachers' strikes.

40. In addition to innovating the coaching model itself, there was also an innovative approach to delivering skills, knowledge, and professional support to teachers in Lebanon. Rather than a one-time training session for the teachers, the intervention adopted a continuous professional development approach facilitated by the coaches. These coaches individually visited teachers and dedicated 20 minutes to teach them a specific element of the model, such as the steps of Instructional Dialogue. The teachers were then encouraged to apply this element in their classrooms during the following week. The coach would return in the subsequent week to observe the teacher's implementation and effectiveness of the model, followed by a 20-minute discussion

to reflect on their learning. This process was repeated for each new element introduced, with a minimum of four to six visits per teacher.

41. The delivery method followed a cascade or training of trainers (ToTs) model, where initial coaches were trained in the complete model and then shared their knowledge and skills with other teachers. This created a personalized and ongoing support system, allowing teachers to practice and integrate each element of the coaching model before progressing. Gradually introducing the components facilitated behavior change and effective implementation in classrooms.

42. During the implementation phase of the Project, coaches trained in the model worked with two teachers, conducting visits from November 2021 to April 2022. Each visit consisted of observing the teacher in the classroom, followed by a meeting to review the previous lesson and discuss its implementation. Additionally, a new lesson was introduced during each visit. Although the original plan included a minimum of 6 visits, the shortened implementation period and concerns about disruptions led to a reduction to 4 visits, covering the essential elements of the coaching practice. Figure 7 below illustrates the schedule of coach visits and the proposed topics to be covered.

Figure 7: Implementation Schedule and Model dissemination adjusted for 4 visits

V. Evaluation Design

43. The World Bank conducted an Impact Evaluation study to measure the effectiveness of a new teaching approach. They used a randomized control trial (RCT) with two groups: the treatment group, which received training and implemented the new model, and the control group, which continued with traditional teaching methods. The study aimed to observe changes in teacher behavior in the classroom as they were trained on the new teaching modules. Coaches in English,

French, and Biology were randomly selected for the intervention and assigned fewer teachers than usual, visiting each teacher at least four times. Baseline data from treatment and control groups (teachers, coaches, principals, and students) was collected in early 2020, while endline data was collected in spring 2022 after significant delays caused due to COVID-19.

Figure 8: Working through the guiding principles of new teaching approach

44. In January-February 2020, 56 coaches across the three subjects underwent intensive training on the coaching model and the evaluation process. The main goals of the training were to learn and use an observation protocol, deepen understanding of Instructional Dialogue, practice consultation routines with colleagues (Critical Friends Group), and employ debriefing methods with teachers. The training followed a spiral curriculum, revisiting key content at multiple points throughout the training period. A spiral curriculum helps participants build a deeper understanding of the content over time through repeated exposure and practice.

Challenges

45. The complex local environment, coupled with the rigid education system Lebanon has known for decades, presented some challenges for implementation. While modifications could address some of these during the scale-up phase, many are exogenous. Key challenges included:

a. Some Teachers are resistant to adopting the new model or changing their traditional teaching methods, citing several impediments (e.g., hesitation to try something new, inability to devote energy to new things in the current environment, already having too much to focus on, simply trying-to-get-through-the-day, and other reasons);
b. Some Teachers left the school, shifted to different positions, changed subjects taught, or retired.
c. Some teachers were not informed about the new model, so it took some time to explain
d. Some classrooms had larger-than-normal numbers of students and were still hard to manage
e. Some teachers were absent during school visits, delaying the implementation timeline;
f. The political and economic context is challenging coaching efforts (strikes, road blockage)
g. The time needed for the debrief session with the teacher, mainly for deliberate practice, may be short, or the teacher is not available as they must run to the next class/school;
h. Some visits are scheduled too closely, leaving little time for teachers to practice;
i. Inclusive schools where students have special needs – teachers are challenged to introduce new models;
j. Misconceptions about the selection of teachers and correlation with their current performance (am I being coached because I'm doing poorly?);
k. Some teachers asked why they weren't selected to learn the new model.
l. Disruptions to education due to COVID-19-related school closures presented new challenges to teachers' already complex roles. These challenges included added strain in assessing students' learning loss, adjusting teaching to the differentiated needs of learners, and attending to students' socioemotional well-being and academic needs. Confronting the deepened learning crisis in the

face of the pandemic will require immediate and significant global investments to support teachers better and improve teaching quality[7].

Sampling Strategy

46. The evaluation employed a longitudinal sampling strategy, tracking the same coaches from baseline to endline. In January 2020, 111 coaches were selected and assigned to two schools and two teachers per school for data collection. The baseline sample consisted of 129 control group schools (258 teachers) and 92 treatment group schools (189 teachers).

47. For the endline data collection in May 2022, 107 of the original 111 coaches were surveyed again, resulting in a low attrition rate of under 4%. However, due to the pandemic and economic crisis in Lebanon, there was a relatively high attrition rate at the teacher level. Although approximately 90% of the treatment and control schools remained in the study, it was observed that only about 65% of the original teacher sample remained in the study between baseline and endline. This may be due to teachers who moved schools, exited the profession, or emigrated due to economic challenges. To maintain the sample size, the study team replaced the attrite teachers with their actual replacements at the schools. Despite the attrition, the random assignment of coaches to treatment and control groups ensured sample balance and equivalence at the school and teacher levels.

Evaluation Strategy

48. To assess the effects of the new coaching model, a randomized controlled trial (RCT) involving 111 coaches employed by the Lebanese Ministry of Education and Higher Education (MEHE) was conducted. Among them, 56 coaches were randomly selected to receive training in the new coaching model and implement it with their assigned schools and teachers. This group was considered the treatment group. The remaining 55 coaches formed the control group, representing the typical practices in the education system. By comparing the outcomes of the treatment and control groups, it was possible to determine the impact of the intervention.

49. The random assignment of coaches to the treatment and control groups ensured that they were similar in terms of their characteristics[6], both observed and unobserved. This design allowed us to compare coach and teacher behavior and outcomes under normal conditions (control group) with those under the new coaching intervention (treatment group).

50. The evaluation employed a difference-in-differences linear regression model that enabled us to compare changes in outcomes of treatment coaches between baseline and endline to those of coaches in the control group during the same period. Formally, we estimate the following linear regression equation:

$$Y_{ijt}=\alpha+\beta T_i+\gamma Post_t+\delta T_i.Post_t+\varepsilon_{ijt} \quad (1)$$

[7] Hammoud & Shuayb. (2021). *The impact of covid 19 lockdown on access and quality of education: reflections from teachers and students in Lebanon*. Centre for Lebanese Studies CLS, LAU Lebanese American University, LAU.

51. Yijt represents the outcomes of interest to the evaluation, including classroom observation scores, teacher classroom practices, teacher perceptions of coaching, self-efficacy, job satisfaction, etc., associated with coach i assigned to school/teacher j and observed in period t. Ti is an indicator that takes on a value of 1 if the coach belongs to the treatment group, 0 otherwise. Postt is an indicator variable that takes on a value of 1 if the period observed is the endline, 0 for the baseline. εijt denotes the idiosyncratic error term.

52. For ease of interpretation, we will denote the expected value of Y for the treatment group in the baseline and endline as Y(T | 0) and Y(T | 1), respectively. The expected value of Y for the control group in the baseline and endline will be represented as Y(C | 0) and Y(C | 1), respectively. Using this formulation, we can determine the baseline to endline growth for the treatment group as:

$$Y(T \mid 1)-Y(T \mid 0)=(\alpha+\beta+\gamma+\delta)-(\alpha+\beta)=\gamma+\delta \quad (2)$$

While the reciprocal for the control group is:

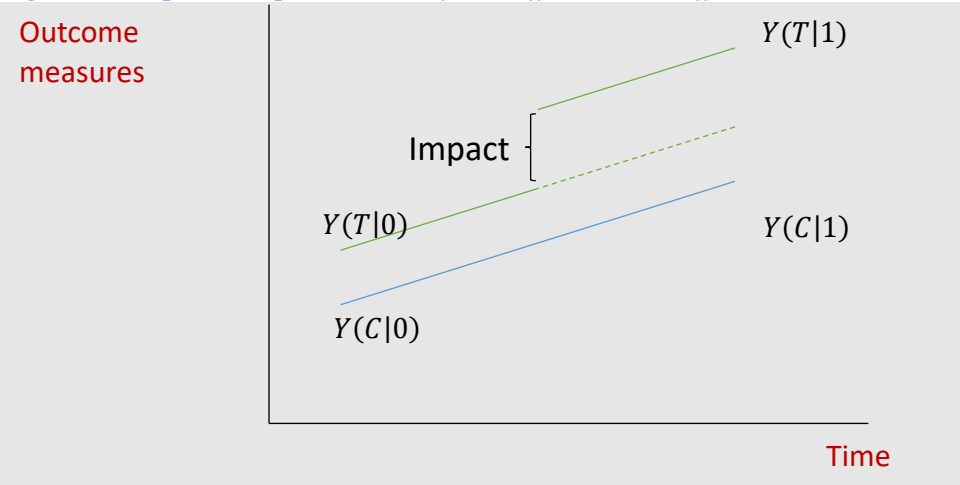
$$Y(C \mid 1)-Y(C \mid 0)=(\alpha+\gamma)-(\alpha)=\gamma \quad (3)$$

53. Lastly, the causal impact of the treatment can, thus, be computed as the difference between the two previous equations representing the differential growth between baseline and endline for the treatment group relative to the control group as follows:

$$\{Y(T \mid 1)-Y(T \mid 0)\}-\{Y(C \mid 1)-Y(C \mid 0)\}=\gamma+\delta-\gamma=\delta \quad (4)$$

54. As such, the parameter of primary interest to this evaluation is δ, which can be interpreted as the effect of the new coaching model on coach and teacher outcomes, all else being held equal. Figure 8 illustrates how the treatment effect is identified graphically.

Figure 9: Graphical Representation of the Difference-in-Differences Model



Measurement Instruments

55. The impact evaluation utilized several measurement instruments to assess the effects of the coaching intervention. These instruments were administered to different cohorts involved in the study, including coaches, teachers, principals, and students. Additionally, a classroom observation tool called the Protocol for Language Arts Teaching Observation (PLATO) benchmark tool was developed to measure the extent to which teachers implemented the new coaching model they had received. The data collected from coaches, teachers, and the classroom observation tool were analyzed to evaluate the impact of the intervention on teacher behavior and outcomes.

Link between the PLATO tool and the DOPS tool

The Protocol for Language Arts Teaching Observation (PLATO) is a classroom observation tool developed by Pam Grossman and her colleagues from Stanford University. Originally, the Plato tool was designed for use in English/Language Arts (ELA) classrooms and for students in grade 3 through 9 with an objective to study the relationship between teacher practices and student outcomes. It has since been used in different areas including research, teacher professional development. The PLATO tool usually captures 12 elements of instruction: Purpose, Intellectual Challenge, Representations of Content, Connections to Prior Academic Knowledge, Modeling and Use of Models, Strategy Use Instruction, Feedback, Classroom Discourse, Text-Based Instruction, Accommodations for Language Learning, Behavior Management and Time Management. These elements are divided into 4 categories: Representation and use of content, Disciplinary demand, Instructional scaffolding, and Classroom environment. Teachers are assessed for each element from 1 ('provide no evidence') to 4 ('provide consistent strong evidence').

The DOPS (Department of Scholastic and Pedagogical Guidance) also has a classroom observation tool which captures elements of instruction such as: level of teacher questioning, response to student answers, opportunities for all students to engage in dialogue and the extent to which students are participating in thinking activities. These elements are divided into 2 categories: Instructional Dialogue and Critical Thinking. Each element is also scored from 1 ('the teacher never or rarely engaged in that element of instruction') to 4 ('the teacher consistently engages in that type of instruction').

In short, the PLATO and the DOPS tool are both classroom observation tools and may be used in an additional manner.

56. The evaluation collected data at multiple times to gather information about the participants and their experiences. Baseline data was collected in January-February 2020, while interim data was collected from coaches throughout the implementation phase. The final round of data collection, known as endline data, took place in May 2022 after updating the baseline instruments.

57. The surveys administered during data collection aimed to gather both general data and demographic information from the respondents. This included details such as age, gender, and years of experience. Coaches were specifically asked about their prior training, experience as a coach, interactions with teachers, professional development needs, and interactions with other coaches. Similarly, teachers and principals were asked questions about their general information, experience working with principals, perceptions of teacher effectiveness, and the perceived

usefulness of coaching visits and feedback. These surveys helped fully understand the participants' backgrounds and perspectives on the coaching intervention.

Coach Sample Description and Balance

58. Table 2 provides an overview of coaches' average demographic and background characteristics in both the treatment and control groups, along with a test for sample balance at baseline. The analytic sample consists of 111 coaches, with 65 in the control group and 46 in the treatment group.

59. The results indicate that the coaches in the treatment and control groups are similar across most background characteristics. These characteristics include age, education level, whether they coached in the previous year, years of teaching experience, coaching experience, and whether they hold additional or supplemental employment. On average, the coaches are around 46 years old, have 24 years of teaching experience, and almost seven years of coaching experience. The majority of the coaches are women.

60. The data also reveal that approximately 26-29% of the coaches have a master's degree or higher educational attainment. Interestingly, nearly one-fifth of all coaches have other employment in addition to their coaching engagements. These findings highlight the composition and background of the coach sample, demonstrating similarity between the treatment and control groups at baseline.

Table 2: Demographics of coaches at baseline and the difference between the two groups

Demographics and General Data	Control	Treatment	Difference
Age	46.82	46.34	0.48
Education (master's or higher)	0.29	0.26	0.03
Gender (Females)	0.81	0.93	-0.12*
Coached last year?	0.52	0.55	-0.03
Years of teaching experience	24.02	23.26	0.76
Years of coaching experience	6.63	6.70	-0.07
Other employment	0.18	0.21	-0.03

Note: Figures in the table represent mean values for the control group, treatment group, and the mean difference between the two groups at baseline. Asterisk (*) denotes statistical significance at the 10% level.

61. The primary objective of this analysis is to assess the validity of the random assignment of coaches to the treatment and control groups. The results indicate that there were no statistically significant differences between the two groups for the majority of characteristics, suggesting that the assignment mechanism was not biased or purposefully targeted toward a specific group of coaches.

62. However, one notable difference was observed in the proportion of women in the treatment and control groups, with the treatment group having a 12% higher proportion of women, and this difference was statistically significant. Although this gender composition imbalance may introduce some potential bias in the impact estimates, it can be addressed by including coach-level covariates in the regression analysis. By controlling for these covariates, the potential bias arising from the gender difference can be mitigated, and it can be ensured that any observed impacts of the coaching intervention are not solely attributable to the gender composition of the groups.

Teacher Sample Description and Balance

63. The random selection of teachers as an extension of the random assignment of coaches into treatment and control groups can be interpreted as cluster randomization, where coaches represent a cluster of two schools and four teachers. This approach helps ensure that any differences in outcomes between the treatment and control groups are not due to systematic differences in the teachers' characteristics, as they were assigned based on the random assignment of coaches. The table below summarizes the demographics of teachers involved in the impact evaluation, taken at baseline.

Table 3: Demographics of teachers at baseline and the difference between the two groups

	Control	Treatment	Difference
Gender (Female)	0.96	0.94	0.02
Second Language of Instruction is English	0.61	0.64	-0.03
Employment status: Permanent	0.40	0.38	0.03
Employment status: Contract	0.49	0.50	-0.01
Employment status: Seconded	0.10	0.11	-0.01
The teacher has a bachelor's degree	0.61	0.56	0.05
The teacher has a master's degree	0.19	0.19	-0.00
The teacher has a teaching diploma	0.31	0.40	-0.09*
Years teaching	19.87	20.44	-0.57
Teach at other institutions or school	0.28	0.29	-0.01

Note: Figures in the table represent mean values for the control group, treatment group, and the mean difference between the two groups at baseline. Asterisk (*) denotes statistical significance at the 10% level.

64. The teachers in the treatment and control groups exhibit overall similarity in their observed characteristics, including gender, the second language of instruction, employment status, educational attainment, teaching experience, and whether they teach at other institutions or schools. The only significant difference was in the proportion of teachers holding a teaching diploma, with 40% of the treatment group reporting having a teaching diploma compared to 31% in the control group.

65. The validity of the random assignment was confirmed, indicating that there was no purposeful selection based on teacher characteristics. However, to address the baseline difference regarding teachers with a teaching diploma, covariate adjustments can be incorporated into the regression estimation process to mitigate the potential bias and ensure unbiased impact estimates.

For additional information on the coaching landscape and profile of coaches, see Annex 1.

VI. Evaluation Results

66. The previous section demonstrated how the treatment and control groups were randomly chosen and had similar socio-demographic characteristics at the beginning of the study. This section focuses on the results of the intervention.

67. Firstly, the study compares the average scores along various indicators of the treatment and control groups. These include coaching practices, the process of coaching, job satisfaction, and the content of discussions during classroom observations or teacher debriefing sessions.

68. The following section examines the impact of the coaching model by looking at how much the treatment group improved from the beginning of the study to the end, compared to the control group. The study used a statistical model called Ordinary Least Squares (OLS), which helped analyze the data. The model considers fixed effects at the coach or school levels, depending on whether the study measures outcomes for coaches or teachers/schools.

Descriptive statistics: Difference in means of outcome variables

69. The report presents findings from two perspectives: coaches and teachers. This means it analyzes the data separately for the coaches and the teachers involved in the study. Furthermore, it presents the results obtained from two specific classroom observation tools: PLATO and DOPS. These tools were used to assess the quality of classroom practices and provide valuable insights into the effectiveness of the coaching intervention.

Summary of the findings

70. The number of coaching visits per teacher decreased significantly from baseline to endline due to factors like the Covid-19 pandemic and other crises. Both the treatment and control groups were affected similarly. However, the treatment group coached fewer teachers, allowing for more focused training. The reduction in the number of teachers coached by the treatment group was aligned with the project's objective to improve education quality.

71. The treatment and control groups had similar focus areas during teacher debrief meetings, with slight differences observed. However, according to teachers' feedback, the treatment group showed a positive increase in pedagogical practices. Teachers felt that the coaching model focused more on pedagogical approaches and teaching methods, which aligned with the goal of improving teaching quality.

72. The treatment group coaches perceived an increase in effective communication, support, knowledge, and comfort working with teachers compared to the control group. They also reported

that their feedback was more valuable and practical. The intervention positively impacted coaches' perceptions of their coaching process and self-efficacy.

73. Both treatment and control coaches faced similar challenges at baseline, including large class sizes and insufficient time for debriefing. At the endline, treatment coaches still saw large class sizes as a challenge and reported receiving more support from principals. They also reported other challenges related to their additional responsibilities, time constraints, and professional development opportunities to a slightly greater extent than the control group.

74. Overall job satisfaction among coaches was high at baseline, except for satisfaction with their salaries. However, at the endline, there was a significant decrease in salary satisfaction among coaches, with only 2% reporting satisfaction. There were no significant differences in job satisfaction between the treatment and control groups.

75. In summary, the intervention had a positive impact on coaching practices, with a focus on pedagogical practices and teaching methods. Coaches in the treatment group perceived improvements in communication, support, knowledge, and self-efficacy. Challenges related to large class sizes and other responsibilities were more pronounced for the treatment group. However, job satisfaction, particularly regarding salary, decreased significantly for all coaches regardless of their treatment status.

76. Analyzing the proportions of teachers implementing specific practices, it was found that the coaching intervention had a significant impact on goal setting at the beginning of classes and encouraging critical thinking tasks. The treatment group showed an increase in these practices compared to the control group. However, there were no significant differences between the groups regarding small group work, classroom disciplinary management, and classroom discussions. The findings indicated positive outcomes aligned with the intervention's objectives.

77. The number of coach observations and debrief sessions received by teachers showed notable differences between the treatment and control groups. Treatment teachers received more coach observations and debrief sessions compared to the control group at the endline. The difference was statistically significant, suggesting a higher level of coaching engagement in the treatment group.

78. At the endline, treatment teachers had more positive perceptions of the benefits of coaches' feedback in terms of identifying areas of improvement, improving pedagogical practices, and enhancing student results compared to control teachers. Treatment teachers also reported a higher level of support, motivation, trust, and collaborative work with coaches during coaching sessions. These findings indicated that the coaching intervention had a significant positive impact on teachers' perceptions of the benefits they received from coaches' feedback and the coaching process.

79. Overall, the coaching intervention positively influenced goal-setting, critical thinking tasks, and teachers' perceptions of the benefits they received from coaches' feedback and the coaching process. The following section analyzes each area in detail.

The following sections delve into each area in detail.

Coaching practice

Changes in Coaches' Perspective and Behavior

80. The analysis of the coach survey data provides insights into various aspects of coaching practice, including workload, teacher debriefs, content of teacher feedback, self-reported considerations during class observations, self-efficacy, perceived challenges, and job satisfaction.

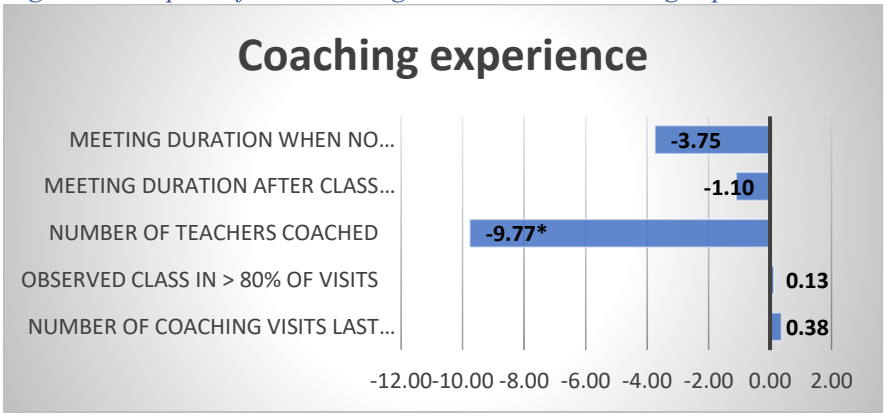
81. Starting with coaching practice, Table 4 (Annex 1) presents the findings. It shows the average number of coaching visits completed per teacher during the intervention timeline, the percentage of visits that included lesson observations (at least 80% of visits), the number of teachers coached, and the duration of teacher debriefs.

82. Both the treatment and control coaches completed around 30 visits per teacher at the beginning of the study (baseline), but this number decreased to an average of 6 visits per teacher at the end of the study (endline). The decrease in visits can be attributed to several factors, such as the impact of the Covid-19 pandemic, economic crises, teacher strikes, and rising fuel prices. These factors hindered many coaches from carrying out their duties effectively. It is important to note that the Covid-19 crisis occurred during the project's implementation phase, resulting in school closures and restrictions imposed by the Government.

83. The absence of a significant difference in the number of coaching visits per teacher between the treatment group and the control group at the endline indicates that both groups were equally affected by the various crises mentioned earlier. However, there was no significant difference between the treatment and control groups regarding the number of teachers coached. Before the project implementation (at baseline), there was no significant difference in the number of teachers coached by each group (on average, approximately 50 teachers). However, at the endline, the treatment group showed a significant reduction in the number of teachers coached compared to the control group. This reduction aligns with one of the project's objectives, which is to improve the quality of education by providing focused training to a smaller group of teachers. By reducing the number of teachers coached, the coaches were able to allocate more time and attention to each teacher, thereby enhancing their teaching practices. Specifically, the number of teachers coached by the treatment group decreased from an average of 51.20 teachers to 26.63 teachers.

84. In conclusion, the intervention reduced the number of teacher assignments for the coaches in the treatment group compared to their counterparts in the control group. Although the study expected the duration of the debrief meetings to have a positive and significant impact, it did not find any significant effect or difference between the coaches and teachers, regardless of their treatment status. This can be attributed to the reduced visits and interactions between people during the Covid-19 pandemic. Consequently, the duration of meetings between teachers and coaches decreased for both groups at the endline.

Figure 10: Impact of the coaching model on the coaching experience



Teachers and Coaches' Debrief Focus:

85. Tables 5 (Annex 1) and 12 (Annex 1) provide insights into the topics which coaches in the treatment and control groups focused on during their meetings with teachers. These topics include classroom management, pedagogical practices, content and quality of assessments, use of formative assessments, subject knowledge, and proficiency in the language of instruction. It is important to note that coaches could report spending the most time on multiple focus areas simultaneously when providing feedback to teachers.

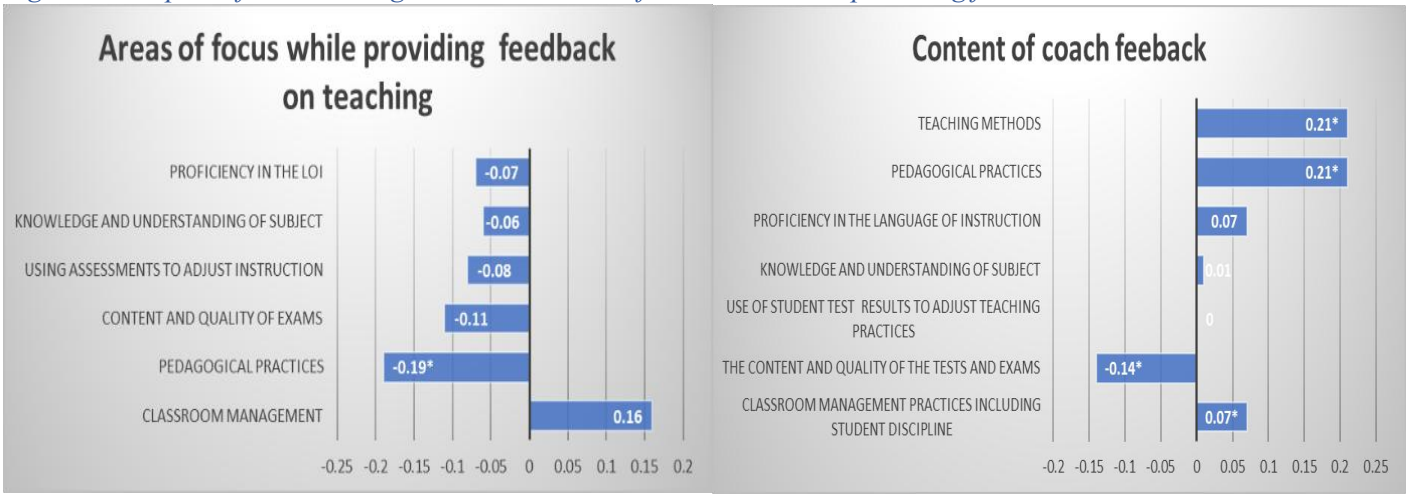
86. At both the baseline and endline, there were no significant differences between the treatment and control groups, except for pedagogical practices at the endline. This means that the two groups had similar reported areas of focus during teacher debrief meetings at the beginning of the study, and none of the baseline differences were statistically significant. On average, the treatment coaches showed a slightly increased focus on classroom management and a decreased focus on pedagogical practices, quality of exams, and language proficiency when providing teacher feedback.

87. However, Table 12 (Annex 1) reveals that according to teachers' feedback, there was a positive increase in pedagogical practices at the endline for the treatment group. In this table, the teachers were asked to report on the topics and content of the coaches' feedback during their meetings or visits. The teachers reported whether their coach discussed or provided input on various aspects, such as classroom practices, exam content, formative assessments, subject knowledge, language proficiency, pedagogical practices, and teaching methods. The coaching model aims to improve teaching quality, so teachers' perceptions of the intervention are crucial. The objective was achieved if teachers felt that coaches focused more on pedagogical practices and teaching methods.

88. The subsequent figures illustrate the impact of the coaching model when examining teachers' and coaches' feedback on the content. From the coaches' perspective, the study did not find a significant effect of the coaching model. However, from the teachers' perspective, the study observed the coaching model's positive and significant effect on pedagogical practices and teaching methods. Teachers reported that coaches primarily focused on pedagogical practices and teaching methods during their interactions. This outcome aligns with the expected effect of the coaching model and will likely positively influence students' achievements. In fact, teachers were

asked to assess whether the coach's feedback helped them identify areas of improvement, enhance their pedagogical practices, and improve student results [as shown in Table 13 (Annex 1)]. Teachers from the treatment group were more likely to state that the coaching was helpful and that they noticed an improvement in student results. This effect was indeed statistically significant.

Figure 11: Impact of the coaching model on coaches' focus areas while providing feedback



Coaching Process and Self-Efficacy:

Table 6 (Annex 1) presents coaches' self-reported perceptions of their coaching process and self-efficacy

89. They rated themselves on their level of knowledge and comfort working with teachers, level of communication, level of support provided to teachers, and whether they thought the feedback they provided was useful and practical. Here are the key findings:

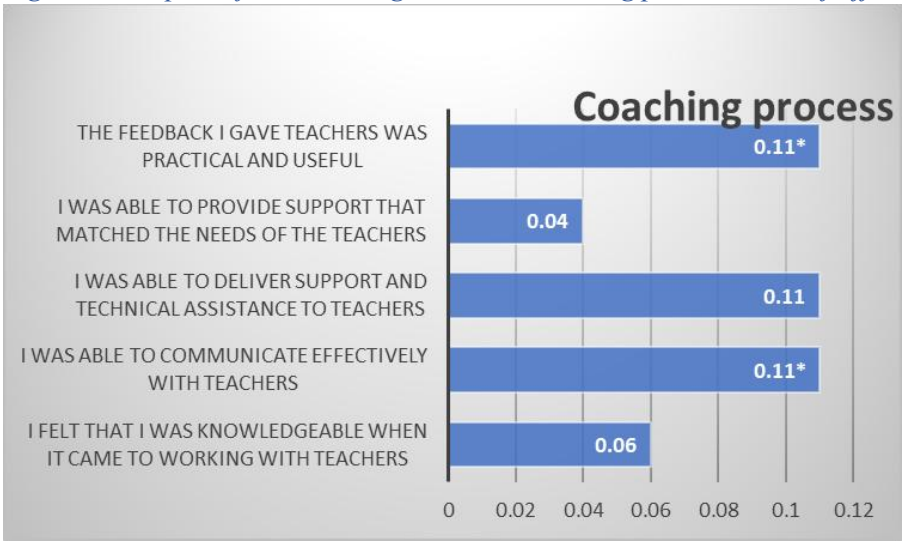
1. Communication and Support: At baseline, there were no significant differences between the treatment and control groups regarding coaches' perceived level of communication with teachers, the level of support and technical assistance provided to teachers, and whether the support matched the teachers' needs. However, at the endline, the treatment group showed an increase in the probability of perceiving effective communication, delivering support and technical assistance, and delivering support that matched the teachers' needs compared to the control group.
2. Knowledge and Comfort Working with Teachers: The control group had higher self-perceptions of their level of knowledge and comfort working with teachers at baseline. However, at endline, their perceptions decreased while the treatment group's perceptions increased.
3. Usefulness and Practicality of Feedback: At baseline, the control coaches had a higher probability of perceiving their feedback to teachers as useful and practical compared to the treatment group. However, at the endline, the treatment coaches had a higher self-perception in terms of the usefulness of the feedback they provided.

90. The figure below illustrates the impact of the coaching model on coaching process and self-efficacy. It demonstrates a positive and significant effect of the intervention, except for the perception of providing support that matched the needs of teachers. The treatment group outperforms the control group in terms of coaches' perceptions. For example, the intervention increased the likelihood that coaches perceive themselves as knowledgeable in their ability to coach teachers by approximately 31 percentage points. It also improved communication between

coaches and teachers by 18 percentage points and increased the likelihood of delivering support to teachers by 19 percentage points. Additionally, the treatment coaches' self-perception of the usefulness of their feedback increased by 11 percentage points compared to the control group.

91. Overall, the intervention had a positive impact on coaches' perceptions of their coaching process and self-efficacy. Coaches felt they had greater value, knowledge, and comfort in their interactions with teachers, and they had designated support to share with teachers regarding teaching.

Figure 12: Impact of the coaching model on coaching process and self-efficacy



Coaches' Perspective on Challenges:

92. The coaches were asked to report the challenges or barriers they face in performing their coaching responsibilities at the school and professional levels. Table 8 (Annex 1) presents the proportion of coaches in each outcome variable.

93. Both treatment and control coaches had similar perceptions of the challenges they faced at baseline. The main challenges reported by nearly 75% of coaches from both groups were large class sizes, insufficient time to debrief with teachers, and teachers lacking adequate teaching skills.

94. At endline, the only significant shift in coaches' perceptions of challenges was related to large class sizes. The treatment coaches still viewed large class sizes as a challenge, with 79% of them reporting it compared to only 60% of the control group.

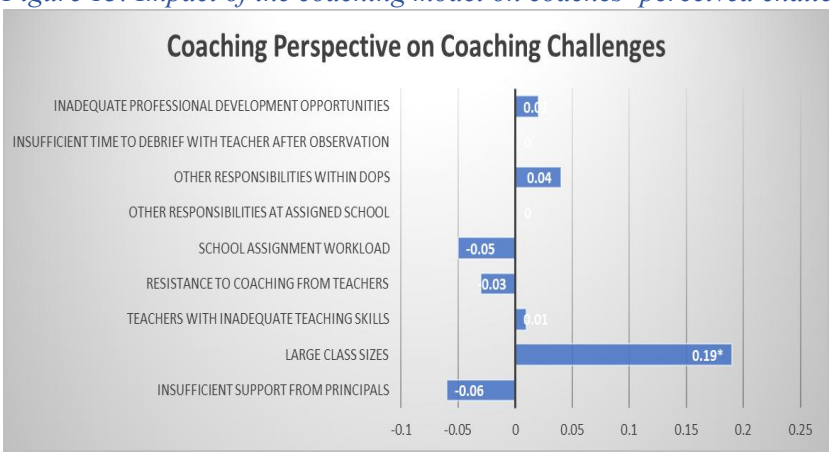
95. The impact of the intervention indicates that treatment coaches experienced specific changes in their perceptions of challenges compared to the control group. Specifically:

1. Support from Principals: Treatment coaches felt they had more support from principals and administration than the control group. They were 18 percentage points less likely to report insufficient support from principals.

2. Large Class Sizes: Treatment coaches were 29 percentage points more likely to consider large class sizes as a significant challenge to their coaching ability. This suggests that the new coaching model may have introduced elements that make managing larger class sizes more challenging.
3. Other Challenges: Treatment coaches also reported other challenges at a slightly higher rate compared to the control group. These challenges included their own additional responsibilities within DOPS, insufficient time to debrief with teachers following observations, and inadequate professional development opportunities. The differences ranged from 9 to 13 percentage points.

96. Overall, the intervention had an impact on coaches' perceptions of challenges. Treatment coaches felt they received more support from principals, recognized the challenge of large class sizes, and reported other challenges related to their own responsibilities, time constraints, and professional development opportunities to a slightly higher extent compared to the control group.

Figure 13: Impact of the coaching model on coaches' perceived challenges



Job Satisfaction:

97. The coaches were asked about their job satisfaction and various aspects related to their coaching role, including their career choice, overall job satisfaction, satisfaction with their salaries, and satisfaction with the support received from their coaching center and subject coordinators. Here are the findings from Table 9 (Annex 1):

98. **Baseline Job Satisfaction**: Both treatment and control coaches reported elevated levels of satisfaction with their job during the baseline period. Approximately 80% of coaches from both groups expressed satisfaction with all aspects of their job, except for their salary.

99. **Salary Satisfaction:** Before the pandemic, around 24-28% of coaches reported being satisfied with their coaching salary. However, at endline, only 2% of coaches reported being satisfied with their salary. This indicates a significant decrease in salary satisfaction among coaches.

100. Notably, there were no significant differences between the treatment and control groups in terms of job satisfaction at both baseline and endline. This means that the intervention did not have a significant effect on coaches' job satisfaction.

Figure 14: Impact of the coaching model on coaches' levels of job satisfaction

Changes in Teachers' Perspectives and Behavior

101. The survey data collected from both the treatment and control groups of teachers aimed to understand their classroom practices, coaching experiences, self-efficacy, and perceived challenges. Descriptive statistics were used to provide an overview of these variables and analyze any changes or differences over time. The statistics served as a starting point for further analysis of the impact of the coaching intervention on teachers' experiences and outcomes.

102. The purpose of collecting this information is to understand how the coaching intervention and other factors may have influenced teachers' classroom practices and behaviors, their confidence in their abilities (self-efficacy), and the challenges they faced.

Teacher Interpersonal and Focal Practices:

103. Table 10 (Annex 1) presents the proportions of teachers from both the control and treatment groups who frequently or always implement specific practices during their lessons. Here are the key findings:

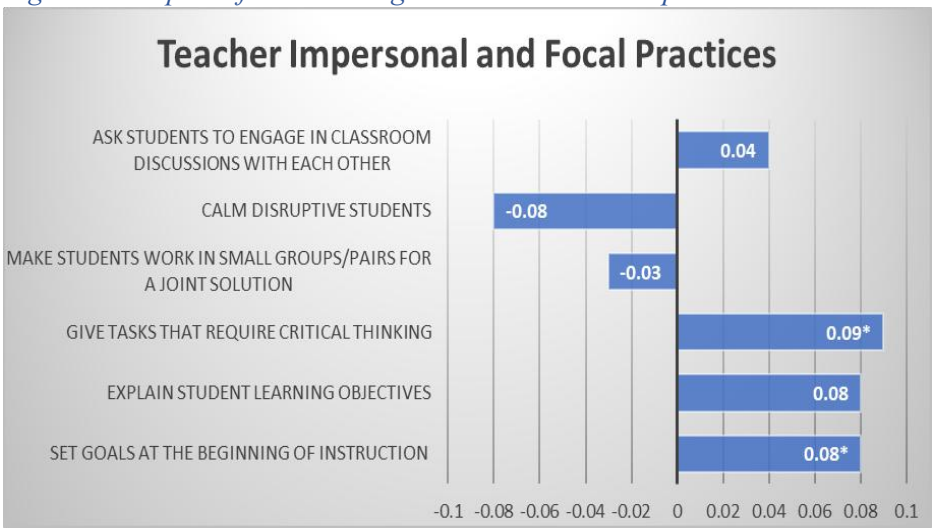
1. Setting Goals at the Beginning of Class: At baseline, approximately 66% of teachers from both groups reported setting goals at the beginning of their classes. However, after the intervention, there was a significant difference between the treatment and control groups. Specifically, 76% of teachers from the treatment group now set goals at the beginning of the class, compared to 68% in the control group. This represents a significant difference of 8 percentage points at baseline.
2. Engagement in Children's Cognitive Development Activities: Initially, there was no significant difference between the treatment and control groups in terms of engagement in children's cognitive development activities. However, at the endline, there was a significant increase in the proportion of teachers from the treatment group who frequently or always gave tasks requiring critical thinking. This proportion increased to 33% in the treatment group, while it decreased to 24% in the control group. This finding aligns with the intervention's objective of promoting increased critical engagement and thinking among students during lessons.
3. Small Group Work, Classroom Disciplinary Management, and Classroom Discussions: There were no statistically significant differences between the treatment and control groups in terms of the proportion of teachers who frequently or always made students work in small groups, engaged in

classroom disciplinary management, or asked students to engage in classroom discussions. These variables did not show significant differences at either the baseline or endline.

4. Opportunities for Dialogue Among Students: The treatment group teachers were 4 percentage points more likely to give opportunities for dialogue among students (students-to-students) compared to the control group.

104. These findings provide insights into the implementation of specific practices by teachers in both the treatment and control groups. The intervention had a significant impact on goal setting and encouraging critical thinking tasks, while there were no significant differences in other areas, such as small group work, classroom disciplinary management, and classroom discussions. The increase in goal-setting and critical thinking tasks among the treatment group indicates positive outcomes aligned with the intervention's objectives.

Figure 15: Impact of the coaching model on Teacher Impersonal and Focal Practices



Coaching feedback to teachers

105. According to Table 11 (Annex 1), teachers received an average of 2 coach observations per school year at baseline. However, by endline, there were some notable differences between the treatment and control groups.

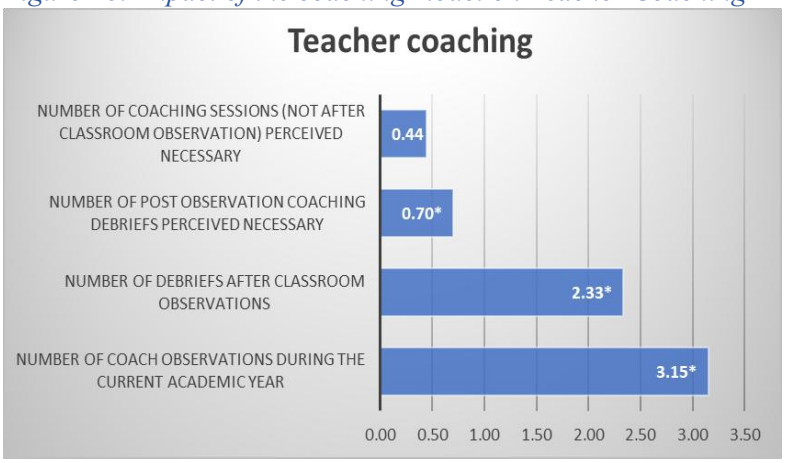
106. Coach Observations: Control teachers received almost one fewer observation per year than baseline, while treatment teachers received an additional 2 observations. This means that treatment teachers received, on average, 3 more coach observations than the control group at the endline.

107. Debrief Sessions: Both groups reported receiving between 2.2 and 2.5 debrief sessions after classroom observations at baseline. However, at the endline, the treatment group reported receiving 4.3 debriefs on average, while the control group reported receiving just under 2 debriefs. This indicates that treatment teachers received, on average, 2 more debrief sessions than the control group at the endline.

108. In terms of teachers' perceptions, both treatment and control teachers indicated that an average of 3 coaching sessions were necessary to improve their practice, both at baseline and endline. However, the number of necessary debrief sessions differed between the two groups. At baseline, both groups reported needing about 3.4 debriefs to improve their practice. But at the endline, treatment teachers reported needing an average of 4.3 debriefs, while control teachers reported needing an average of 3.6 debriefs.

109. Overall, the figure indicates that at the endline, treatment teachers received more coach observations and debrief sessions compared to the control group. The difference between the two groups in terms of the number of coach observations and debriefs received is statistically significant.

Figure 16: Impact of the coaching model on Teacher Coaching



Perceived benefit of the coaches' feedback and follow-up:

110. Table 13 (Annex 1) highlights the differences between treatment and control teachers at baseline and endline regarding their perceptions of the benefits of coaches' feedback and coaching sessions. The following observations can be made:

1. Coaches' Feedback:

Baseline: There was no significant difference between the treatment and control groups in terms of their perceptions of coaches' feedback.

Endline: Significant differences were observed between the two groups. The proportion of treatment teachers who reported that the feedback was immensely helpful in identifying areas of improvement, improving pedagogical practices, and improving student results increased compared to the control group. This difference was statistically significant, with an increase of 21 and 22 percentage points for improving pedagogical practices and improving student results, respectively.

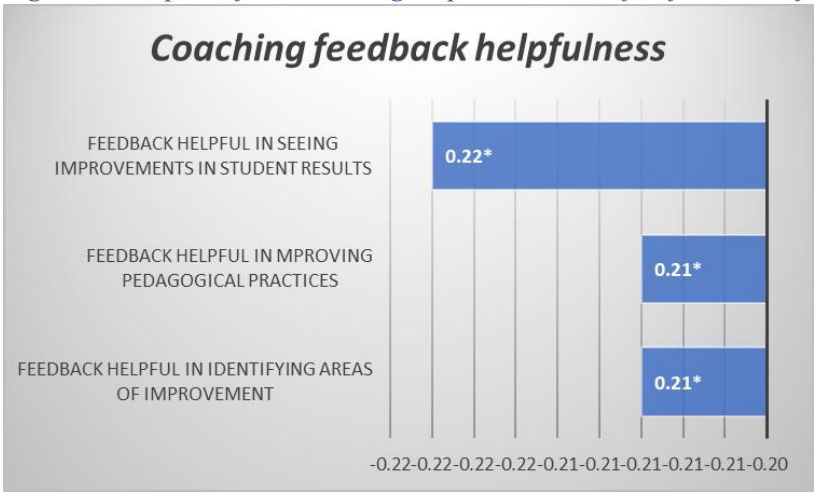
2. Coaching Sessions:

Endline: Treatment teachers were more likely to agree that coaches supported their instruction by providing constructive feedback, suggesting alternative practices or approaches, and offering motivation and encouragement during coaching sessions. They also reported a higher level of mutual trust and collaborative

work with coaches, including the identification of goals to focus on. The difference between the treatment and control groups was statistically significant.

111. In summary, at endline, treatment teachers had more positive perceptions of the benefits of coaches' feedback in terms of identifying areas of improvement, improving pedagogical practices, and improving student results compared to control teachers. Additionally, treatment teachers reported more support, motivation, trust, and collaborative work with coaches during coaching sessions. These findings suggest that the coaching intervention had a significant positive impact on teachers' perceptions of the benefits they received from coaches' feedback and the coaching process.

Figure 17: Impact of the coaching on perceived benefit of coaches' feedback to the teachers



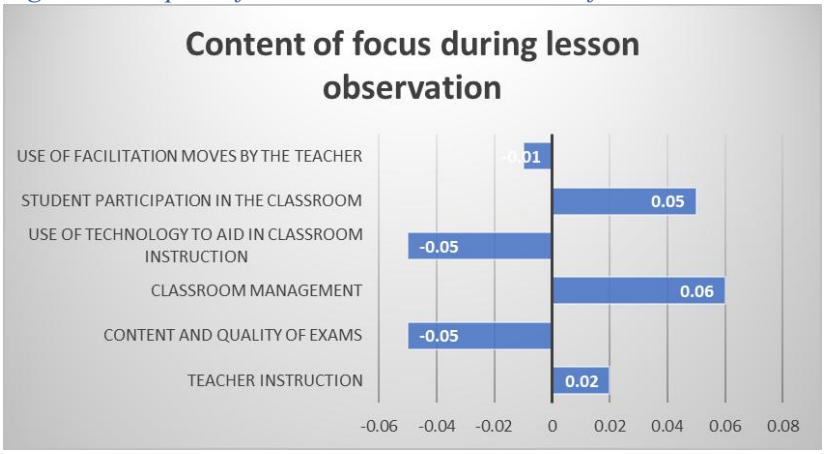
Classroom Observations – Changes in teaching practices

112. When coaches observed teachers' lessons, they paid attention to various aspects such as classroom practices, use of technology, and specific teaching techniques. Table 7 (Annex 1) provides insights into whether coaches consistently focused on and analyzed certain elements during their observations, including teacher instruction, assessment content, classroom management, student participation, and facilitation moves.

113. The results indicate that there was no significant difference between the treatment and control groups at both the endline and baseline. However, it was found that coaches in the treatment group were slightly more likely to prioritize teacher instruction and classroom management compared to the control group at the endline. This finding aligns with the coaches' own perceptions, as reported in Table 5 (Annex 1), where they indicated a higher likelihood of focusing on classroom management when providing feedback on teaching practices.

114. Additionally, the figure below illustrates the impact of the intervention on coaches' focus areas, namely student participation in the classroom, classroom management, and teacher instruction. Although the intervention had a positive effect on these aspects, the differences observed were not statistically significant. The magnitude of the impact estimates was also relatively small. This outcome may not be surprising since the new coaching model did not explicitly modify the coaches' focus during classroom observations.

Figure 18: Impact of the intervention on coaches' focus areas



PLATO and DOPS Classroom observation tools (monitoring for uptake of the model using a 3rd party independent tool)

115. The classroom observation process aimed to capture the various activities undertaken by teachers during their lessons and assign scores based on the PLATO and DOPS scoring rubrics. Independent observers, rather than coaches, conducted these classroom observations. Each observation consisted of two 15-minute segments. After the first segment of 15 minutes of teaching, observers assessed the structures of the activities and whether a teacher spent at least 2-3 on various elements such as lecturing, question and answer sessions, small group and whole group discussions, student presentations, independent student work, and teacher-led small group discussions. The same process was repeated in the second segment. After each segment of 15 minutes, observers also scored teachers on items from the DOPS classroom observation and PLATO tools, such as intellectual challenge, classroom discourse, intellectual dialogue, and critical thinking.

116. Table 15 presents data on the lesson structures observed during baseline and endline. It shows that both treatment and control teachers had similar patterns in structuring their lessons. Lecturing and engaging in brief student responses to teacher questions were the most common activities, occurring for at least 2-3 minutes in over 75% of the observed lessons. Student independent work and whole group discussions were also present in around 20-29% of the observed lessons.

117. These findings suggest that both treatment and control teachers had comparable lesson structures at baseline and endline. The data provides insight into the common activities implemented by teachers during their lessons, highlighting the prevalence of lecturing and teacher-led question-and-answer sessions. The PLATO and DOPS observation tools facilitated the monitoring of these activities and allowed for the assessment of intellectual challenge, classroom discourse, intellectual dialogue, and critical thinking.

Table 15: Classroom Observation

	Baseline			Endline		
Segment 1	Control	Treatment	Difference	Control	Treatment	Difference
Teacher Lecture	0.76	0.75	0.01	0.81	0.87	-0.06

Short Student Responses to Teacher Questions	0.86	0.85	0.01	0.81	0.82	-0.01
Small Group or Partner Discussions	0.15	0.09	0.06*	0.08	0.04	0.03
Whole Group Discussion	0.28	0.31	-0.03	0.09	0.29	-0.20*
Student Presentations	0.05	0.07	-0.02	0.08	0.05	0.03
Student Independent Work	0.20	0.27	-0.06*	0.24	0.21	0.03
Teacher-led small Group	0.06	0.05	0.01	0.02	0.02	0.00
Segment 2						
Teacher Lecture	0.69	0.67	0.02	0.79	0.82	-0.03
Short Student Responses to Teacher Questions	0.83	0.81	0.02	0.77	0.78	-0.02
Small Group or Partner Discussions	0.19	0.21	-0.02	0.16	0.10	0.06
Whole Group Discussion	0.29	0.29	0.00	0.12	0.31	-0.20*
Student Presentations	0.08	0.08	0.00	0.12	0.07	0.06*
Student Independent Work	0.34	0.32	0.02	0.38	0.27	0.11*
Teacher-led small Group	0.08	0.07	0.02	0.02	0.03	0.00

118. Lastly, let's examine the classroom observations of teachers from the treatment and control groups. We specifically looked at how well teachers engaged in intellectual challenge activities, encouraged classroom discourse, fostered intellectual dialogue, and promoted critical thinking. These concepts were rated on a scale of 1 to 4, with higher scores indicating better implementation. The observations were conducted twice during each lesson to assess consistency.

119. Overall, the teachers received average scores ranging from 1.75 to 2.27 out of 4, except for student-to-student dialogue, which averaged between 1.23 and 1.46. These scores suggest that both the treatment and control groups have room for improvement in applying these activities and techniques. The control group did not show considerable progress in these areas at the endline, indicating that their usual approach did not focus on these concepts. In contrast, the treatment group scored higher on almost all concepts, with averages ranging from 2.2 to 2.34 (except for student-to-student dialogue), while the control group averaged between 1.69 and 1.98 on the same concepts.

120. In summary, the treatment group demonstrated better implementation of intellectual challenge activities, classroom discourse, intellectual dialogue, and critical thinking compared to the control group. However, there is still room for improvement for all teachers in effectively incorporating these practices into their lessons.

Table 16: Impact of Model Uptake

PLATO scoring & DOPS Scoring[9]	Baseline			Endline		
Segment 1	Control	Treatment	Difference	Control	Treatment	Difference
Intellectual Challenge	1.83	1.81	0.02	1.66	2.2	-0.53*

CD-Uptake	1.83	1.90	-0.06	1.81	2.22	-0.40*
CD-Opportunities	2.16	2.19	-0.02	1.77	2.23	-0.46*
ID-Level of Questioning	1.76	1.82	-0.07	1.71	2.23	-0.52*
ID-Response to Student Answers	1.94	2.08	-0.14*	1.92	2.34	-0.43*
ID-Student to Student Dialogue	1.23	1.27	-0.03	1.3	1.41	-0.11
Critical Thinking	1.75	1.78	-0.03	1.69	2.18	-0.50*
Segment 2						
Intellectual Challenge	1.91	1.97	-0.05	1.87	2.25	-0.38*
CD-Uptake	1.81	1.93	-0.11*	1.91	2.22	-0.30*
CD-Opportunities	2.22	2.27	-0.05	1.9	2.28	-0.39*
ID-Level of Questioning	1.82	1.84	-0.03	1.81	2.25	-0.44*
ID-Response to Student Answers	1.93	2.03	-0.10*	1.98	2.32	-0.34*
ID-Student to Student Dialogue	1.42	1.46	-0.03	1.37	1.56	-0.19*
Critical Thinking	1.89	1.93	-0.04	1.85	2.24	-0.39*

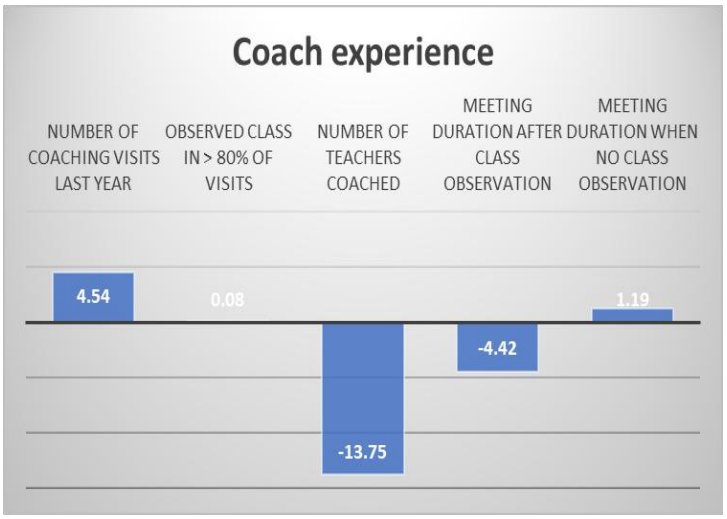
Relevance to existing coaching system

121. This section of the report focuses on presenting the estimated impacts of the new coaching model on both coaches and teachers. Specifically, it examines the effects on pedagogical practices, teaching methods, and student outcomes, providing insights into how the implementation of the coaching model has influenced these key areas.

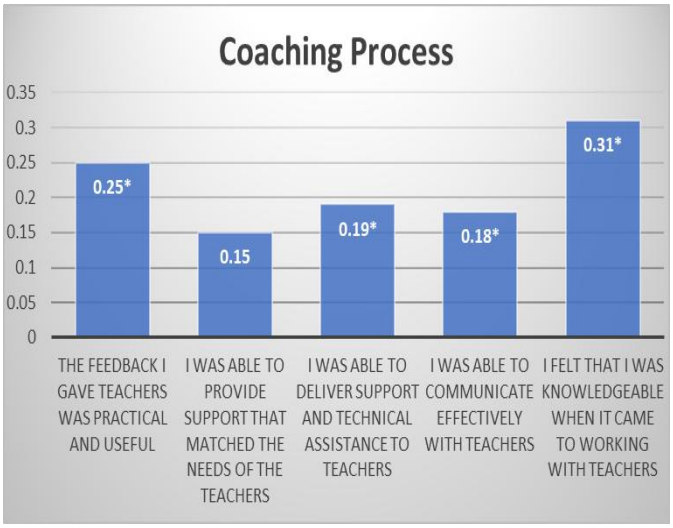
1. Coaches

122. This report subdivided the impacts of the coaching model on coaches into six categories. On average, the coaching model did not show significant effects on coaches, except for variables related to the "coaching process." This outcome is expected since the main focus of the coaching model is on teachers and students. The results are presented in the graphs below as the difference between the treatment and control group at the endline.

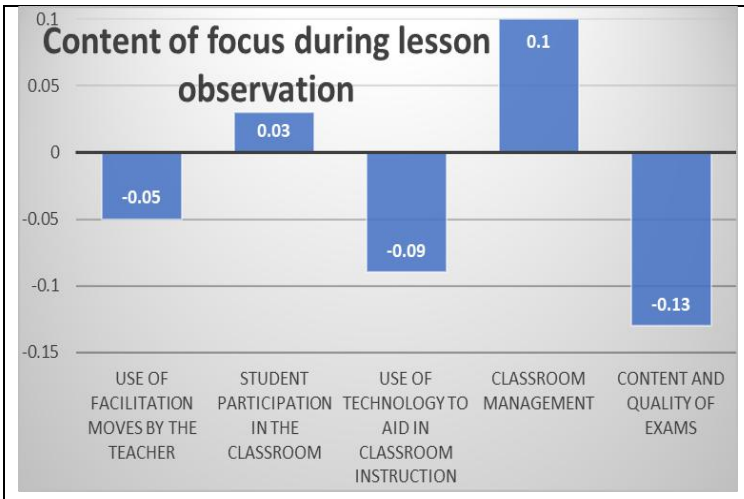
123. In terms of **coach experience**, the findings indicate no significant effects. However, there were positive effects observed in terms of reducing the number of teachers assigned to each coach (by 13.75 teachers) and increasing the number of coaching visits per teacher (by 4.54 visits). This implies that by reducing the coaches' workload, they could provide more frequent visits and dedicate more time to each teacher outside of classroom observation.

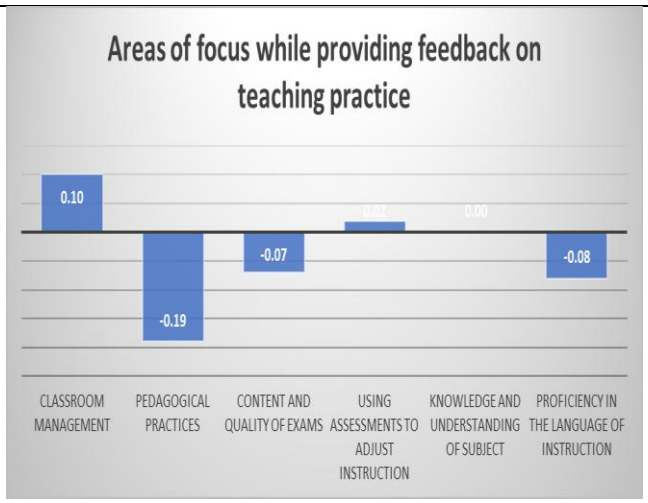


124. Concerning the **coaching process**, all variables except for "the ability to provide support to teachers that matched their needs" showed positive and noteworthy results. The coaching model improved communication between coaches and teachers, the delivery of support and technical assistance to teachers, and coaches' knowledge about working with teachers. Specifically, coaches' perceptions of effective communication with teachers and delivering support and technical assistance to teachers increased by 18 and 19 percentage points, respectively. Their perception of knowledge about working with teachers also increased by 31 percentage points. Furthermore, coaches' perceptions of feedback being practical and useful significantly increased by 25 percentage points.



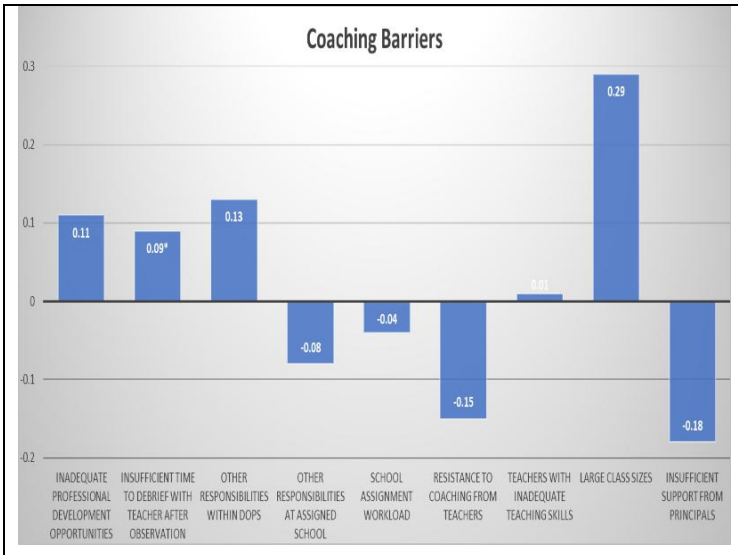
125. The results regarding **areas of focus during lessons' observation and feedback with teachers** were not significant but still provided some valuable insights. Coaches were more likely to focus on classroom management (an increase of 10 percentage points) and students' participation in the classroom during lessons' observation (an increase of 3 percentage points). These findings align with the objective of the coaching model to equip teachers with tools to encourage student participation in the classroom.

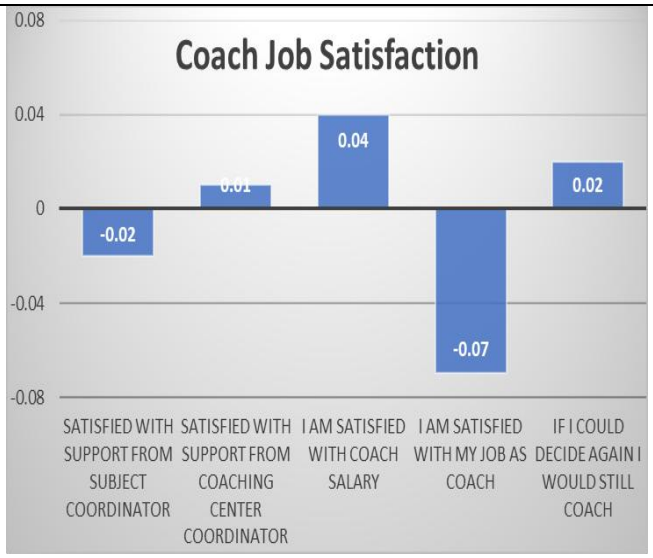




126. The coaching model shed light **on barriers that hinder coaches' ability** to effectively coach teachers, although the effects were not significant. Factors such as large class sizes, teachers with inadequate teaching skills, and limited professional development opportunities continue to pose challenges for coaches, which can also impact students' learning outcomes.

127. Additionally, the new coaching model had a positive impact (although not significant) on **support from principals and teachers**. Coaches reported receiving increased support from principals, with a decrease of 18 percentage points in coaches reporting insufficient support. The percentage of coaches reporting teachers' reluctance to coaching also decreased by 15 percentage points. One notable impact of the coaching model on coaching barriers was the duration of debrief sessions after classroom observations. Coaches expressed a desire for more time during debrief sessions, which may be attributed to the reduced number of teachers coached by each coach and the teachers' increased openness to coaching.

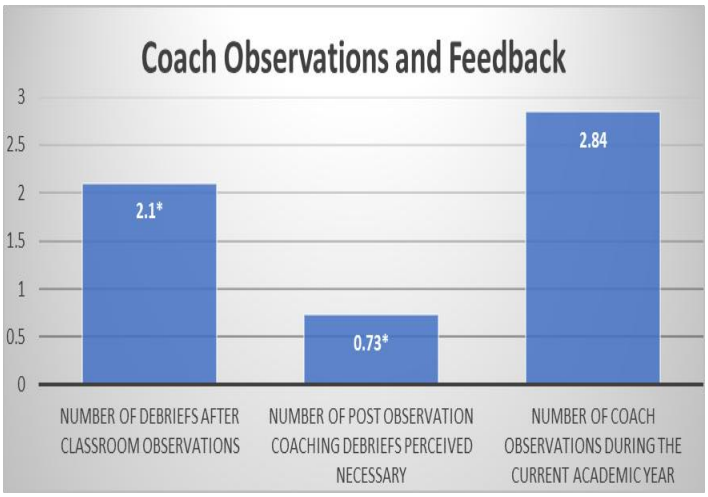




128. Overall, the results indicate various positive impacts of the coaching model on coaches, albeit with some nonsignificant findings. These findings provide valuable insights into the implementation and effectiveness of the coaching model in improving coaching processes and support for teachers.

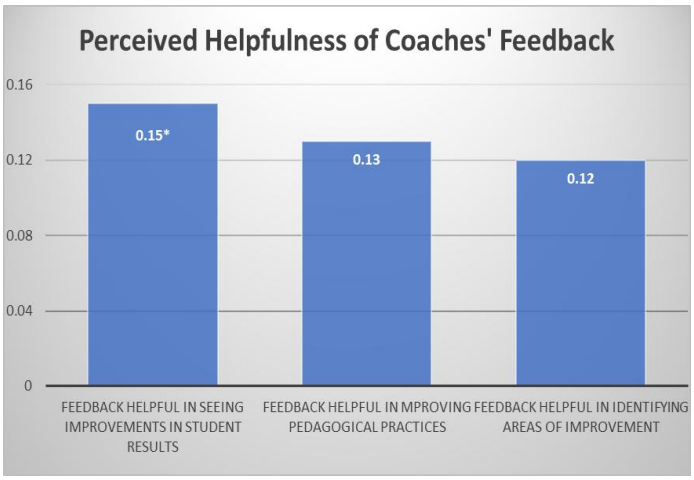
2. Teachers

129. Several expected effects were observed in our estimates of the intervention impacts on teachers using the difference-in-differences model. Firstly, teachers received more visits and feedback from coaches, with the number of coach observations and debrief sessions increasing by 2.84 and 2.1, respectively. Additionally, there was a significant increase in the number of debriefing sessions following classroom observations, indicating a greater emphasis on reflective practices and collaborative discussions.

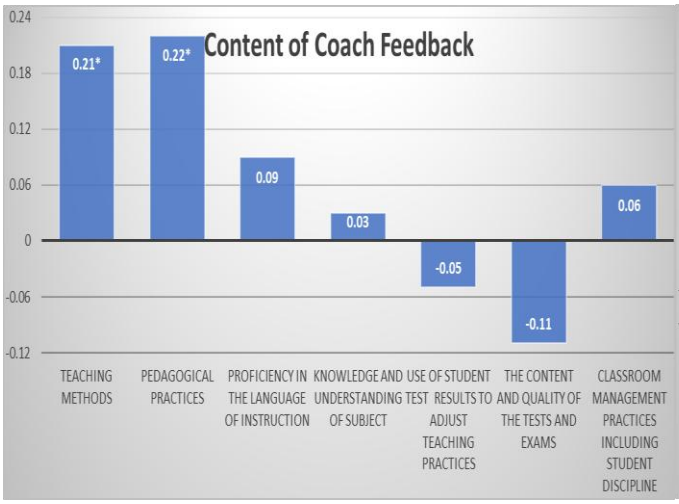


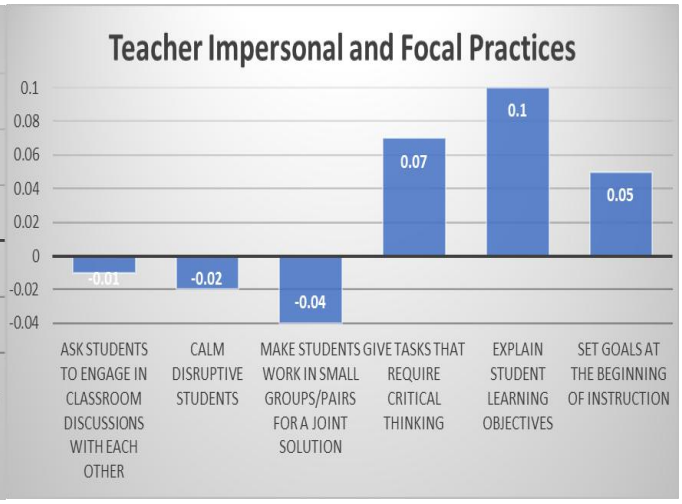
130. The impact of the new coaching model on the perceived necessity of debriefing sessions by teachers was positive and significant. This finding is particularly interesting as it sheds light on three key factors. Firstly, before implementing the intervention, teachers felt they were receiving fewer debriefs than they deemed necessary, underscoring the importance they placed on receiving feedback and support. Secondly, it highlights the significant value that teachers attribute to coaching as a means of professional development. Lastly, the new coaching model proved effective in helping teachers improve their teaching methods and pedagogical practices, further reinforcing the benefits of the intervention.

131. Moreover, the intervention had a positive and significant impact on the likelihood that teachers reported the coaches' feedback as helpful in seeing improvements in student results. This increase of 15 percentage points indicates that the coaching model played a valuable role in assisting teachers in fostering positive outcomes for their students.



132. Furthermore, although not statistically significant, positive impacts were observed in the improvements of pedagogical practices and the identification of areas for improvement. Specifically, the new coaching model significantly increased the likelihood that coaches focused on teaching methods and pedagogical practices during feedback sessions with teachers, with estimated impacts of 22 and 21 percentage points, respectively. This suggests that the coaching model facilitated a more targeted and effective approach to supporting teachers in refining their instructional techniques and addressing areas that required improvement.





133. It is important to note, however, that these estimated impacts, while demonstrating positive trends, did not reach statistical significance. Therefore, cautious interpretation of these findings is warranted. Nonetheless, the overall results suggest that the new coaching model holds the potential for enhancing teaching practices and fostering positive student outcomes, even if the statistical evidence falls short of significance.

3. Changes observed

134. When examining the intervention impacts on lesson activities, a significant difference is found in the aspect of engaging in whole group discussions. The estimation reveals that lessons in the treatment group were 18 percentage points more likely to involve whole group discussions compared to the control group.

135. Furthermore, the impact estimates consistently align with the differences observed at the endline assessment, with effect sizes ranging from 0.24 to 0.55 out of 4, all of which are statistically significant across both segments of lesson observation. Notably, the largest effect sizes were observed in the intellectual challenge, critical thinking, and the level of questioning, with effect sizes of 0.55, 0.44, and 0.45, respectively. These findings indicate that the coaching model has a substantial impact on fostering greater intellectual engagement, critical thinking skills, and increased opportunities for classroom discourse.

136. It is important to note that the effect on student-to-student dialogue, while positive in magnitude, did not reach statistical significance. Despite this, the overall pattern of results suggests that the new coaching model is indeed leading to meaningful changes and shifts in classroom practices among the treatment group teachers, consistent with the underlying theoretical framework of the coaching model being implemented.

137. In summary, the analysis indicates a significant impact of the coaching model on promoting whole-group discussions in lessons. The effect sizes in various dimensions of classroom practice further support the notion that the coaching model is effectively translating into tangible changes in teaching approaches and classroom dynamics, as intended by the theoretical framework underlying the intervention.

Limitations of the Study

138. The evaluation of the new coaching model using a randomized controlled trial design has certain limitations that should be acknowledged. One major limitation is the absence of an examination of the intervention's impact on student learning outcomes, which serves as the ultimate desired outcome of the intervention. While the study provides valuable insights into the effects of the coaching model on coaches and teachers, including their uptake of the model and perceived effectiveness, it does not allow for an assessment of the direct impact on student academic performance. This restricts our ability to evaluate the overall effectiveness of the coaching model in improving student learning outcomes.

139. Another limitation that may have influenced the implementation fidelity and efficacy of the intervention is the delays in implementation and the extended timeframe of the intervention. While the randomized trial design enables a comparison between the intervention group and a counterfactual representing the state of the educational system without the intervention, external factors such as the pandemic and economic crisis could have interacted with the intervention effects in a non-additive manner. It is important to recognize that these external shocks may have affected all participants equally but could have potentially influenced the estimated intervention impact. Therefore, it is challenging to determine whether the observed effects would remain consistent under more stable or "normal" conditions. Conducting future replications of this evaluation under different environmental conditions could help validate the results obtained in this study.

140. Furthermore, there are opportunities for future research to enhance our understanding of the coaching model. Supplementing the quantitative analysis with qualitative interviews of coaches, teachers, and students would provide additional insights into their experiences and perspectives regarding the coaching model's implementation and impact. Additionally, extending the analytic timeframe would enable the identification of any long-term impacts that the coaching model may have on teachers' practices and student outcomes.

In conclusion, while the randomized controlled trial design offers valuable insights into the effects of the coaching model on coaches and teachers, limitations regarding the assessment of student learning outcomes, external shocks, and the potential for qualitative and long-term analysis should be considered. Addressing these limitations and conducting further research can contribute to a more comprehensive understanding of the coaching model's effectiveness and its potential for improving educational outcomes.

VII. Report Recommendations

141. The evaluation of the Lebanon Coaching model provides evidence that certain aspects of the model have the potential to enhance teachers' approaches to lessons and teaching in key subject areas for students in grades 4-8. These findings suggest that the coaching model can be a valuable tool in addressing the challenges faced by Lebanon's education system, which are often complex and influenced by external factors such as political and socio-economic issues.

142. It is worth noting that despite these challenges, Lebanon benefits from strong community support for education and a high value placed on the education of children by parents. This supportive environment creates opportunities for the successful implementation and scalability of the coaching model.

143. **Consistent and re-enforced teacher support from coaches:** Based on the findings of the evaluation, it is recommended to establish a policy that includes a minimum number of visits per coach to each teacher, accompanied by a comprehensive support framework. This policy would ensure that teachers receive regular and meaningful support from coaches, which has been shown to have a positive impact on teaching practices. Implementing such a policy would require careful planning and coordination to address the specific needs and contexts of the Lebanese education system. By providing consistent support to teachers through the coaching model, the aim is to further enhance the quality of education and improve student outcomes.

144. One measure of ensuring consistent support even during challenging times is to introduce flexible coaching models incorporating hybrid (online and in-person) options. The implementation of the Lebanon coaching model faced serious disruptions due to COVID-19-related school closures, teacher strikes, and socio-political instability. These challenges significantly reduced the number of coaching visits and debrief sessions, limiting the effectiveness of the intervention. Although the model was adapted for partial virtual delivery, this shift was reactive and lacked a comprehensive design for hybrid implementation. To improve resilience and scalability, future coaching models should proactively incorporate hybrid delivery modes. This would allow coaching to continue uninterrupted during emergencies and enable more flexible scheduling for teachers and coaches. Structured virtual modules, complemented by periodic in-person visits, can ensure continuity while reducing logistical burdens.

145. **Institutionalized feedback loops where teachers evaluate the coaching experience to inform continuous model refinement.** The evaluation revealed a disconnect between coaches' and teachers' perceptions of the coaching experience, particularly regarding whether teachers' views were incorporated into feedback. While 86% of coaches reported integrating teacher input into their feedback, only 68% of teachers agreed. This suggests a need to systematically capture and respond to teacher feedback. Institutionalizing feedback loops—through anonymous surveys, focus groups, or structured reflection tools—would enable ongoing improvements to the coaching content and process. This participatory approach not only enhances trust and accountability but also ensures that coaching is responsive to teachers' evolving needs and school realities.

146. Furthermore, while the intervention included observations using PLATO and DOPS rubrics, these were primarily for evaluation purposes rather than routine quality assurance. The findings highlighted that that coaches often had multiple competing responsibilities, and the quality or consistency of coaching delivery was not systematically monitored. Introducing a formal coaching quality assurance mechanism—such as periodic external observers or structured peer review systems—would help maintain high standards and at the same time provide developmental feedback to coaches themselves, encouraging continuous improvement in their facilitation, communication, and feedback techniques.

147. **Coaching models differentiated by subject and school context to account for variation in needs and instructional styles.** The pilot coaching model targeted English, French, and Biology teachers across grades 4–8, yet the evaluation did not differentiate findings or strategies by subject. However, pedagogical practices, classroom dynamics, and instructional challenges can vary significantly by subject and context (e.g., inclusive schools, large class sizes, rural vs. urban settings). A one-size-fits-all coaching model may overlook these nuances. Tailoring coaching strategies to specific disciplines and school environments can make the support more relevant and actionable. For example, science teaching may benefit from more emphasis on hands-on experiments and inquiry-based learning, while language instruction may require more focus on dialogue and comprehension strategies.

148. **Coaching practices that are integrated into teacher career progression frameworks to incentivize participation and professional growth.** Despite overall satisfaction with the coaching experience, teachers' participation was sometimes hindered by misconceptions—some believed they were being coached because of poor performance. Moreover, teachers' workloads and low morale due to systemic challenges affected their engagement. Embedding coaching within a broader professional development and career advancement framework would help reposition it as a valuable opportunity rather than a remedial intervention. Recognizing coaching participation and outcomes in teacher evaluations, promotions, or access to further training can foster greater motivation and legitimacy. This approach also aligns with the Education Reform Plan's focus on strengthening the teaching profession and professionalizing teacher support systems.

149. **A support network or circle for teachers to allow sharing of good practices:** the evaluation provides evidence of the potential effectiveness and feasibility of the reformed Lebanon Coaching model. Leveraging the community's support for education and implementing a policy that establishes minimum coach-to-teacher visits, along with a robust support framework, can contribute to the continuous improvement of teaching practices and the overall education system in Lebanon.

FIGURES

TABLES

ANNEX 1- Detailed of Impact Evaluation Results: Baseline vs Endline

Table 4: Coaching practice

	Baseline			Endline		
Coaching Practice	Control	Treatment	Difference	Control	Treatment	Difference
# Of coaching visits last year (per teacher)	32.97	28.81	4.16	5.78	6.16	-0.38
Observed class in > 80% of visits	0.58	0.64	-0.06	0.67	0.80	-0.14
# Of teachers coached	47.23	51.20	-3.98	36.40	26.63	9.77*
Meeting duration after class observation	28.15	31.47	-3.32	27.16	26.06	1.10
Meeting duration when no class observation	31.68	26.73	4.95	26.18	22.43	3.76

Table 5: Most focused area (time wise) while providing feedback on teaching practice

	Baseline			Endline		
	Control	Treatment	Diff	Control	Treatment	Diff
Classroom management	0.20	0.26	-0.06	0.23	0.39	-0.17
Pedagogical practices	0.37	0.38	-0.01	0.43	0.24	0.18*
Content and quality of exams	0.43	0.38	0.05	0.40	0.29	0.11
Using assessments to adjust instruction	0.31	0.21	0.10	0.26	0.18	0.08
Knowledge and understanding of subject	0.32	0.26	0.06	0.28	0.22	0.06
Proficiency in the Language of Instruction	0.34	0.36	-0.01	0.41	0.34	0.07

Table 6: Coaching process and coach self-efficacy

	Baseline			Endline		
Coaching Process	Control	Treatment	Difference	Control	Treatment	Difference
I felt that I was knowledgeable when it came to working with teachers	0.78	0.53	0.24*	0.76	0.82	-0.06
I was able to communicate effectively with teachers	0.95	0.89	0.07	0.82	0.93	-0.11*
I was able to deliver support and technical assistance to teachers	0.94	0.86	0.08	0.80	0.91	-0.11
I was able to provide support that matched the needs of the teachers	0.92	0.82	0.10	0.80	0.84	-0.04
The feedback I gave teachers was practical and useful	0.94	0.80	0.14*	0.82	0.93	-0.11*

Table 7: Coaches always consider or analyze the following during lesson observations

	Baseline			Endline		
	Control	Treatment	Diff	Control	Treatment	Diff
Teacher instruction	0.21	0.18	0.03	0.22	0.24	-0.02
Content and quality of exams	0.35	0.43	-0.08	0.34	0.29	0.05
Classroom management	0.16	0.12	0.04	0.12	0.18	-0.06
Use of technology to aid in instruction	0.14	0.18	-0.04	0.14	0.09	0.05
Student participation in the classroom	0.11	0.14	-0.03	0.08	0.13	-0.05
Use of facilitation moves by the teacher	0.05	0.09	-0.04	0.10	0.09	0.01

Table 8: Coaches' Perspectives on Challenges

	Baseline			Endline		
Coaching Barriers	Control	Treatment	Difference	Control	Treatment	Difference
Insufficient support from principals	0.33	0.45	-0.12	0.29	0.23	0.06
Large class sizes	0.75	0.64	0.11	0.60	0.79	-0.19*
Teachers with inadequate teaching skills	0.81	0.81	0.00	0.85	0.86	-0.01
Resistance to coaching from teachers	0.24	0.36	-0.12	0.29	0.26	0.04
School assignment workload	0.47	0.46	0.00	0.40	0.35	0.05
Other responsibilities at assigned school	0.22	0.31	-0.09	0.21	0.21	-0.00
Other responsibilities within DOPS	0.14	0.05	0.10	0.10	0.14	-0.04
Insufficient time to debrief with teacher after observation	0.76	0.67	0.09	0.58	0.58	0.00
Inadequate professional development opportunities	0.60	0.51	0.09	0.38	0.40	-0.02

Table 9: Coaches' Job satisfaction

	Baseline			Endline		
Job Satisfaction	Control	Treatment	Difference	Control	Treatment	Difference
If I could decide again I would still coach	0.83	0.85	-0.03	0.79	0.84	-0.05
I am satisfied with my job as coach	0.86	0.98	-0.11*	0.77	0.81	-0.04
I am satisfied with coach salary	0.28	0.24	0.04	0.02	0.02	-0.00
Satisfied with support from coaching center coordinator	0.77	0.88	-0.11	0.79	0.91	-0.12
Satisfied with support from subject coordinator	0.83	0.90	-0.08	0.88	0.93	-0.06

Table 10: Teachers self-reporting the frequency of their actions

	Baseline			Endline		
Frequency = Always	Control	Treatment	Difference	Control	Treatment	Difference
Set goals at the beginning of instruction	0.63	0.66	-0.04	0.68	0.76	-0.09*

Explain student learning objectives	0.53	0.51	0.02	0.47	0.55	-0.08
Give tasks that require critical thinking	0.27	0.30	-0.03	0.24	0.33	-0.09*
Make students work in small groups/pairs for a joint solution	0.21	0.23	-0.02	0.14	0.11	0.02
Calm disruptive students	0.50	0.44	0.06	0.43	0.35	0.08
Ask students to engage in classroom discussions with each other	0.25	0.31	-0.06	0.24	0.28	-0.04

Table 11: Teacher coaching

	Baseline			Endline		
	Control	Treatment	Difference	Control	Treatment	Difference
Number of coach observations during the current academic year	2.00	2.30	-0.30*	1.16	4.31	-3.14*
Number of post observation coaching debriefs perceived necessary	3.43	3.40	0.03	3.62	4.32	-0.70*
Number of debriefs after classroom observations	2.22	2.46	-0.24	1.95	4.28	-2.33*

Table 12: Coach feedback

	Baseline			Endline		
Content of coach feedback	Control	Treatment	Difference	Control	Treatment	Difference
Classroom management practices including student discipline	0.18	0.19	-0.01	0.08	0.15	-0.07*
The content and quality of the tests and exams	0.85	0.82	0.03	0.64	0.50	0.14*
Use of student test results to adjust teaching practices	0.25	0.31	-0.05	0.21	0.21	0.00
Knowledge and understanding of the subject	0.46	0.44	0.02	0.34	0.35	-0.01
Proficiency in the language of instruction	0.32	0.29	0.02	0.11	0.18	-0.07
Pedagogical practices[7]	0.66	0.65	0.00	0.52	0.73	-0.21*
Teaching methods[8]	0.52	0.52	0.00	0.50	0.71	-0.21*

Table 13: Perceived benefit of the coaches' feedback

	Baseline			Endline		
Coaching feedback helpfulness	Control	Treatment	Difference	Control	Treatment	Difference

Feedback helpful in identifying areas of improvement	0.39	0.46	-0.07	0.33	0.53	-0.21*
Feedback helpful in improving pedagogical practices	0.36	0.41	-0.06	0.31	0.52	-0.21*
Feedback helpful in seeing improvements in student results	0.34	0.39	-0.05	0.28	0.50	-0.22*

Table 14: Coaching session and follow-up

	Baseline			Endline		
Coaching sessions and follow up	Control	Treatment	Difference	Control	Treatment	Difference
Did a DOPS coach observe your classroom in the past?	0.58	0.34	0.24*	0.46	0.8	-0.34
Time with coach in coaching session was enough to have a constructive disc	0.69	0.67	0.02	0.72	0.83	-0.11*
Able to reflect on teaching practice with the coach during coaching session	0.8	0.75	0.05	0.82	0.94	-0.13*
Coach provided encouragement and motivation in the coaching session	0.9	0.85	0.05	0.89	0.95	-0.06*
During the coaching sessions the coach supported my instruction	0.82	0.85	-0.03	0.8	0.9	-0.10*
The coach followed up with me to further support my instruction	0.68	0.65	0.03	0.68	0.94	-0.26*
To improve my practice the coach and I agreed on what goals to focus on	0.76	0.75	0.01	0.77	0.93	-0.16*
The coach and I trust one another	0.86	0.87	-0.01	0.83	0.95	-0.12*
The coach was easy to meet and work with	0.92	0.86	0.05*	0.9	0.93	-0.03
The coach and I worked together collaboratively	0.85	0.88	-0.02	0.87	0.96	-0.09*
The coach showed a sincere desire to understand and improve my classroom	0.88	0.88	0.01	0.85	0.94	-0.08*
The coach incorporated my views	0.66	0.71	-0.04	0.78	0.84	-0.06

Table 15: Classroom observation

	Baseline			Endline		
Segment 1	Control	Treatment	Difference	Control	Treatment	Difference
Teacher Lecture	0.76	0.75	0.01	0.81	0.87	-0.06
Short Student Responses to Teacher Questions	0.86	0.85	0.01	0.81	0.82	-0.01
Small Group or Partner Discussions	0.15	0.09	0.06*	0.08	0.04	0.03

Whole Group Discussion	0.28	0.31	-0.03	0.09	0.29	-0.20*
Student Presentations	0.05	0.07	-0.02	0.08	0.05	0.03
Student Independent Work	0.20	0.27	-0.06*	0.24	0.21	0.03
Teacher-led small Group	0.06	0.05	0.01	0.02	0.02	0.00
Segment 2						
Teacher Lecture	0.69	0.67	0.02	0.79	0.82	-0.03
Short Student Responses to Teacher Questions	0.83	0.81	0.02	0.77	0.78	-0.02
Small Group or Partner Discussions	0.19	0.21	-0.02	0.16	0.10	0.06
Whole Group Discussion	0.29	0.29	0.00	0.12	0.31	-0.20*
Student Presentations	0.08	0.08	0.00	0.12	0.07	0.06*
Student Independent Work	0.34	0.32	0.02	0.38	0.27	0.11*
Teacher-led small Group	0.08	0.07	0.02	0.02	0.03	0.00

ANNEX 2- Comparable Coaching Program Reviewed

As part of the process by which we generated our recommendation, we reviewed several coaching programs—implemented in countries around the world—that have an evidence base of effectiveness. A full report on these coaching systems can be found in Appendix A of this report. However, we summarize each system here along with its primary components.

1.) **My Teaching Partner (United States)**
A coaching intervention developed at the University of Virginia Curry School of Education, based on Classroom Assessment Scoring System (CLASS) observation protocol. The program has been implemented in the US in classrooms from preschool to secondary level. The coaching intervention uses the CLASS instructional framework to focus on improving interactions between students and teachers, usually through videos of the teacher's classroom. Each Coach and Teacher participate in six to nine coaching cycles in an academic school year.

2.) **Coaching system in rural multi-grade primary schools (Peru)**
A national coaching system implemented in Peru starting in 2010. Experienced teachers serve as coaches to provide pedagogical support and regular feedback to primary school teachers in rural areas. Coaches making nine monthly visits to teachers; each visit lasts a full day, with five hours of observation and three hours of reflection, discussion, feedback, and practice. Teachers also participate in eight "micro-workshops" during the school year, provided by coaches.

3.) **Virtual coaching using practical classroom techniques (Brazil)**
This program focused on supporting high school teachers with lesson planning, classroom management and student engagement. Information from baseline observations were used to motivate teachers to change; coaches worked with teachers on classroom techniques focused on maximizing student engagement and learning and provided teachers with access to video examples from Brazilian classrooms. Each coach held four, two-hour coaching sessions during the school year with each of the ~30 schools under his/her supervision.

4.) **Comparison of On-site and Virtual Coaching Models (South Africa)**
A program for public schools to deliver structured pedagogy of English as an additional language in the first three grades of primary school through two coaching modalities: (1) on-site coaching and (2) virtual coaching. All teachers received lesson plans and educational materials, but those in the on-site group received printed materials and those in the virtual group received materials on a tablet. Teachers from both groups attend two days of center-based training—those in the virtual coaching group spend a third day learning how to use the tablet.

5.) **Primary Math and Reading (PRIMR) program (Kenya)**
The Primary Math and Reading (PRIMR) program in Kenya focused on improving literacy and numeracy outcomes for students in grades one and two. Although the program has many features, a key component is teacher coaching. Coaches receive 15 days of training on math and literacy pedagogies, and teachers receive 10 days of training by coaches. Teachers receive monthly coaching to support teachers' use of the pedagogical practices, in addition to student textbooks and teacher guides.

1. United States: My Teaching Partner

What is it?

A coaching intervention developed at the University of Virginia Curry School of Education, based on Classroom Assessment Scoring System (CLASS) observation protocol. The program has been implemented in the US in classrooms from preschool to secondary level.

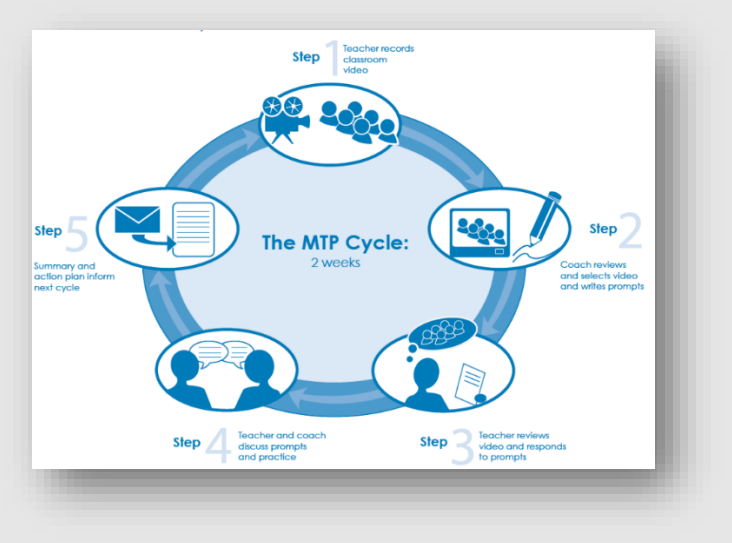
How does it work?

This intervention is based on improving practices in the CLASS instructional framework and is usually video-based. Each Coach and Teacher participate in six to nine coaching cycles in an academic school year.

The **Coaching Cycle** consists of:

1) The teacher videos themselves teaching and sends it to their coach.
2) The coach selects a 1–2-minute segment from the teacher's video about a CLASS dimension they are working on & poses questions to the teacher.
3) The teacher watches the video segment & answers the coach's questions.
4) The coach and teacher discuss the segment and instructional practice being worked on.
5) The coach sends a summary & action plan for moving forward to the teacher. (This helps inform the next coaching cycle.)

Each coaching cycles last approximately two weeks. During each two-week cycle, coaches spend 4–5 hours on activities per teacher. Teachers spend 2–3 hours every two weeks engaging in ongoing coaching cycles during the program.



Coaches use a **specific set of prompts to tailor feedback** during coaching cycles. (Early et al. 2017)

1) "Nice Work". Using this prompt, coaches focus on what the teacher did well in the video they submitted.
2) "Consider This". Using this prompt, the coach helps the teachers develop classroom observation skills, focusing specifically on how teachers' words and actions impact students.
3) "Making the Most" Using this prompt, the coaches' comments focus on the *instructional support* domain of CLASS

The topics coaches work with teachers on moves through the CLASS domains; starting with Emotional Support, then Classroom Organization, and finally Instructional Support. The amount of time spent on each domain varies based on specific teacher's needs.

Professional Development Supports

- Teachers in My Teaching Partner have access to an online video library that has over 400 short clips displaying best practices.
- There is also a college course that teachers in the MTP program enroll in. The three-credit college course focuses on improving teachers' knowledge of effective interactions, their skills in identifying effective interactions, and applying those skills to their own classrooms.

What evidence is there of effectiveness?

The average student in a classroom where the teacher is enrolled in MTP sees a 9% increase in their test scores in one year. The program is especially effective in classrooms that serve higher proportions of students in poverty. This program is best researched in preschool classrooms, where studies have found that participating in MTP increased teachers' ability to read and respond to student cues, to use a variety of instructional methods, and to stimulate their students' language development. In Secondary classrooms, students of teachers in MTP coaching saw an increase in their test scores from the 50th to 59th percentile.

2. Peru: Coaching system in rural multi-grade primary schools

What is it?

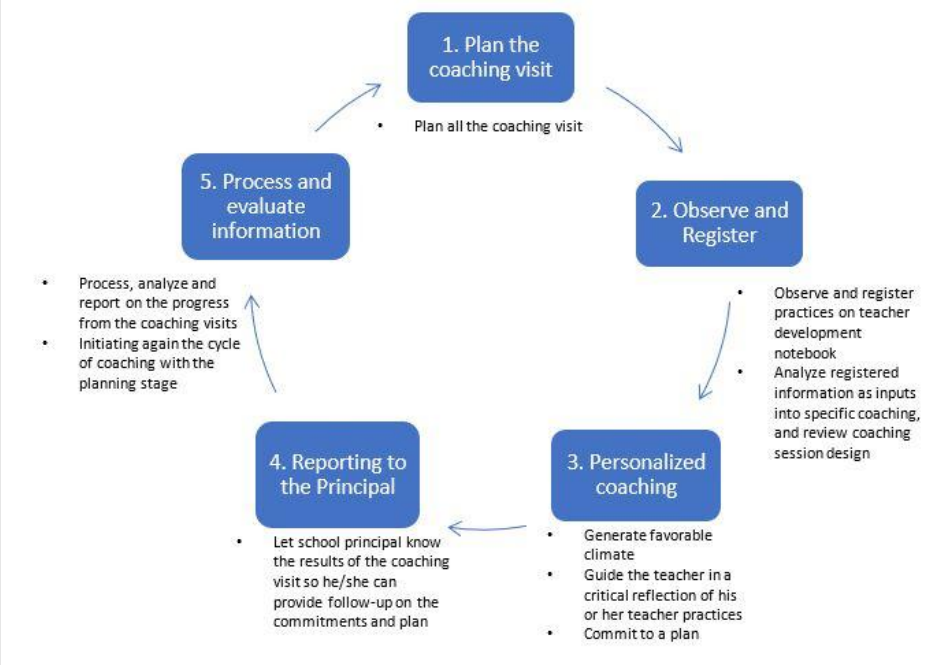
A national coaching system implemented in Peru starting in 2010. While the coaching program was designed centrally, local school boards hire experienced teachers to serve as coaches to provide pedagogical support and regular feedback throughout the year to multi-grade primary school teachers in rural areas.

How does it work?

The program is implemented nationally and has several key characteristics:

1) **Coaching visits:** Coaches making nine monthly visits to the teachers they coach. Each coaching visit lasts a full day, with five hours of observation and three hours of reflection, discussion, feedback, and practice of new routines. The first visit is diagnostic in nature and focuses on figuring out where the teacher's skills are. The coach uses a rubric to assess the teacher and with this data draws up a coaching plan for the school year. Coaches write a diagnostic report about the teacher and then present it to their colleagues. The report includes general data, goals for the year, an analysis of the basic profile (areas for improvement), results of the individual teacher compared to their peers, general conclusions, difficulties encountered during the diagnostic visit, and copies of the data collected. Then, during the seven middle visits coaches observe the teacher and provide training and feedback.

   Each coaching cycle follows five phases:



2) **Group Micro-workshops:** coaches also provide teachers with eight group "micro-workshops" throughout the year for all teachers they coach (each one is four hours long). These workshops are a space to discuss pedagogical practices and for teachers to exchange ideas. Coaches plan them based on a common problem they noticed during their observations with their teachers.
3) **Supervised by pedagogical specialists:** Coaches are trained and supervised by pedagogical specialists who accompany them on their coaching visits twice a year.
4) **Emphasis on Pedagogical Practices:** Coaches don't focus with teachers on content knowledge of the material but rather on improving their pedagogical practices and their ability to auto-reflect and analysis. Coaches also support the skills of teachers "related to understanding the student in her context, curricular planning, guiding the learning, school environment, and evaluating student learning."

What evidence is there of effectiveness?

Using experimental and quasi-experimental designs Majerowicz and Montero (2018) confirm that this coaching system improved student learning outcomes considerably. Specifically, the program had positive effects in the range of between 0.24 and 0.34 standard deviations on second grade standardized test scores on second grade. The results suggest that students of low, medium, and high initial ability benefit from the program. Students of younger teachers have stronger impacts on their learning, which suggests that older teachers may be less amenable to change.

3. Brazil: Virtual coaching using practical classroom techniques

What is it?

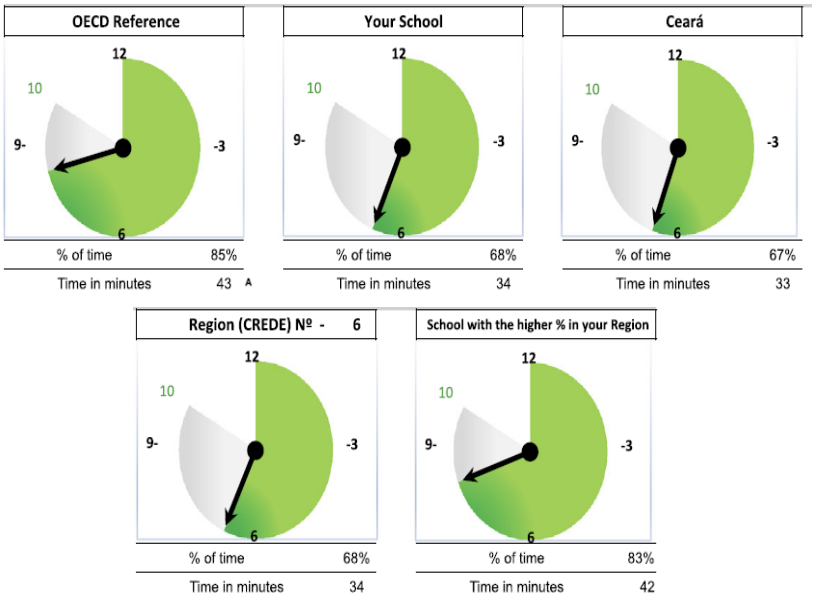
A 9-month training and coaching program developed by a local organization in the state of Ceará, Brazil that assisted high schools by supporting teachers on lesson planning, classroom management and student engagement. The intervention observed teachers using a classroom observation protocol at the beginning of the year, and then had coaches work with teachers on classroom techniques focused on maximizing the engagement and learning of all students. High-school teachers were generally enthusiastic about these techniques, stating that their pre-service training failed to offer practical techniques.

How does it work?

The intervention had four parts:

1) Classroom observation and feedback: At the beginning of the year, and after being observed using a classroom observation protocol, teachers were provided with a short infographic bulletin that included good practice benchmarks on instructional time use. This information shock for teachers was intended to generate demand for coaching support and improvement in teaching practices.

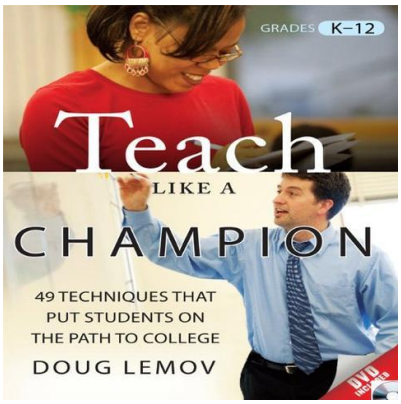
Classroom time on instruction: The good practice benchmark is a well-prepared teacher that can devote 85% of classroom time (out of a 50-minute lesson) to instructional activities.



2) Self-help materials: Teachers and coaches (pedagogical coordinators) were provided with a book (*Teach Like a Champion* by Doug Lemov) covering practical classroom techniques focused on maximizing the engagement and learning of all students.

Categories of the classroom techniques

- Setting high academic expectations
- Planning that ensures academic achievement
- Structuring and delivering lessons
- Engaging students in the lessons
- Creating a strong classroom culture
- Setting and maintaining high behavioral expectations
- Building character and trust



Additionally, school teaching personnel had access to video examples, a website with good practice video examples from Brazilian classrooms, and a system to upload their own videos of good teaching practices and good coaching practices from their school.

3) Initial Training for Coaches: The local organization that developed the program also trained the coaches (pedagogical coordinators) on 3-day long sessions, which covered: i) the goals of the program; ii) the *Teach Like a Champion* book; iii) how to work with the firm's master coaches during the year; iv) how to help teachers experiment with the book's techniques; and v) how to hold individual coaching sessions with teachers and provide feedback. Coaches were also asked to upload their own coaching sessions with teachers so master coaches could review them.
4) Coaching via Skype: Each ELOS master coach was responsible for holding four two-hour coaching sessions during the school year with each of the 30 or so schools under his supervision. Master coaches reviewed coaching videos from coaches and provided feedback as well as helped with planning.

What evidence is there of effectiveness?

The results of a randomized controlled trial (Bruns, Costa and Cunha 2018) shows that over a single school year the program increased teachers' time on instruction and student engagement. It also produced statistically significant gains in student learning on both the national secondary school exit exam (0.04 to 0.06 standard deviations) as well as the state's assessment (0.05 to 0.09 standard deviations). The impact of the program was stronger among schools that implemented it with greater fidelity. The fact that coaching was done via Skype kept the cost of the program low.

4. South Africa: Comparison of On-site and Virtual Coaching Models

What is it?

A program for public schools to deliver structured pedagogy to learn English as an additional language in the first three grades of primary school through two modalities of coaching: an on-site coaching intervention and a virtual coaching intervention. The program is currently being implemented in 180 public schools in low-income communities in rural areas of the Mpumalanga province in South Africa.

How does it work?

All teachers in the intervention schools receive lesson plans and educational materials, but those with on-site coaches get printed materials and those with virtual coaches get them on an application in a tablet. Teachers from both coaching interventions attend center-based training for 2 days—those with virtual coaches spend a third day learning how to use the tablet.

On-site coaches are reading experts who work face to face with teachers to support mastery of instructional practice, teacher content knowledge, and professional confidence. There are approximately 32 teachers per coach, and the coach visits them roughly every 3 weeks. Coaching sessions entail lesson demonstrations, classroom observation, discussions on how to assess their practice and student outcomes. While they perform classroom observations, coaches provide feedback on the lesson and model the practice the teacher is trying to master. Coaches also provide emotional support and motivation.

Virtual coaches coach teachers via one-on-one or group text messaging and phone calls on WhatsApp. Every two weeks virtual coaches follow up on training focus areas with individual teachers to improve instructional practice and to check they are covering the curriculum. Motivational support is provided via group message once a week. Virtual coaches also share short videos on the most difficult teaching areas for teachers.

Intervention Design

	On-site coaching	Virtual coaching
Lesson plans	Paper-based	Electronic: tablet
Learning and teaching support materials	Paper-based	Paper-based Include videos, sound clips, and photos of example writing
Coaching	In classroom visit Once every three weeks	Via cell phone calls and instant messaging Once every two weeks
Training	Initial training: Two-day block training Quarterly training: One day at the start of each term Needs-based cluster training	Initial training: Three-day block training Quarterly training: One day at the start of each term Additional videos developed as required

Source: Table 1 from Kotze, Fleisch and Taylor (2019)

What evidence is there of effectiveness?

Based on the results of a randomized controlled trial (RCT) and after one year of program implementation, the evidence suggests that both on-site and virtual coaching are equally effective, as teachers under both interventions have students with equally improved English oral proficiency (Kotze, Fleisch and Taylor 2019). These students perform considerably better than those students in the control group. The costs per student for each coaching model are similar. This study builds on previous studies that use structured pedagogy and coaching (Fleisch et al., 2016; Fleisch et al. 2017; and Cilliers et al, 2019).

5. Kenya: Primary Math and Reading (PRIMR) program

What is it?

The Primary Math and Reading (PRIMR) program in Kenya (funded by USAID and DFID) focused on improving literacy and numeracy outcomes for students in grades 1 and 2. Although the program had many features, such as teacher training, provision of textbooks and teachers' guides in some iterations, one of the key parts is a teacher coaching component.

How does it work?

The PRIMR program has been implemented since the early 2010s and based on the results of various studies and evaluations it has been modified several times. A comprehensive version of PRIMR included the following:

1) **Coaches allowed to prioritize coaching:** Tutors who were responsible to provide coaching were overburdened with administrative responsibilities that took more than 60% of their time. Under the program, these tutors, which were mostly experienced and recognized teachers, were relieved of many of their administrative duties so they could coach more with a focus on instruction. The tutors were renamed curriculum support officers and conducted coaching visits every day and were responsible for about 15 to 30 schools.
2) **Training for Coaches:** 15 days of training for pedagogical coaches on the reading and math content and pedagogy mostly at the beginning of the year.
3) **Training for Teachers by the Coaches:** 10 days of training for teachers on the reading and math content and pedagogy mostly at the beginning of the year.
4) **Coaching:** Regular coaching visits to teachers about once a month to reinforce pedagogy, including reflection, discussions, application of a classroom observation tool, modelling of the new practices by the coach either in the classroom or after it, etc. Coaches also assessed math and reading knowledge of selected students to track progress.
5) **Transport Stipends:** Coaches were provided with modest transport stipends, so they could reach their assigned schools and conduct coaching visits with their teachers.
6) **Student Textbooks:** Provision of textbooks with new curriculum on math and reading for all students
7) **Teachers' Guides:** Provision of structured teachers' guides with detailed information on how to carry out the lessons.

What evidence is there of effectiveness?

Several studies of PRIMR implemented in different regions or with variations had showed impacts on teacher instructional practices and student learning outcomes (Piper, Zuilkowski, and Mugenda 2014; Piper and Zuilkowski 2015). To understand what major components of the program were driving the impact that had been observed in previous studies, and, in turn, to inform future decision-making and planning, Piper et al. 2018 conducted one study with three versions of the program. The setup for that study was the following:

It was carried out in rural zones in the counties of Bungoma and Machakos and included randomized assignment of zones to three different versions (or treatment groups) of the program and a business-as-usual control group to be able to compare results using a randomized controlled trial.

The three treatment groups were:

Treatment 1: Teacher Training (10 days per year) + Coaching (instructional support every month). Teachers supported to use more effectively existing classroom materials

Treatment 2: Same as (1) + Revised Textbooks for every student. Teachers encouraged to develop own lesson plans ensuring alignment with revised textbooks

Treatment 3: Same as (2) + Teachers' Guides. Teachers' guides included partially scripted lessons for 150 days fully articulated with revised textbooks

The intervention was intended to test at a medium scale and through the government's regular education system with regular employees providing support and professional development to teachers.

However, the treatment group 2 had the greatest impact on student learning outcomes. Treatment 3, although the most expensive, was also the most cost-effective because it also raised learning outcomes the most.

REFERENCES

Béteille, Tara, and David Evans. 2021. Successful Teachers, Successful Students: Recruiting And Supporting Society's Most Crucial Profession.

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 (June): 214–50. https://doi.org/10.1016/j.econedurev.2018.03.003.

Cilliers, Jacobus, Brahm Fleisch, Cas Prinsloo, Vijay Reddy, and Stephen Taylor. 2019. "How to Improve Teaching Practice? Experimental Comparison of Centralized Training and in-Classroom Coaching." *Journal of Human Resources*.

Early, Diane M., Kelly L. Maxwell, Bentley D. Ponder, and Yi Pan. 2017. "Improving Teacher-Child Interactions: A Randomized Controlled Trial of Making the Most of Classroom Interactions and My Teach Effect size comparisons by treatment group, subject, and grade ıary): 57–70

Fleisch, Brahm, Volker Schöer, Gareth Roberts, and Amy Thornton. 2016. "System-Wide Improvement of Early-Grade Mathematics: New Evidence from the Gauteng Primary Language and Mathematics Strategy." *International Journal of Educational Development* 49: 157–174.

Fleisch, Brahm, Stephen Taylor, Volker Schöer, and Thabo Mabogoane. 2017. "Failing to Catch up in Reading in the Middle Years: The Findings of the Impact Evaluation of the Reading Catch-Up Programme in South Africa." *International Journal of Educational Development* 53: 36–47.

Global Education Evidence Advisory Panel (GEEAP). 2023. 2023 Cost-Effective Approaches to Improve Global Learning - What does recent evidence tell us are "Smart Buys" for improving learning in low- and middle-income countries?. https://thedocs.worldbank.org/en/doc/231d98251cf326922518be0cbe306fdc-0200022023/related/GEEAP-Report-Smart-Buys-2023-final.pdf

Kotze, Janeli, Brahm Fleisch, and Stephen Taylor. 2019. "Alternative Forms of Early Grade Instructional Coaching: Emerging Evidence from Field Experiments in South Africa." *International Journal of Educational Development* 66 (April): 203–13. https://doi.org/10.1016/j.ijedudev.2018.09.004.

Majerowicz, Stephanie, and Ricardo Montero. 2018. "Can Teaching Be Taught ? Experimental Evidence from a Teacher Coaching Program in Peru." Harvard University.

Piper, Benjamin, Stephanie Simmons Zuilkowski, Margaret Dubeck, Evelyn Jepkemei, and Simon J. King. 2018. "Identifying the Essential Ingredients to Literacy and Numeracy Improvement: Teacher Professional Development and Coaching, Student Textbooks, and Structured Teachers' Guides." *World Development* 106 (June): 324–36. https://doi.org/10.1016/j.worlddev.2018.01.018.

Piper, Benjamin, and Stephanie Simmons Zuilkowski. 2015. "Teacher Coaching in Kenya: Examining Instructional Support in Public and Nonformal Schools." *Teaching and Teacher Education* 47 (April): 173–83. https://doi.org/10.1016/j.tate.2015.01.001.

Piper, Benjamin, Stephanie Simmons Zuilkowski, and Abel Mugenda. 2014. "Improving Reading Outcomes in Kenya: First-Year Effects of the PRIMR Initiative." *International Journal of Educational Development* 37 (July): 11–21. https://doi.org/10.1016/j.ijedudev.2014.02.006.

Snilstveit, B., Stevenson, J., Menon, R., Phillips, D., Gallagher, E., Geleen, M., & Jobse, H. (2016). The Impact of Education Programmes on Learning and School Participation in Low and Middle-Income Countries: A Systematic Review Summary Report. Systematic Review Summary no. 7, International Initiative for Impact Evaluation (3ie).; Beteille, Tara. & Evans, D.K. (2018). Successful Teachers, Successful Students: Recruiting and Supporting Society's Most Crucial Profession. The World Bank