USAID
FROM THE AMERICAN PEOPLE

Maktabgacha
va maktab
ta'limi vazirligi

ULA and Mathematics Teaching and Learning Materials Uptake Study

Phase 1–2 Report

UZBEKISTAN
EDUCATION
for
EXCELLENCE
PROGRAM

RTI INTERNATIONAL

FLORIDA STATE UNIVERSITY 1851

MISSISSIPPI STATE UNIVERSITY

Uzbekistan Education for Excellence Program

ULA and Mathematics Teaching and Learning Materials Uptake Study
Phase 1–2 Report
Cooperative Agreement No. 72011519CA00004

Submitted to:

Abdugani Bazarov
Agreement Officer's Representative
USAID/Central Asia/Uzbekistan
3 Moyqorghon Street
Tashkent, Uzbekistan 70093
Tel.: (99871) 140 2486
Fax: (99871) 120 6309
abarzarov@usaid.gov

Submitted by:

RTI International
3040 East Cornwallis Road
Research Triangle Park, NC 27709

Dr. Carmen Strigel
Project Manager
cstrigel@rti.org

Authors: Sonia Arias, M.Ed.; Temur Rakhmatov

September 5, 2023

CONTENTS

FIGURES AND TABLES

List of Figures

List of Tables

LIST OF ACRONYMS

EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
FA	Field Assistant
MoPSE	Ministry of Preschool and School Education
Q&A	Question and Answer
RQ	Research Question
RTI	RTI International
SCS	Student-Centered Strategy
STB	Student Textbook
TG	Teacher's Guide
TLM	Teaching and Learning Material
TPD	Teacher Professional Development
ULA	Uzbek Language Arts
USAID	United States Agency for International Development

INTRODUCTION

BACKGROUND

Result 1 of the U.S. Agency for International Development (USAID) Uzbekistan Education for Excellence Program focuses on improved Uzbek Language Arts (ULA) and Mathematics outcomes in grades 1–4. The Program theory of change states that improvements in curriculum products combined with teacher professional development (TPD) will translate into improvements in student achievement over time. To achieve this aim, the Program in Years 2 and 3 completed the following activities as outlined in the Program Description:

- Developed relevant and appropriate student learning standards for ULA and Mathematics.
- Developed scopes and sequences to guide student textbooks (STBs) and teacher guides (TGs) development.
- Developed STBs and TGs.
- Conducted quality assurance review (international and national subject matter specialist review, Product Review Group, Scientific Council Review).

The Program focuses on implementation science to look closely at what is working, how and why, and what effect the changes are having on improving teaching and learning. As such, the Program's pilot approach is designed to study the effectiveness of the TGs and STBs in terms of uptake of new teaching and learning materials (TLMs) and new student-centered strategies (SCSs), the accompanying TPD approach, and the effect on student learning (impact assessment following the December 2021 Early Grade Reading Assessment [EGRA] and Early Grade Mathematics Assessment [EGMA] baseline). The TLM Uptake Study is complemented by the ULA and Mathematics TLM Desk Review, which assessed adherence to best practices in layout, design, and content.

PURPOSE

The goal of this study was to learn whether and how teachers in grades 2 and 4 were using the new ULA and Mathematics TG and STBs and applying the SCSs therein. The study consisted of two phases. The Program used Phase 1 results and ensuing recommendations to inform the development and finalization of materials and ongoing TPD efforts. Phase 2 results served to determine whether teachers became more proficient over time in their application of the SCSs embedded in the TGs.

RESEARCH DESIGN AND QUESTIONS

This was a mixed method, partially longitudinal, study that drew on qualitative and quantitative data and sought to answer the following research questions (RQ)[1]:

[1] The Program modified the RQs from the original Program concept note and Activity Monitoring and Evaluation Plan to better explore the nuances of teacher behavior change in Uzbekistan. The revised RQs provided the Program with the opportunity to (1) conduct a more nuanced desk review, (2) address the Ministry's questions, and (3) better understand teachers' interpretations of the TGs and the SCSs therein.

RQ1: To what extent do the content, instruction, and design of the ULA and Mathematics TGs and STBs support the delivery of lessons and adhere to best practices?[2]

RQ2: How are teachers modifying ULA and Mathematics lessons as they teach with the new TLMs in the classroom?

RQ3: To what degree are ULA and Mathematics teachers applying selected SCSs in the classroom?

CONCEPTUAL FRAMEWORK

RTI's essential guidance and best practices for developing TGs served as the guiding conceptual framework for this ULA and Mathematics TLM Uptake Study.[3] This study applied the framework in several ways: to assess the quality of the TLMs; to determine teachers' modifications to lessons and activities when using TLMs in the classroom to inform the final versions of TLMs; and to track progressive changes over time in teachers' observable behaviors based on four stages of change: Not yet started, Novice, Emerging, and Proficient.[4]

[2] Reference to Desk Review document.

[3] Piper, B., Sitabkhan, Y., Mejía, J., & Betts, K. (2018). *Effectiveness of Teachers' Guides in the Global South: Scripting, Learning Outcomes, and Classroom Utilization.* RTI Press Publication No. OP-0053-1805. Research Triangle Park, NC: RTI Press. https://doi.org/10.3768/rtipress.2018.op.0053.1805

[4] Sitabkhan, Y., Harden, K., & Slade, T. (2022). Teaching by the book: Teacher decision-making while using structured lesson plans. *IARTEM e-journal*, *14*(1), 1-12.

METHODOLOGY

DATA COLLECTION METHODS

Data collection methods **(Table 1)** for this study were informed by the above framework and were implemented as follows:

- **Phase 1 (November 2022)** focused on data collection that informed eventual TLM finalization and ongoing TPD. The Program analyzed the layout and structure of the TLMs and observed how and why teachers modified TG lessons when implementing them in the classroom.5 Teacher classroom observations also looked at the extent to which teachers were applying SCSs from the TGs. During this phase, the Program also conducted a teacher self-administered self-efficacy survey.
- **Phase 2 (May 2023)** focused only on detecting progressive changes made by teachers in the application of SCSs and their correlation with teachers' self-efficacy was an additional strand of investigation.

Table 1. Research Questions, Timing, and Tools

Research Focus	Timing	Tools
TLM content, instruction, and design (RQ1)	November 2022–January 2023	RTI rubrics
Teacher lesson modifications (RQ2)	November 20022	Lesson modifications tool (qualitative)
Teacher application of SCSs (RQ3)	November 20022 and May 2023	Teacher observable behaviors tool (quantitative) and self-efficacy tool (quantitative)

To answer RQ1 in Phase 1, the Program conducted a desk review of the TG and STBs. This consisted of adapting and applying evaluation rubrics to determine the degree to which the Program applied evidence-based best practices in the development of the TLMs. Findings are summarized in this study, and complete findings are available in the final Desk Review Report. [6]

To answer RQ2 and also inform TLM finalization and ongoing TPD activities, in Phase 1, the Program administered a proven qualitative paper-based lesson modification tool to understand how teachers were making sense of the new SCSs within the TGs over time, based on training and increasing familiarity with and confidence in the TLMs. The tool captured changes to lessons made by teachers in the process of teaching, and why they were making these changes.

In both Phases 1 and 2, the Program administered a quantitative electronic teacher classroom observable behaviors tool to answer RQ3. This tool was designed to track whether teachers were applying SCSs from within the TGs over time. Based on findings, the

[5] Reference TLM Uptake Study Briefer Phase 1 here.
[6] Reference TLM Desk Review

Program assigned teachers to the categories of *Not yet started*, *Novice*, *Emerging*, and *Proficient* and determined the extent to which teachers moved from one category to the next. **Table 2** presents the descriptions for each of the proficiency categories. In addition to this observation tool, the Program also administered an electronic self-efficacy tool, which included specific questions to determine teachers' perceived self-confidence in the teaching of the ULA and Mathematics subjects.

Table 2. General Descriptions of Proficiency Categories

Category	Description
Not yet started	The teacher has not yet started to apply the new methodology.
Novice	The teacher is just beginning to apply the new methodology. The teacher may apply some components of the methodology but not consistently.
Emerging	The teacher is regularly applying several elements of the new methodology but is not yet demonstrating consistent proficiency.
Proficient	The teacher expertly applies almost all elements of the new methodology with consistency across the lesson.

TOOL DEVELOPMENT

The Program developed and deployed five tools for this study, and they are described briefly in **Table 3**.

Table 3. Tool Details

Tool title	Data collection platform	Number or type of items in the tool	Type of questions
RTI TLM review rubric for TG	Paper	Structured rubric to determine whether TLMs address MoPSE and evidence-based criteria and whether they are addressed adequately and consistently throughout the TLMs	Qualitative
RTI TLM review rubric for STB	Paper	Same as above	Qualitative
Lesson modifications tool	Paper	Structured observation protocol to capture details of all modifications teachers made to a Mathematics or ULA lesson	Qualitative
Teacher observable behaviors tool	Tangerine tablet application	Mathematics 33 items ULA 55 items	Quantitative
ULA and Mathematics self-efficacy tools	Tangerine tablet application	13 ULA and 13 Mathematics items	Multiple choice, quantitative

The two TLM review rubrics, one for TGs and the other for STBs, were based on RTI's best practice criteria and are described in **Table 4**.

Table 4. Best Practice TLM Review Criteria

TG CONTENT, INSTRUCTION, DESIGN	STB CONTENT, INSTRUCTION, DESIGN
▪ Layout ▪ Student standards and lesson objectives ▪ Cohesion and clarity ▪ Front and back matter ▪ Formatting ▪ One-page layout ▪ Instructional supports ▪ Physical features applied to all TGs	▪ Layout ▪ Scaffolding of skills ▪ Lesson objectives ▪ Formatting ▪ Language ▪ Illustrations ▪ MoPSE life skills ▪ Social-emotional learning ▪ Critical thinking ▪ Creativity ▪ Problem solving ▪ Individual, group, pair work

The Program further strengthened RTI's extensive and proven best practices evaluation rubric for developing TLMs by including gender and inclusive elements and conducting interviews with subject matter experts.[7]

The Program tested the electronic quantitative teacher observable behaviors tool for both subjects as well as the paper-based teacher lesson modifications tool with four teachers, each conducting two ULA and Mathematics lessons (a total of eight lessons). Based on the testing, the Program amended the tools to ensure easier use and understanding by non-subject matter experts who were then trained to collect the data for this study. RTI's Tangerine® data collection app served as the principal tool for all quantitative data collection activities.

The Program tested the ULA and Mathematics self-efficacy tools[8] with trainers and teachers. The Program found that the ULA questions, once translated into Uzbek, were too complicated and difficult for teacher/trainer respondents to grasp. Respondents assessing the Mathematics tool, on the other hand, found all but three Mathematics self-efficacy questions easy to understand. The Program eliminated these three unclear Mathematics

[7] RTI International. (2015). *A guide for strengthening gender equality and inclusiveness in teaching and learning materials*. A report for the EdData II: Data for Education Research and Programming (DERP) Program. Washington, DC: U.S. Agency for International Development.
Bulat, J., Dubeck, M., Green, P., Harden, K., Henny, C., Mattos, M., Pflepsen, A., Robledo, A., & Sitabkhan, Y. (2017). *What we have learned in the past decade: RTI's approach to early grade literacy instruction.* RTI Press Publication No. OP-0039-1702. Research Triangle Park, NC: RTI Press. https://files.eric.ed.gov/fulltext/ED582383.pdf
Nakamura, P., Shone, R., & Saidoshurov, S. (2016). *Teaching and learning material in Tajikistan: How do they align with reading research?* American Institutes for Research (AIR) report on U.S. Agency for International Development (USAID) Quality Reading Project. https://www.air.org/resource/report/teaching-and-learning-materials-tajikistan-how-do-they-align-reading-research

[8] Enochs, L., Smith, P. L., & Huinker, D. (2000). Establishing factorial validity of the mathematics teaching efficacy beliefs instrument. *School Science and Mathematics 100(4)*, 194–202.

questions and adapted the ULA tool to match the revised Mathematics tool. The Program piloted the two tools again until all questions were clear enough for participants to understand and respond to.

The Program trained its subject matter experts, data collectors, and ministry representatives to administer the tools. The Program trained 16 field assistant (FA) data collectors and two Ministry of Preschool and School Education (MoPSE) representatives on the electronic quantitative teacher observable behaviors tool as well as on the paper-based lesson modification qualitative tool. Prior to Phase 2 data collection, the Program conducted a refresher training with the same FAs on the electronic quantitative teacher observable behaviors tool.[9] Both trainings included inter-rater reliability testing through the joint observation, review, and analysis of classroom videos by the observers with over 90% internal consistency with the desired performance, which is considered reliable.

Table 5. Number of Teachers Observed for Quantitative Data Collection, by Region, Grade, and Phase

		Namangan	Sirdaryo	Total	Total
Phase 1	Grade 2	20	19	39	**76**
	Grade 4	20	17	37	
Phase 2	Grade 2	20	19	39	**76**
	Grade 4	20	17	37	

SAMPLING

Quantitative data: Given that the focus of this study was to inform Program adaptation, the Program drew a non-representative sample of 80 teachers from 40 randomly selected Program schools. The Program selected a school sample that presents variability of demographics such as districts and size of schools. Once schools were selected, the Program identified two teachers per school, one from grade 2 and one from grade 4, for a total teacher sample of 80 teachers. However, because some teachers were not available during the data collection process, the Program was only able to collect data from 76 teachers out of the 80 total originally planned. **Table 5** presents a summary of the sample demographics for the quantitative data collection.

The main consideration for this sample size was to accommodate questionnaires with multiple possible responses. Considering that classroom observation data were mainly Yes/No responses to questions of whether teachers demonstrated certain SCSs and desired teaching SCSs behaviors or not, a data collection of 40 teachers per grade (80 total) would likely provide enough precision on individual item responses.

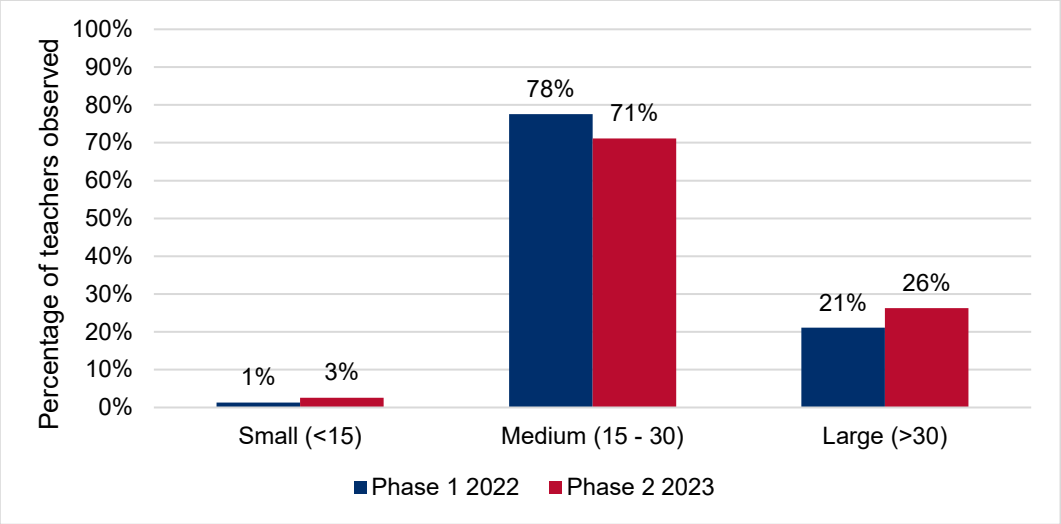
Although the Program selected 40 schools randomly, the selection of teachers within the schools for Phase 1 data collection was based on one convenience sampling criteria: teachers' availability at the time of data collection. The Program observed the same teachers in Phase 1 and Phase 2, and thus observations in each school were determined by teachers' teaching schedule. In Phase 2, the Program adapted the data collection routing plan to meet

[9] No MoPSE representatives were able to attend this training due to the ongoing reorganization process within the ministry.

teachers' teaching schedule to be able to collect data from the same teachers and to not impact teachers' schedules.

The Program also collected information on class sizes to see if average teacher performance in applying SCSs was different across different class sizes. However, as the Program could not determine specific class sizes in advance, it was not able to differentiate the sample prior to visiting schools and thus was not able to control this parameter in the data collection process. Nevertheless, most schools had medium-sized classes (between 15 and 30 students), and around a quarter of the observed classrooms were large (above 30 students). Hardly any classrooms could be categorized as small (fewer than 15 students), **(Figure 1)**.

Figure 1. Proportion of Small, Medium, and Large Classes



Qualitative data: the qualitative teacher lesson modification sample is a sub-sample of the quantitative sample. The Program observed a smaller sample of 12 teachers (6 grade 2 and 6 grade 4 teachers) twice: once for Mathematics and once for ULA. The Program also observed an additional 6 grade 1 teachers for ULA, given the very different lesson structure of grade 1 TLMs for ULA. In total, the Program observed teachers making 55 ULA lesson modifications and 53 Mathematics lesson modifications.

DATA ANALYSIS

TLM Desk Review

Based on RTI's established TLM review criteria, the Program analyzed 153 lessons (TGs and STBs) as well as the front matter pages of the grade 1–4 TGs. Weekly sets of lessons were randomly selected from each grade level for each of the yearly terms. See the Desk Review Report for complete list of lessons reviewed.[10]

Qualitative Data Analysis of Teacher Lesson Modifications

The Program analyzed qualitative data for patterns and trends. Specifically, the Program iteratively coded teachers' observed lesson modifications and their explanations for these modifications until a final coding scheme emerged. To ensure 90% inter-rater reliability, coders double-coded 25% of all items and resolved all discrepancies. The Program then

[10] Reference desk review report

used descriptive statistics (i.e., analyzed data descriptively as frequencies of occurrences) for clearer rendering of findings.

Quantitative Analysis of Teacher Observable SCSs and Behaviors Over Time

The Program mapped the desired SCSs and behaviors to scores, whereby teachers who demonstrated certain student-centered behaviors received a score. Teachers accumulated total scores depending on which part of the observed lessons they were teaching. **Table 6** provides examples of scores assigned to teacher SCSs and behaviors for a Mathematics independent work lesson section.

Table 6. Score Assignment for Behavior Elements in Independent Work Section of Mathematics Lesson

Mathematics Subscale 3: Independent Work, Total: 4 points			
Questions	**Yes**	**No**	**N/A**
Did the teacher ask the students to work independently or in small groups during the independent work part of the lesson?	1	0	0
Did the teacher check the students' work while they were working independently?	1	0	0
Did the teacher give extra help to struggling students?	1	0	0
Did the teacher give a more difficult task to the students who completed the task?	1	0	0

Following the score assignments, the Program mapped the score ranges to the proficiency categories of *Not yet started*, *Novice*, *Emerging*, and *Proficient*. **Table 7** provides an example of a teacher's assigned proficiency categories when teaching a Mathematics independent work section of a lesson.

Table 7. Score Mapping to Proficiency Categories for Independent Work Section of Mathematics Lesson

Mathematics Subscale 3: Independent Work		
Proficiency category	**Description of category**	**Points**
Not yet started	The teacher did not ask students to work independently.	0
Novice	The teacher asked students to work independently.	1
Emerging	The teacher asked students to work independently and made efforts to monitor their work.	2–3
Proficient	The teacher asked students to work independently, monitored their work, and made efforts to differentiate instruction.	4

In Phases 1 and 2, the Program looked at individual observed items to better understand how teachers' application of SCSs and respective behaviors were progressing when teaching lessons and individual parts of the lessons. The Program also analyzed changes in the percentage of teachers in each proficiency category and at the differences between self-efficacy survey responses, grade, and class size.

The Program analyzed teachers' self-efficacy survey responses by calculating a self-efficacy score. The calculated score ranged from 12 to 60 (12–28 indicating low confidence, 28–44 indicating medium confidence, and 44–60 indicating high confidence), and correlated responses with classroom teacher observable behavior findings.

LIMITATIONS

One of the limitations of this study is that the sample was not sufficiently large to be representative of all 919 schools and 9,000 teachers targeted by the Program in the Sirdaryo and Namangan pilot regions. However, the sample size was sufficiently robust to be able to draw general conclusions on whether teachers applied SCSs in the classroom and how those applications evolved over two time points (Phase 1 and Phase 2).

It is also important to note that the Program did not conduct this study with a control group. Thus, the findings of this study cannot conclusively indicate that the progress made was entirely due to the Program's various interventions.

Another limitation of this study is that unlike Mathematics lessons, ULA lessons differ in terms of structure from day to day throughout the week, whereby if one lesson contains a modeling and student discussion part, the next lesson may contain mainly independent work. For purposes of data collection, this posed a problem in that the Program was not able to completely align Phase 1 school visits with Phase 2 school visits and lessons. This resulted in not having a sufficient sample size for the *independent work* component of ULA lessons.

Lastly, the self-efficacy surveys collected self-reported data, and respondents may have provided biased estimates or exhibited social desirability bias, aiming to present themselves in a favorable light. While participation was voluntary and anonymous, these factors may have influenced responses.

ULA AND MATH TEACHING AND LEARNING MATERIALS UPTAKE FINDINGS

The findings of Phases 1 and 2 of the TLM Uptake Study are presented below and are organized by RQ.

RQ1: TO WHAT EXTENT DO THE CONTENT, INSTRUCTION, AND DESIGN OF THE ULA AND MATHEMATICS TGS AND STBS SUPPORT THE DELIVERY OF LESSONS AND ADHERE TO BEST PRACTICES?

The ULA and Mathematics TGs and STBs meet the rigorous, evidence-based evaluation criteria for content, instruction, design, and adherence to student standards. The findings of the desk review are summarized below.

- **Page layout**: The ULA and Mathematics TGs and STBs page layout is consistent throughout.
- **Formatting**: All reviewed ULA and Mathematics TG lessons follow the preferred formatting guidance throughout. For example, TLMs have colorful bolded headings throughout.
- **Standards and lesson objectives**: TG lesson design includes a ULA theme or a Mathematics unit title. These are clearly labeled at the beginning of a lesson along with the individual learning standard addressed in the lesson. These set the overall objectives of the activities for each lesson.
- **Clarity and cohesion**: The Program found consistent application and effective use of icons, placement of headings, subheadings, textboxes, and images as well as alignment of TGs with STBs.
- **Instructional supports**: TG design considers teacher instructional support by using heavier scripting at the beginning of the TG, which is subsequently reduced as the year progresses and teachers become familiar with the lesson structure and accompanying methodology.
- **Structure, front matter content, functionality**: The Program found that the Mathematics and ULA TGs follow consistent practices regarding the structure, front matter content, and functionality of TGs.
- **Scaffolding of skills**: The desk review revealed that STB content progresses from simple to more difficult learning tasks, and skills are presented in a spiral format in that they are presented throughout the year.
- **Language/Text**: The Program found the STBs to be age-appropriate, and they effectively communicate appropriate messages. However, gender equity and inclusion of people with disabilities and persons from different regions and socio-economic and cultural/ethnic backgrounds are lacking in most Mathematics and ULA lessons.

- **Illustrations and graphics**: The desk review found the use of graphics and illustration to be effective. For example, illustrations and graphics accurately reflect and support the lesson content.
- **Social**-emotional learning skills: These are addressed in most lessons, specifically teamwork, relationships, responsible decision-making, and self-awareness.
- **Critical thinking**: The use of critical thinking to evaluate evidence to solve problems or complete activities independently and with others, and to interpret graphics, problems, and questions is present in most activities in the TLMs.
- **Creativity**: Lessons provide students with opportunities to explore and use their own creativity in relevant, interesting, and worthwhile ways throughout the reviewed TLMs.
- **Problem solving**: The desk review found problem solving to be present in all reviewed lessons. For example, ULA lessons include graphic organizers, Venn diagrams, word study, and comprehension questions to encourage and support problem solving.
- **Individual, gr**oup, and pair work**:** Both Mathematics and ULA TGs include group and pair work, but only in a few lessons.

In the areas of language, objectives, and clarity and cohesion, the desk review identified some areas of improvement and needed adaptations. The Program subsequently addressed these and other recommendations during the development and revision of the TLMs. These are described in detail the Desk Review Report.[11]

RQ2: HOW ARE TEACHERS MODIFYING ULA AND MATHEMATICS LESSONS AS THEY TEACH WITH THE NEW TLMS IN THE CLASSROOM?

Mathematics

As teachers began to use the TLMs in Phase 1, they modified Mathematics lessons mostly because they struggled to become familiar with the new methodology of giving students time to think and explain independently.

Sampled teachers made six types of Mathematics lesson modifications. Sixty percent of teachers' lesson modifications indicated that teachers did not provide students with sufficient time to complete individual, independent work. Another significant type of lesson modification included teachers' not providing students with the opportunity to discuss or explore solutions or wrong answers (47%). Teachers also modified lessons by adding content that was aligned to the lesson objectives (25%) and content that was not aligned to lesson objectives (6%). Table 8 lists all six coded types of modifications made by teachers and provides descriptions of each modification together with illustrative examples.

Table 8. Teacher Mathematics Lesson Modification Types, Frequencies, Descripion, Examples

Modification type	Frequency	Description	Exemplar
Individual work	60%	Teacher does not allow students to work independently.	Instead of independent work, teacher asked student to come to the front of the class to solve the problem.

[11] Reference desk review report.

Table 8. Teacher Mathematics Lesson Modification Types, Frequencies, Descripion, Examples

Modification type	Frequency	Description	Exemplar
Discussion	47%	Teacher explains without asking students for solutions to mathematics problems. Teacher does not discuss wrong answers.	Teacher explained before letting students discuss possible solutions.
Similar addition	25%	Teacher adds tasks that are aligned to the lesson objective.	In symmetric shapes activity, teacher asked extra questions such as “find symmetric shapes in the classroom.”
Reorganization	11%	Teacher changes order of activities.	Teacher combined problem presentation activity with discussion activity.
Not similar addition	6%	Teacher adds tasks that are not aligned to the lesson objective.	Teacher asked students to write the number 18 “beautifully.”
Other	28%	Teachers omits or substitutes content, skips activity, or misallocates time.	Teacher spent too much time reviewing previous lesson.

In sum, at the beginning of the school year, teachers struggled to reorient their deeply ingrained approach of presenting problems to students at the beginning of the lesson and not letting students explore the problem without prior teacher explanation. As we will see in RQ3, below, teachers’ practice improved in this regard by the end of the school year.

Teacher explanations: To delve deeper into why teachers made the above lesson modifications, the Program asked teachers to explain the reasoning behind their modifications. Teachers’ explanations were varied, including expressing the need to be mindful of all or some students’ learning needs (43% and 11%, respectively), wanting to keep to their former teaching methodologies (11%), or lacking time to properly prepare for the lesson (11%). **Table 9** lists the coded teacher explanations as well as examples for each.

Table 9. Teacher Mathematics Explanation Types, Frequencies, Description, and Examples

Modification type	Frequency	Description	Exemplar
All students’ knowledge	43%	Teacher indicates that students either need more practice or more time, or already know the content.	“Today's lesson is not new for my students. My students know new content so I don't think I need to provide time after presenting the problem.”
Individual students’ knowledge	11%	Teacher differentiates between “good” and “slow” students (individually or by groups).	“Some students may not understand. That's why I always ask them to solve problems after discussing problems together.”
Old methodology	11%	Teacher uses methodology that she already knows, or that she is more comfortable with.	“I always use this method.”
Preparation	11%	Teacher states that she does not know how to do an activity or a	“Honestly, I do not know how to orginize the lesson.”

Table 9. Teacher Mathematics Explanation Types, Frequencies, Description, and Examples

Modification type	Frequency	Description	Exemplar
		certain lesson section. Or she does not feel prepared and confuses sections.	
Other	23%	Teacher indicates challenges with classroom or time management or does nott answer.	"If I discuss each wrong answer, there isn't enough time."

Teachers' explanations point to a pattern. Teachers strove to be attentive to students' needs and made decisions about what and how to teach based on their perceived students' needs.

ULA

At the beginning of school year, teachers mostly modified ULA lessons by either omitting or adding content as they struggled with the timing of ULA lessons and the activities therein.

The most predominant type of ULA lesson modifications that teachers made was content omission (42%). In addition, teachers also added content (24%) and changed the teaching methodology (27%). **Table 10** lists all four coded types of ULA modifications made by teachers and provides descriptions of each modification together with illustrative examples.

Table 10. Teacher ULA Lesson Modification Types, Frequencies, Description, Examples

Modification type	Frequency	Description	Exemplar
Content omission	42%	Teacher omits content from entire or part of an activity.	Did not ask students to make up phrases/sentences using new vocabulary.
Content addition	24%	Teacher adds tasks that are aligned to the lesson objective.	After reading the 'teacher read aloud' the teacher summarized the story herself.
Change of methodology	27%	Teacher changes the nature of an activity or part of an activity.	Instead of students making up sentences orally, they wrote them down.
Other	9%	Teachers has issues with classroom management or timing.	Spent 3 minutes instead of 15 on an activity.

Almost 70% of ULA lesson modifications were comprised of either content omissions or content additions. This can be attributed in part to the fact that teachers struggled with time management (see below). Even though the TGs include timing for each activity, teachers rushed to complete lessons and in so doing found themselves either having to skip activities or to add activities at the end of the lesson because they finished too soon. The Program addressed these issues of timing during subsequent trainings.

Teacher explanations: Similar to Mathematics teachers' explanations of their modifications, ULA teachers' explanations also included being mindful of students' needs (29%). Lack of preparation was the most used explanation by teachers (31%) as corroborated by Program staff and subsequently addressed in future trainings. **Table 11** lists the coded teacher explanations as well as examples for each.

Table 11. Teacher ULA Explanation Types, Frequencies, Description, and Examples

Modification type	Frequency	Description	Exemplar
Preparation	31%	Teacher states that she is nervous, not ready.	"I was nervous because I was not ready for the lesson."
Knowledge of students	29%	Teacher is mindful of what her students needs.	"My students love this activity. It helps them to learn how to make up words from letters."
Time	15%	Teacher either took too long or rushed or finished ahead of schedule.	"I don't always manage to complete the 'writing for purpose task' and so I assign it for homework."
Other	25%	Teacher states having personal life issues to address.	"I lost my close realtive and I was busy with family gatherings. I could not prepare for the lesson."

The Recommendations section below describes how the Program addressed teacher modifications and explanations.

RQ3: TO WHAT DEGREE ARE ULA AND MATHEMATICS TEACHERS APPLYING SELECTED STUDENT-CENTERED STRATEGIES IN THE CLASSROOM?

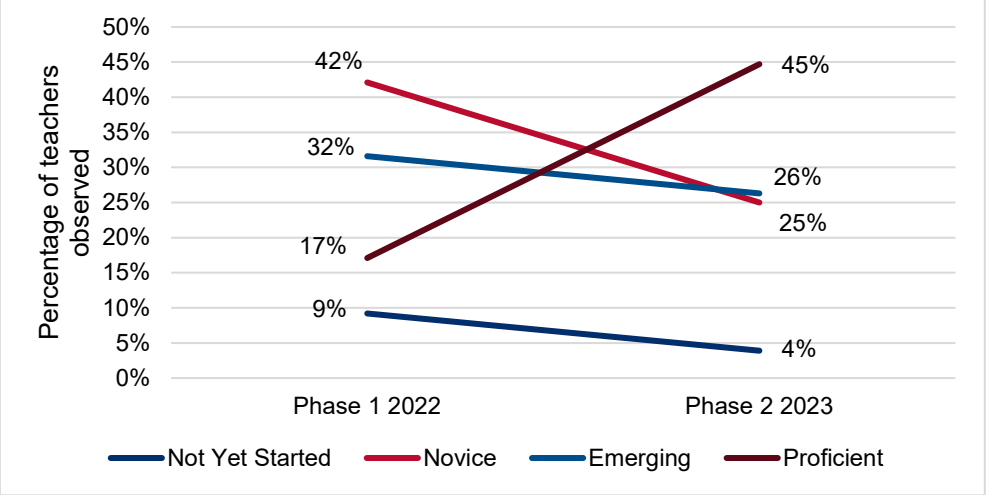
Mathematics

Overall findings

Teachers' support of student explanation and independent work types of SCSs improved significantly between Phases 1 and 2. However, teachers' modeling of Mathematics problem solving strategies remained unchanged between both phases.

Figure 2 shows the overall percentage change of teacher behavior by proficiency category over time. Overall, the proportion of teachers in the *Proficient* category increased significantly from 17% to 45%. This indicates that between phases, teachers increased their use of the TGs and demonstrated an increase in the application of selected SCSs.

Figure 2.Overall Percentage Change in Teacher Observed Behaviors by Proficiency Category, by Phase (Mathematics)



The Program did not observe any statistically significant differences in teachers' performance across grades, class sizes, and self-efficacy scores. Additionally, there was no correlation between class size (medium and large) and how teachers applied SCSs.

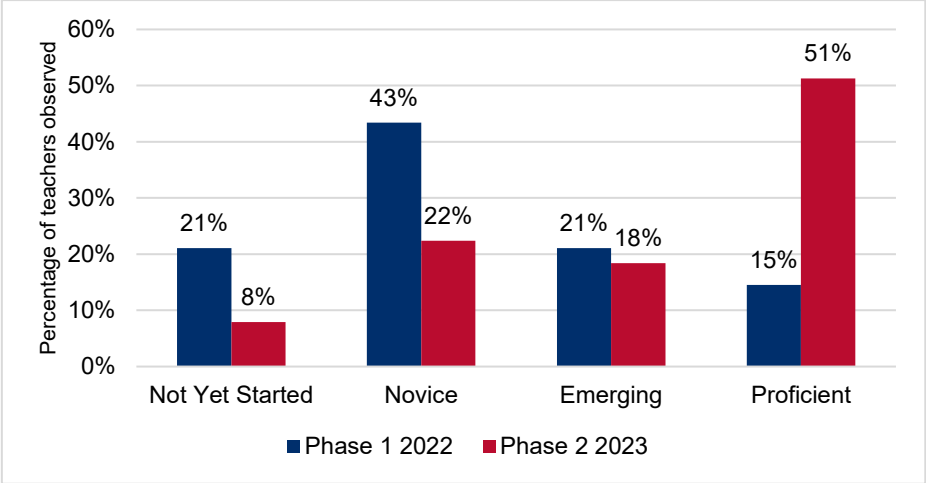
The self-efficacy score analysis shows that a majority of teachers indicated having high self-efficacy in their ability to teach mathematics. This indicator did not change from Phase 1 to Phase 2.

Teacher support of student explanation and engagement

By the end of the school year a larger number of teachers provided their students with opportunities to think about new mathematics problems independently and for students to offer their own explanations. The Program tracked the extent to which teachers supported student explanation and engagement by observing whether teachers (1) provided students with opportunities and time to think and work on new problems independently, (2) asked students for their answers to the questions, and (3) asked students to explain their answers.

Figure 3 presents teachers' proficiency categories for each data collection phase. In Phase 1, most teachers (43%) fell within the *Novice* category and only 15% of teachers fell within the *Proficient* category. By Phase 2, however, 51% of teachers achieved *Proficient* status.

Figure 3. Percentage of Teachers by Proficiency Category, by Phase (Mathematics Problem Presentation and Explanation SCS)



Although in both Phases 1 and 2 teachers did ask students for the answers to the discussion questions (hence the high *Novice* percentage in Phase 1), many teachers only started asking students to *explain* their answers by asking questions like "why and how did you get this answer" in Phase 2.

Teacher modeling

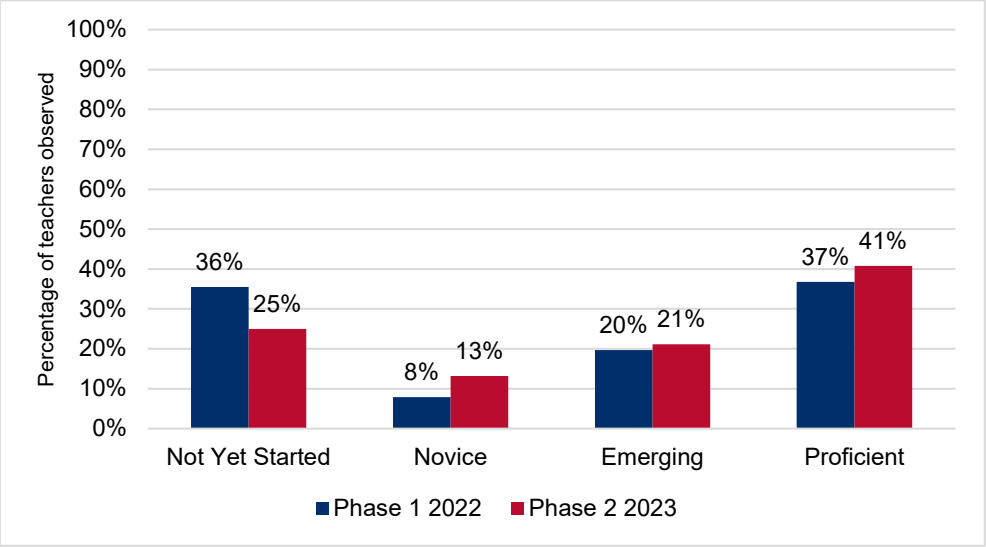
The Program did not observe any significant changes in how teachers modeled problem solving between Phase 1 and 2.

The desired teacher behaviors for modeling problem solving activities are as follows: teachers (1) model how to solve problems in the discussion part of the lesson, (2) engage students in the explanation process, (3) ask yes/no questions, and (4) ask open-ended questions.

Findings indicate that only approximately 40% of teachers achieved the level of *Proficient* in all four of the above-described desired modeling behaviors (**Figure 4**). Program experts

posit that even though modeling was greatly stressed during trainings, teacher modeling did not change significantly because instructions in TGs were not explicit in this regard. Moreover, given the limited time of the pilot, the Program opted to place greater emphasis on student engagement with explanation and justification together with independent work because these were part of lessons that were "newest" to teachers. Teachers have traditionally modeled and modeled well, so even though this behavior can always be improved, it was not an overt focus in the materials.

Figure 4. Percentage of Teachers by Proficiency Category, by Phase (Mathematics Second Discussion Part of the Lesson)

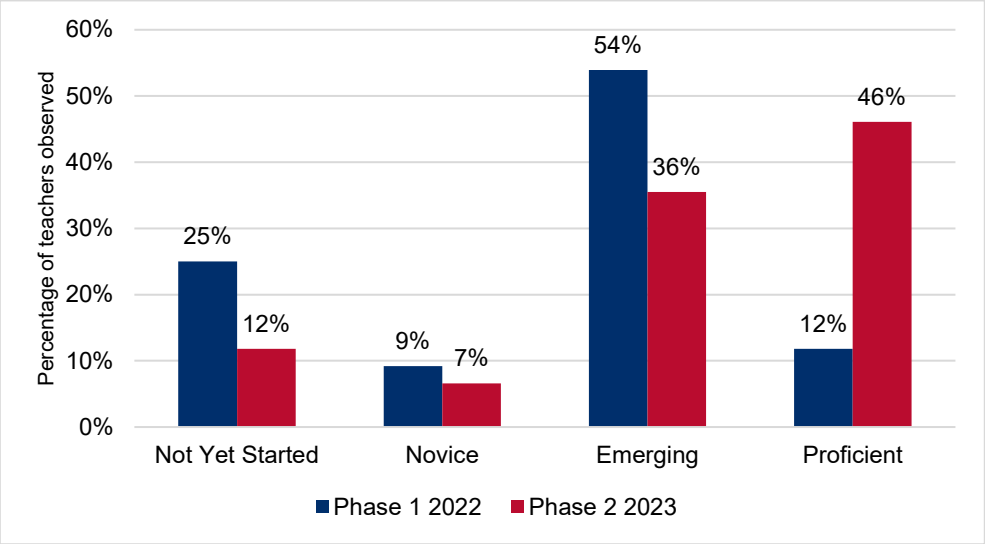


Student independent work and student interactions

Teachers made significant progress in how they conducted student independent work activities. Over 85% of teachers fell within the *Emerging* category or below in Phase 1. By Phase 2, over 45% of teachers achieved *Proficient* status (**Figure 5**).

The Program observed how teachers conducted the mathematics independent work part of the lesson by tracking whether teachers: (1) asked students to do independent work, (2) monitored students while they were working, (3) supported students who were struggling, and (4) and gave more tasks to students who finished the task earlier.

Figure 5. Percentage of Teachers by Proficiency Category, by Phase (Mathematics Student Independent Work Part of the Lesson)



In Phase 1, the Program observed that almost all teachers conducted student independent work but a significant majority (88%) of teachers failed to provide students who finished the task earlier with additional tasks. However, in Phase 2, teachers began to engage fast performing students by providing them with additional tasks and applying a more differentiated approach in the classroom. Thus, more teachers were categorized as *Proficient* by Phase 2.

ULA

Overall findings

Teachers showed significant improvement in supporting student explanations and discussions.

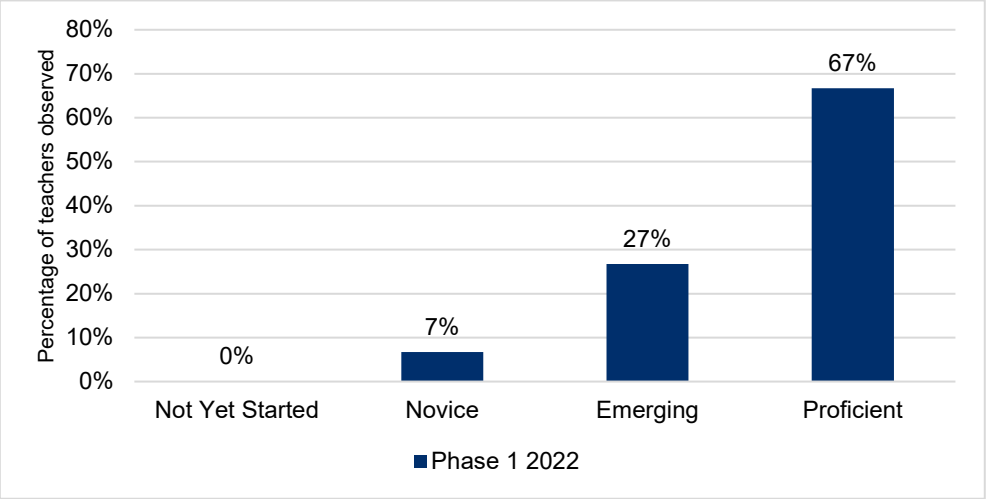
The Program did not observe any significant correlations between teachers' SCSs and behaviors and factors such as grade, class size, and self-efficacy scores. Although teachers' ULA self-efficacy scores increased by a small magnitude, the score did not correlate with proficiency scores. This is likely due to the fact that a majority of teachers had high self-efficacy at the beginning of the pilot.

Teacher modeling

Teacher modeling in ULA was only observed in Phase 1. A majority of teachers performed well in this part of the lesson (27% *Emerging* and 67% *Proficient*).

The Program observed teachers modeling during the phonics activity (reading syllables/words). This was only applicable in the first half of the grade 2 school year. To observe teacher behaviors the Program looked at several elements. These included whether teachers (1) read the syllables/words themselves first (modeling), (2) confirmed students' understanding, and (3) gave students the opportunity to read independently. Some teachers did not achieve the *Proficient* category because they either did not model the task first or because they did not give students the opportunity to read independently (see **Figure 6**). However, over 90% of teachers were *Emerging* or above (27% *Emerging* and 67% *Proficient*). This was due in large part to the fact that SCSs in this section were very similar to the ones in the former curriculum.

Figure 6. Percentage of Teachers by Proficiency Category (ULA Teacher Modeling Part of Lesson, Phase 1)



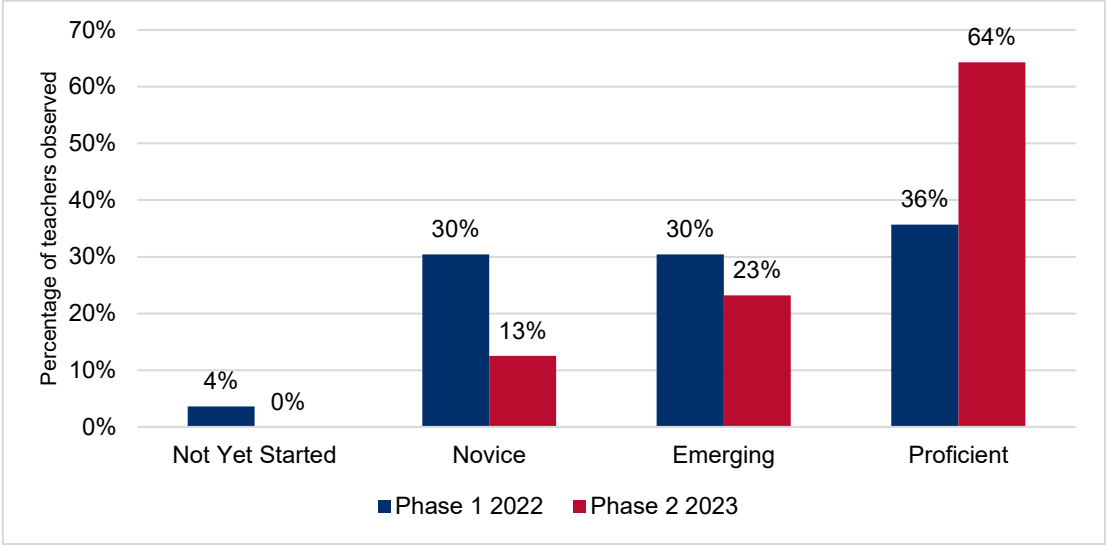
Teacher support of student discussion and engagement

The percentage of teachers categorized as *Proficient in* supporting student discussion and engagement almost doubled by Phase 2.

To observe and measure teachers' behaviors on how they supported students' discussions and engagement, the Program looked at how teachers conducted the reading and listening comprehension part of the ULA lesson. The desired SCSs for this part of the lesson were whether teachers: (1) asked students to predict the content of the text based on the illustrations, (2) asked questions after students read the text, and (3) engaged passive students in the classroom discussion.

In Phase 1 teachers were equally distributed across the categories of *Novice, Emerging,* and *Proficient* (30%, 30%, and 36%, respectively). By Phase 2, however, over 60% of teachers achieved the status of *Proficient* (**Figure 7**). This progress was mainly due to more teachers engaging passive students in the classroom discussion. This teacher behavior increased from 56% in Phase 1 to 85% in Phase 2.

Figure 7. Percentage of Teachers by Proficiency Category, by Phase (ULA Listening and Reading Comprehension Discussions Parts of the Lesson)



Student independent work

The Program observed a significant variation in teachers' observed behaviors when they were conducting the independent writing part of the lesson.

To observe how teachers conducted independent work, the Program looked at the independent writing part of the ULA lesson. As this specific activity is not consistently present in every ULA lesson, at the time of data collection, it was not possible for the Program to predict when or during which part of the day sampled teachers would be teaching the independent part of the lesson. Consequently, the Program could not accumulate a sufficiently large sample size to mitigate the variation within the observations (there were only approximately 20 observations of this particular SCS per phase).

CONCLUSIONS AND RECOMMENDATIONS

This section provides an overview of the main conclusions and recommendations drawn from the above findings and are presented by research question.

RQ1: TO WHAT EXTENT DO THE CONTENT, INSTRUCTION, AND DESIGN OF THE ULA AND MATHEMATICS TGS AND STBS SUPPORT THE DELIVERY OF LESSONS AND ADHERE TO BEST PRACTICES?

With the exception of some required changes in the areas of language, goals and objectives, inclusion, and clarity and cohesion, The ULA and Mathematics TGs and STBs meet the rigorous, evidence-based evaluation criteria for content, instruction, design, and adherence to student standards.

Recommendations

In Phase 1, the above suggested changes were reflected in three sets of recommendations: (1) recommendations based on feedback provided by teachers, the Program's product review group, and the Program technical teams; (2) recommendations that emerged from the desk review; and (3) longer-term recommendations for MoPSE to consider prior to eventual reprinting. All of these recommendations can be found in the final Desk Review Report.[12]

RQ2: HOW ARE TEACHERS MODIFYING ULA AND MATHEMATICS LESSONS AS THEY TEACH WITH THE NEW TLMS IN THE CLASSROOM?

- More than half of teachers' lesson modifications did not provide students with sufficient time to complete Mathematics individual and independent work. Teachers struggled to reorient their deeply ingrained approach of presenting Mathematics problems to students at the beginning of the lesson and were not letting students explore the problem without prior teacher explanation. By the end of the school year teachers improved in this regard.
- Teachers' lesson modifications showed that they struggled with the new methodology in the Mathematics books. As such, the Program added several features to the TGs to support teachers in applying the new methodology more effectively; for example, an icon was added to remind teachers when to let students work independently.
- In Phase 1, almost 70% of ULA lesson modifications made by targeted teachers were comprised of either content omissions or content additions. This can be attributed in part to the fact that teachers struggled with time management. Even though the TGs include timing for each activity, teachers rushed to complete lessons and in so doing skipped or added activities at the end of the lesson because they finished too soon. The Program addressed these issues of timing during subsequent trainings.

[12] Reference final desk review report.

- When teaching both Mathematics and ULA, teachers' modifications showed that they strove to be attentive to students' needs and made decisions about what and how to teach based on their students.

Recommendations

In Phase 1, the Program addressed the above findings by implementing the following selected recommendations.

- Added TPD content on effective differentiated instruction.
- Added an icon in TGs to remind teachers when to let students work independently.
- Focused training on providing all students with time to solve problems.
- Embedded reminders in the TGs to allow students adequate time to solve problems independently.
- Focused training on practicing how to ask questions and review.
- Included questions and structured feedback in the quality assurance tools.
- Reinforced the writing lesson procedure in trainings and highlighted key elements of the writing lesson procedure through Telegram posts.

The complete list of findings and recommendations can be found in the TLM Uptake Study Phase 1 Briefer.[13]

RQ3: TO WHAT DEGREE ARE ULA AND MATHEMATICS TEACHERS APPLYING SELECTED SCSs IN THE CLASSROOM?

Overall, it is possible to conclude that teachers' uptake of the ULA and Mathematics TLMs was successful, especially considering the short one-school-year period in which this pilot took place. This is of even greater note given the magnitude of shift required of teachers in terms of having to absorb a new standards-based curriculum, new content, and new SCSs that were significantly different from their previous experience. Specific conclusions follow.

- By Phase 2, overall, the proportion of teachers in the *Proficient* category increased significantly from 17% to 45%. This indicates that between phases, teachers increased their use of the TGs and demonstrated an increase in the application of selected SCSs. This is particularly significant in that teachers substantially increased their proficiency in only one year.
- By Phase 2, teachers, whose proficiency in providing students with opportunities to think about new mathematics problems independently and for students to offer their own explanations improved significantly, went from 43% *Novice* to 51% *Proficient.*
- The Program did not observe any significant changes over time on how teachers modeled mathematics problem solving. This is likely because instructions in TGs were not explicit in this regard and the Program placed greater emphasis on student engagement with explanation and justification together with independent work, as these were parts of the lessons that were "newest" to teachers.
- Teachers made significant progress in how they conducted Mathematics student independent work activities. Over 85% of teachers fell within the *Emerging* category or below in Phase 1. By Phase 2, over 45% of teachers achieved *Proficient* status.

[13] Reference TLM Phase 1 Briefer.

- A majority of teachers performed well when modeling the phonics part of the ULA lessons (27% *Emerging* and 67% *Proficien*t). This was due in large part to the fact that SCSs in this section were very similar to the ones in the former curriculum.
- The percentage of teachers categorized as *Proficient in* supporting student ULA discussion and engagement almost doubled by Phase 2. This was mainly because more teachers engaged passive students in the ULA classroom discussions. The Program observed a significant variation in teachers' observed behaviors when they were conducting the independent writing part of the lesson. The sample was not large enough to mitigate the variation and provide sufficient statistical significance.

Recommendations

- Given that changes in teacher behavior of this magnitude take time and commitment for full uptake at scale to take root, MoPSE should continue to monitor teachers' application of new SCSs with similar longitudinal studies to better understand the relationship between new materials, TPD, integration of new SCSs into classroom practices, and timing.
- The Program's robust evidence-based piloting approach, which was based on existing education systems and priorities, has proven to be an enormously powerful strategy for improving TLMs and strengthening teachers' student-centered teaching behaviors. MoPSE should consider adapting and/or adopting a similar robust pilot approach when introducing new materials and accompanying TPD.
- Future interventions should ensure teacher modeling instructions are clear and focus on the more challenging teaching strategies.
- Given that teachers were well versed in phonics instruction, future interventions should better identify teachers' strengths prior to developing materials.[14] In so doing, future interventions can be more strategic on what new SCSs to focus on.
- If the lesson structure of TLMs is very varied, future studies should ensure a sufficiently large sample to accommodate for this variation and to ensure that all parts of the lessons are observed.
- To better determine whether progress is achieved due to the Program's interventions or due to other intervening variables, future studies would ideally include a control group.

[14] The Program had originally planned to conduct a classroom observation as part of the Status of Instruction Study but was not able to do so because of coronavirus disease pandemic travel restrictions at the time. The TLM Uptake Study provided an opportunity to introduce classroom observations.

REFERENCES

Bulat, J., Dubeck, M., Green, P., Harden, K., Henny, C., Mattos, M., Pflepsen, A., Robledo, A., & Sitabkhan, Y. (2017). *What we have learned in the past decade: RTI's approach to early grade literacy instruction.* RTI Press Publication No. OP-0039-1702. Research Triangle Park, NC: RTI Press. https://files.eric.ed.gov/fulltext/ED582383.pdf

Enochs, L., Smith, P.L, & Huinker, D. (2000). Establishing factorial validity of the mathematics teaching efficacy beliefs instrument. *School Science and Mathematics 100(4)*, 194–202.

Nakamura, P., Shone, R., & Saidoshurov, S. (2016). *Teaching and learning material in Tajikistan: How do they align with reading research?* American Institutes for Research (AIR) report on U.S. Agency for International Development (USAID) Quality Reading Project. LOCATION: PUBLISHER.

Piper, B., Sitabkhan, Y., Mejía, J., & Betts, K. (2018). *Effectiveness of teachers' guides in the Global South: Scripting, learning outcomes, and classroom utilization*. RTI Press Publication No. OP-0053-1805. Research Triangle Park, NC: RTI Press. https://doi.org/10.3768/rtipress.2018.op.0053.1805

RTI International. (2015). A guide for strengthening gender equality and inclusiveness in teaching and learning materials. A report for the EdData II: Data for Education Research and Programming (DERP) Program. Washington, DC: U.S. Agency for International Development.

Sitabkhan, Y., Harden, K., & Slade, T. (2022). Teaching by the book: teacher decision-making while using structured lesson plans. *IARTEM e-journal* 14(1), 1-12.