SCHOOL-TO-SCHOOL INTERNATIONAL

SCIENCE OF TEACHING

TANZANIA NUMERACY RESEARCH PROJECT: WHAT'S THE BUZZ? STUDENTS DISCUSSING MATH IDEAS IN GROUPS

Final Report

TANZANIA NUMERACY RESEARCH PROJECT:

WHAT'S THE BUZZ? STUDENTS DISCUSSING MATH IDEAS IN GROUPS

FINAL REPORT

August 18, 2025

This publication was produced with funding from the Gates Foundation. It was prepared independently by Dr. Parnika Bhatia, Amani Nicolas, and Dr. Mark Lynd of School-to-School International in collaboration with RTI International.

CONTENTS

ACRONYMS

AG	Ability Grouping
CPA	Concrete-Pictorial-Abstract
DEO	District Education Office
FGD	Focus group discussion
FMKT	Foundational Mathematics Knowledge for Teaching
M&E	Monitoring and evaluation
MTEBI	Mathematics Teaching Efficacy Beliefs Instrument
PI	Principal investigator
RTI	Research Triangle Institute International
STS	School-to-School International

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

The Science of Teaching activity, administered by RTI International (RTI) and funded by the Gates Foundation, is focused on compiling, generating, and disseminating evidence on what works to improve foundational literacy and numeracy teaching and learning. As part of its numeracy activities, Science of Teaching has supported research on strategies to improve foundational mathematics instruction in primary schools in Sub-Saharan Africa and South Asia.

Study Purpose, Background, and Research Questions

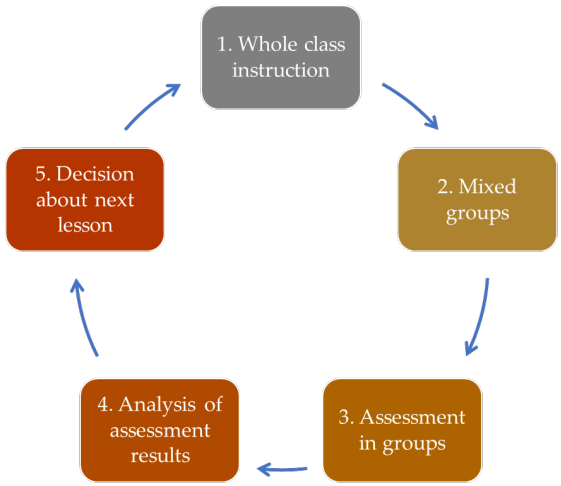
RTI awarded School-to-School International (STS) a contract to implement the Tanzania Numeracy Pilot Research Project to explore the use of Foundational Mathematics Knowledge for Teaching (FMKT) strategies in the context of flexible ability grouping (AG). The goal was to determine whether this combination could strengthen teachers' ability to teach effectively in the early grades in selected schools in Arusha, Tanzania. The Research Project also examines teachers' attitudes toward their roles as student-centered instructors. In this project, student-centered instruction involves using assessment data to tailor teaching and viewing student errors as valuable learning opportunities. Ongoing assessments are used not only to monitor student progress, but also to guide daily teaching decisions. By analyzing student work and responses, teachers can identify learning needs and provide targeted support. Moreover, student errors are seen as chances to deepen understanding: when mistakes are discussed openly in a supportive classroom, they reveal student thinking, uncover misconceptions, and inform future instruction. FMKT denotes both mathematics content knowledge and pedagogy, or understanding how to teach math concepts effectively. The focus of this research was on mathematics pedagogy, including FMKT interventions of training teachers on the use of concrete materials—manipulables such as bottle caps or stones—and strategies for addressing student errors in the classroom. The use of concrete materials in mathematics instruction is emphasized in this project to not only engage students, but more importantly, to build their conceptual understanding. By working with tangible objects, students can grasp abstract mathematical concepts more deeply and meaningfully.

Flexible AG is an instructional strategy in which teachers group students by ability based on ongoing assessment results, then assist those students according to their different levels to improve learning outcomes for all students (Ciccarelli, 2024; Morin, 2024). In Tanzania, STS has employed a flexible AG method for large class sizes that follows the following steps:

1. **Lesson**: The teacher teaches a lesson starting with whole class instruction, then organizes students into mixed-ability groups to continue learning.
2. **Assessment**: After learning in their groups, each student responds to one more assessment question in their copybook, then passes it to a classmate who marks the book according to the teacher-provided correct responses. The students then pass their copybooks to a group leader who tallies the results on a score analysis sheet.

3. **Calculation**: After class, the teacher tallies the number and percentage of students meeting the threshold.
4. **Next lesson**: Depending on the percentage of students meeting the threshold, the teacher either moves on to the next lesson (more than 70 percent), reteaches the lesson (less than 30 percent), or plans a lesson where students are assigned to enrichment and remediation groups (30–69 percent).

Figure 1: Flexible Ability Grouping Steps



Attitudes: This study focused on training teachers in student-centered instruction to shift traditional perceptions of teachers' role from transmitters of knowledge to facilitators of learning. After implementing new AG, FMKT, and student-centered instructional strategies, the Tanzania Numeracy Pilot Research Project explored teachers' beliefs, attitudes, identity, and knowledge through self-reports.

The Research Project's intervention design is based on the theory that if teachers receive training and follow-up support in three areas—AG, FMKT, and student-centered teaching approaches—then they will successfully apply these approaches to address the needs of struggling students, effectively teach foundational math, and exhibit student-centered approaches. The research questions of this study are organized into three main categories:

1. **AG: How well can teachers implement flexible ability grouping during numeracy instruction?**
   - Do teachers implement AG with fidelity?
   - Do teachers adjust instruction according to AG results?
   - Do teachers move students to different groups according to their progress?
2. **FMKT: Do teachers use selected elements of FMKT?**
   - Do teachers use concrete materials appropriately?
   - Do teachers seek to understand the reasons for students' errors?
   - Do teachers effectively address students' errors?

3. **Attitudes: How do teachers view their role?**
    - Do teachers see their role as transmitters of knowledge or facilitators of learning?
    - How does the use of AG and FMKT principles influence teachers' views of their role?

In this project, AG and FMKT are closely linked in supporting and enhancing student-centered instruction. AG enables teachers to tailor instruction to meet diverse student needs, while FMKT provides the deep content knowledge needed to select appropriate concrete materials that build students' conceptual understanding. It also helps teachers effectively use student errors as opportunities for learning. Together, AG and FMKT promote a student-centered approach by guiding teachers to use ongoing assessments to inform instruction, address individual learning gaps, and turn errors into meaningful learning experiences. This integrated approach forms the foundation for the research questions and offers a framework for improving early grade mathematics teaching.

Intervention, Study Design, Methods, and Limitations

Description of the Intervention: In March 2025, STS and Arusha District Education Officers provided a two-day training to Grade 1–3 teachers and head teachers (n=32) from eight high-needs schools in Tanzania's Arusha District. The training focused on three areas:

- The purpose and practice of flexible AG, including the use of a mobile-based app to track student assessment results,
- An introduction to FMKT, with a focus on the use of concrete materials and strategies to work with student errors, and
- The importance of using student-centered approaches in the application of AG and FMKT.

Follow-up support visits to the schools took place in April 2025, offering additional coaching and guidance to all participating teachers and head teachers. Additionally, in each of the 8 schools, STS asked the head teacher to support their teachers and fill in the observation template for three weeks. The template instructed head teachers to observe three teachers from Grades 1-3, hold discussions with these teachers, and submit observation data and a weekly summary of the discussions.

Study Design and Methods: The Tanzania Numeracy Pilot Research Project employed an exploratory, concurrent mixed-methods design. It collected both quantitative and qualitative data at a single time point. Data sources included classroom observations, teacher attitude surveys, head teacher interviews, teacher focus group discussions (FGDs), and teacher interviews. Schools were purposively selected by the District Education Office (DEO) and had not previously received any similar interventions.

Qualitative data were analyzed for themes and patterns, while quantitative data were reviewed for interrater reliability. Findings were triangulated across data sources to enhance validity and capture a comprehensive picture of implementation.

Limitations: This study's findings should be considered in light of several significant limitations. First, the Tanzania Numeracy Research Project was conducted over the course of six months, which limited the time available for schools and teachers to incorporate new instructional strategies fully.

The research involved a small, purposively selected sample of high-needs schools, so results are descriptive and exploratory. They are not generalizable to the broader population. There was no comparison or control group included, which means analysts cannot conclude how outcomes compare to typical teaching practices.

Additionally, the absence of an external baseline made it difficult to attribute observed improvements solely to the intervention rather than to other factors. As a result, the findings should be interpreted as initial insights into the uptake and implementation of the targeted mathematics teaching approaches, rather than as evidence of causal impact.

Findings and Conclusions

Key AG Finding. Moderate but Growing Adoption of Flexible Ability Grouping With Specific Implementation Challenges

Teachers are integrating flexible AG into numeracy instruction. Roughly 30 percent of teachers demonstrating strong fidelity in following all recommended steps and another 60 percent show moderate fidelity by partially implementing core components. Most teachers effectively provide whole-class instruction and form heterogeneous groups, allowing for tailored feedback and peer support. However, structured assessment-based regrouping and consistent use of tracking tools remain limited. This appears to be primarily due to technological barriers, behavior management, and space constraints. Despite these challenges, teachers and head teachers report greater student engagement and classroom management benefits.

Key FMKT Finding: Widespread Use of Concrete Materials and Addressing Student Errors

All teachers are actively use concrete materials and hands-on strategies as part of math instruction, supporting conceptual understanding and student engagement. Observations and interviews confirm that nearly all classrooms use everyday objects for demonstration, with teachers encouraging students to explore, explain their reasoning, and problem-solve in multiple ways. Additionally, identifying and addressing student errors has become a standard part of instruction. Challenges arise from limited and dirty materials, time constraints for lesson planning, and crowded classrooms. Finally, head teachers report increased motivation, effort, and willingness of their teachers to try new techniques.

Key Attitudes Finding: Facilitation and Student-Centered Practice, With Room for Growth

Teachers generally view themselves as facilitators of learning, with classroom observations rating practice as predominantly student-centered. Most teachers report moderate to high self-efficacy in teaching mathematics, particularly with FMKT-aligned strategies. Confidence in more advanced instructional tasks varies.

Beliefs about student learning are broadly optimistic. Teachers express responsibility for all students' success and acknowledge the influence of AG and FMKT on improving their practice. However, there is continued variance in teachers' self-concept as facilitators versus knowledge transmitters, partly attributed to language and training differences. Despite some variation, consensus exists on the value of AG and FMKT in making mathematics teaching more effective.

Recommendations

Based on the strong enthusiasm and modest early uptake of AG and FMKT strategies in high-needs schools—there is potential for these approaches to provide practical, scalable tools for teachers managing large classrooms and working to improve foundational numeracy in Tanzania. STS recommends expanding the model to include a larger number of schools, with the intent of studying its effect using pre/post measures in contrast to comparison schools. In addition, STS recommends building on interventions that showed promise by:

Maximizing Use of Concrete, Locally Available Materials through Community Involvement. Project implementers could encourage schools to actively involve parents and community members in sourcing, preparing, and maintaining hands-on materials for math instruction. Community participation can help ensure a steady supply of clean and appropriate resources, make lessons more engaging, and address challenges such as overcrowded desks or the lack of specific teaching tools.

Engaging School Leaders for Continuous Improvement and Peer Learning. Future implementation could maintain active support and monitoring roles for head teachers, including regular classroom walk-throughs and coaching by DEOs that is focused on AG and FMKT strategies. They could also foster a culture of peer sharing by celebrating innovative teaching, sharing lesson successes, and facilitating professional dialogue during staff meetings. School leadership engagement is crucial for sustaining gains and encouraging teacher experimentation.

Providing Rigorous, Sustained Training and Ongoing Support. Future implementation could offer ongoing, context-specific professional development to deepen teachers' understanding and implementation of AG and FMKT methods. Future training should include practical classroom management techniques, flexible grouping strategies, and methods for tracking student progress in large classrooms. Establishing professional learning communities will encourage collaboration, lesson observations, and joint problem-solving among teachers.

EVALUATION PURPOSE AND EVALUATION QUESTIONS

Evaluation Purpose

The Science of Teaching activity, administered by RTI International (RTI) and funded by the Gates Foundation, is focused on compiling, generating, and disseminating evidence on what works to improve foundational literacy and numeracy teaching and learning. As part of its numeracy activities, Science of Teaching is has supported research on strategies to improve foundational mathematics instruction in primary schools in Sub-Saharan Africa and South Asia.

RTI awarded School-to-School International (STS) a contract to implement the Tanzania Numeracy Research Project to explore the use of Foundational Mathematics Knowledge for Teaching (FMKT) in the context of flexible ability grouping (AG). The Research Project seeks to determine whether this combination could strengthen teachers' ability to teach effectively in the early grades in Arusha, Tanzania. The Research Project also examines teachers' attitudes toward their roles as student-centered instructors.

This final report outlines the research questions, the intervention and its theoretical underpinning, the research framework and methodology, and presents the Research Project's findings and recommendations.

Research Questions

The project examined whether providing teachers with two days of training and site-based follow-up support would result in their use of AG with FMKT to enhance their assessment of students' progress in math and provide teachers with strategies to support struggling students. STS hoped that outcomes would result in increased use of student-centered strategies or more favorable views of student-centered teaching..

The specific research questions are as follows:

1. **Ability Grouping (AG): How well can teachers implement flexible AG during numeracy instruction?**
    - Do teachers implement AG with fidelity?
    - Do teachers adjust instruction according to AG results?
    - Do teachers move students to different groups according to their progress?

The Research Project closely examined how teachers tracked assessment results by student, diagnosed the difficulties of struggling students, and provided appropriate remediation.

2. **Foundational Mathematics Knowledge for Teaching (FMKT): Do teachers use selected elements of FMKT?**

- Do teachers use concrete materials appropriately?
- Do teachers seek to understand the reasons for students' errors?
- Do teachers effectively address students' errors?

The Research Project closely examined the appropriate use of concrete materials and strategies for working with student errors. Strategies for working with student errors are viewed as an example of providing appropriate remediation, through AG, and other types of support.

3. **Attitudes:** How do teachers view their role?
    - Do teachers see their role as transmitters of knowledge or facilitators of learning?
    - How does the use of AG and FMKT principles influence teachers' views of their role?

The Research Project explored teachers' beliefs regarding the use of AG and FMKT as instructional methods and teachers' own perceptions on how they envisioned their role in the classroom.

The Research Team

Three STS technical experts led the design and implementation of the Tanzania Numeracy Research Project, working closely alongside the Arusha DEO and selected schools.

Co-Principal Investigator (Co-PI 1), Amani Nicolas led the design and implementation of the research study in collaboration with the Co-PI 2. He also served as the in-country lead for teacher training and follow-up support as well as daily data quality assurance monitoring. He supported Co-PI 2 with translation, qualitative analysis, and report writing. Mr. Nicolas is a Tanzanian researcher with expertise in science and mathematics that has supported STS's Whole Child Model in the Arusha district with teacher training and research since 2018.

Co-Principal Investigator 2 (Co-PI 2), Parnika Bhatia, Ph.D. led the design and implementation of the research study in collaboration with the Co-PI 1, providing evidence-based expertise to all technical design decisions. She provided technical assistance to the research design remotely led enumerator training for two independent enumerators in Tanzania, led quantitative analysis, and collaborated with the Co-PI 1 on qualitative analysis and report writing. Dr. Bhatia has led numerous impact evaluation studies across Sub-Saharan Africa, including in Malawi, Ethiopia, Sierra Leone, Burundi, Togo, and Guinea-Bissau.

Technical Advisor, Mark Lynd, Ed.D. provided a cost-shared technical support to Co-PIs 1 and 2 in desk review, intervention design, tool development, training of researchers, quality control of analysis and report writing. Dr. Lynd is Co-Founder and President Emeritus of STS and has over 30 years of experience supporting funders, donors, and international education projects in M&E, assessment, teacher training, policy and strategy development, and benchmarking.

Project Background

STS's Prior & Ongoing Work in Arusha, Tanzania

Since 2019, STS has supported eight rural and peri-urban schools in the Arusha Education District through its Whole Child Model, a holistic intervention that aims to improve children's lives and learning through education, health, and community-level interventions. Because STS implemented AG interventions with Whole Child Model schools, STS and the Arusha District Education Office (DEO) determined not to include schools that have received prior interventions in the Tanzania Numeracy Research Project.

Description of the Intervention

On March 22–23, 2025, co-principal investigator Amani Nicolas and two Arusha DEO officers conducted a two-day training for Grade 1–3 teachers and head teachers from the eight schools (n=32). Training consisted of three components:

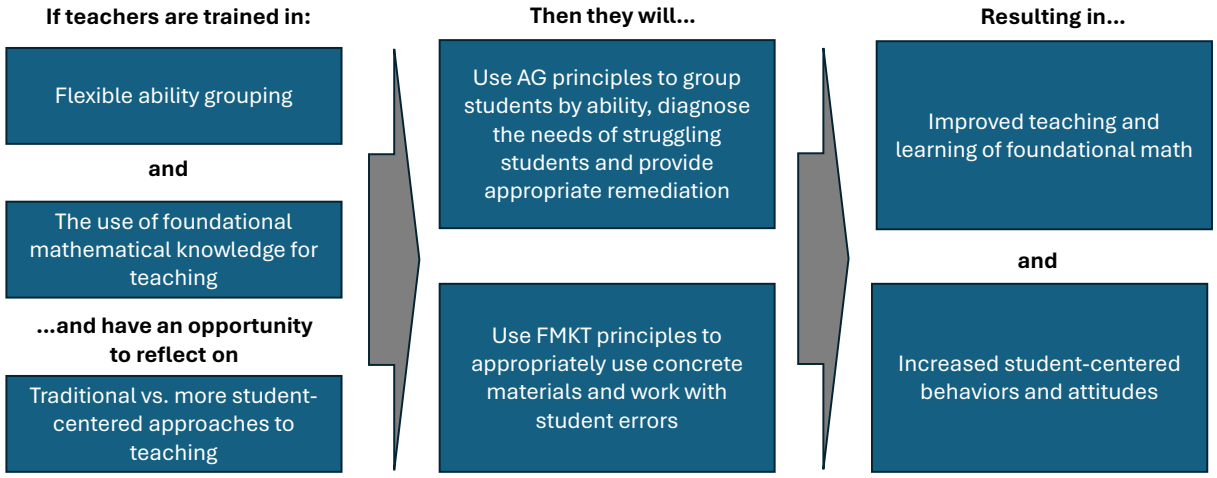
- The purpose and practice of flexible AG, including the use of a mobile-based app to track student assessment results,
- An introduction to FMKT, with a focus on the use of concrete materials and strategies to work with student errors, and
- The importance of using student-centered approaches in the application of AG and FMKT.

From April 8–11, 2025, co-principal investigator Amani Nicolas and the two DEO officers provided follow-up visit support to all Grade 1–3 teachers and head teachers in the participating schools.

Theoretical Underpinning

The intervention design is based on the theory that if teachers receive training and follow-up support in three areas—AG, FMKT, and teaching approaches—then they will successfully apply these approaches, thereby exhibiting greater ability to address the needs of struggling students, effectively teach foundational math, and exhibit student-centered approaches. The theory is depicted in Figure 2.

Figure 2: Theory of Change for Tanzania Numeracy Research Project



This theory of change is predicated on the fact that many teachers in Tanzanian primary schools face two significant problems when teaching mathematics: class size and limited math-specific training.

Over the past several years, STS has worked with Tanzanian teachers to mitigate these problems in two ways. First, STS has trained teachers in the use of flexible AG, a strategy where teachers organize their students into learning groups, then assess and regroup their students based on results. This allows struggling students to receive additional support, even in large classes. Second, STS has trained teachers in the use of best-practice mathematics pedagogy, with a focus on the Concrete-Pictorial-Abstract (CPA) approach in combination with gradual release—"I do, we do, you do."[1] Previous STS-conducted research shows that teachers in supported schools have embraced the AG approach and that students in the early grades of those schools are making progress in mathematics (Lynd & Nicolas, 2023).[2]

Nevertheless, three challenges have arisen. First, teachers have reported that the process of analyzing assessment results and making grouping decisions can be challenging and cumbersome. Second, even when teachers can keep student assessment records, they lack strategies for diagnosing students' difficulties and providing appropriate support. Finally, teachers have requested additional training to reinforce their own mathematical knowledge and teaching skills.

These problems cited are not unique to Tanzania. Across sub-Saharan Africa, teachers can be assigned classes of 100 or more students, especially in rural areas. Faced with these large numbers, teachers report difficulties managing behavior, adapting instruction to individual students' needs, making instruction more interactive, and continuously assessing student

[1] CPA involves moving from concrete materials, to pictorial representations, to abstract symbols and problems whereas "I do, we do, you do" gradually transfers the learning responsibility from the teacher to the student.
[2] Results of Early Grade Mathematics Assessment midline conducted by STS in November 2024. Draft of report in progress.

learning (Adamu, Tsiga, and Zuilkowski, 2022). To cope, teachers resort to strategies such as whole-class instruction, the use of student work at the blackboard, and splitting classes into two streams based on student performance (Naudé and Meier, 2019).

It is in response to this problem that STS has explored its AG approach with its teachers in Tanzania. Monitoring data from STS's work indicates that with this approach, teachers can conduct formative assessment, even in large classes, once per week or more, though tracking individual student progress is challenging. STS's response has been to provide teachers with pre-made assessment trackers—both phone-based and paper-based—where teachers write their students' names once, then fill in a column each time they conduct an assessment. While this strategy has helped, teachers still report it takes time and organization. Additionally, it has only been piloted in classes of 50 students or fewer.

Concerning teachers' need for support in diagnosing students' difficulties and providing appropriate support, the desk review for this project described three approaches to address this issue: Teaching at the Right Level, Luminos's teacher-led assessment system, and Hempel's catch-up program being implemented by RTI in Tanzania. Each of these approaches provides teachers with specific assessment tools to conduct rapid assessments of students' progress in reading or mathematics, a system for assigning them to leveled groups by ability, and level-specific learning activities to conduct with these students. Research on these approaches suggests that while they work well with small class sizes, implementers have not yet found a way to apply them to class sizes of more than 50 students.

One particular issue is that each of these approaches includes a component where students are assessed one-on-one by the teacher or a community facilitator. While this is probably necessary when it comes to diagnosing challenges with struggling students, it may be too time consuming to do with every student every time an assessment is conducted. This limitations supports the possibility that STS's AG approach might suffice as a "first pass" to give teachers, especially in large classes, an initial measure of who is succeeding or struggling in general. Teachers could then conduct rapid one-on-one assessments with struggling students to better understand the nature of each child's challenges.

Concerning teachers' math knowledge and teaching skills, just content knowledge is not enough to be a good math teacher. Instead, to teach math effectively, a teacher needs FMKT, defined as "the knowledge that teachers need to teach mathematics with understanding" (Loewenberg Ball et al., 2008). Following STS's previous numeracy workshops, teachers stated that they understood the value of using concrete materials to explore mathematical notions. However, they indicated that they often lack knowledge concerning the most appropriate materials to use for a given concept, how to ensure students are linking materials to the concept, or how to move students from the concrete stage to the pictorial and abstract stages. In light of these findings, this project focused on two elements of FMKT: the use of concrete materials and working with student errors.

The project centers on implementing student-centered instructional strategies that leverage assessment data to tailor teaching to learners. Ongoing assessments in this approach serve not only to monitor student progress but also to shape daily instructional decisions. A key element of this intervention is treating student errors not simply as mistakes but as powerful opportunities for learning. When errors are openly explored within a supportive classroom environment, they offer insights into student thinking, reveal misconceptions, and inform instructional planning. Moreover, discussing errors collectively helps normalize the learning process, encourages risk-taking, and fosters a growth mindset among students. By making student thinking visible and using assessment to guide instructional decisions, this project aims to foster a more responsive and student-centered learning environment.

Finally, this project is based on the assumption that if teachers successfully use AG and FMKT—that is, if they can use AG to identify and assist struggling students, help students use concrete materials to learn new mathematical concepts, and guide students to understand their errors—then they will exhibit the student-centered approaches envisioned by the Tanzania Numeracy Research Project.

EVALUATION METHODS AND LIMITATIONS

The Tanzania Numeracy Research Project is an exploratory study leveraging a mixed-methods approach. It is defined by collecting, analyzing, and mixing both quantitative and qualitative data in a single study (Creswell & Clark, 2017). Quantitative data includes fidelity of implementation research through classroom observations, use of a teacher attitudes survey, and selected quantitative items from the teacher interview tool. Qualitative data includes head teacher interviews, teacher focus group discussions (FGDs), teacher interviews, as well as selected open-ended items contained within the classroom observation tool. The results of this research can be used to inform two facets:

- Insights around how teachers respond to FMKT and AG principles and their use in large classes, and
- Issues concerning the expansion of AG and FMKT models to other schools.

Sample and Data Collection Tools

The Research Project was conducted in eight primary schools where STS has not previously worked. Schools were selected by the Arusha DEO, which determined that the study should be conducted in the district's schools of highest need. The challenges the selected schools face include their remoteness, high student to teacher ratio, limited infrastructure (e.g., few classrooms), high dropout rates, and low examination pass rates. Three teachers, across Grades 1–3, and one head teacher from each of the selected schools participated in the training, follow-up support, and data collection (n=32: 24 teachers and 8 head teachers).

The Research Project administered five tools:

1) **Teacher attitude survey tool**. The teacher attitude survey tool consisted of the 33 items contained in the validated Mathematical Practices-Teaching Efficacy Belief Instrument (Riggs, 2018).[3] Its focus was on two dimensions of teachers' attitudes toward teaching mathematics: their sense of self-efficacy and outcome expectancy. . STS translated this tool into Kiswahili and piloted it for language and question clarity with two teachers in participating schools.

2) **Classroom observation tool**. The classroom observation tool was designed to be administered during one 30-minute Grade 1–3 math class. It measures both the fidelity of implementation of AG and FMKT processes introduced in training, and it documented teachers' approaches and attitudes—transmission versus student-centered—in the classroom Data were captured through both quantitative and qualitative items. The tool was implemented with all Grade 1–3 teachers in the eight selected schools. To develop these

[3] See Annex III for tool.

tools, STS revised extant AG and classroom observation tools to include elements added in this project.

3) **Teacher interview tool**. The teacher interview tool was designed to be administered after classroom observations by the same enumerator who observed a teacher. It takes 30–45 minutes to administer. It contained pre-set items—both quantitative and qualitative—to track fidelity of AG implementation that go beyond that day's lessons—such as calculation of scores or grouping decisions. The interview tool also captured teachers' perceptions of FMKT strategies and questions arising from classroom observations conducted with the same teachers—for instance, areas where teachers may have struggled with specific steps or chosen to be more directive or facilitative. Finally, the tool contained scenarios that require teachers to choose which approach they would take concerning FMKT and student-centered approaches. The tool was implemented with all Grade 1–3 teachers in the eight selected schools. To develop this tool, STS revised extant AG interview tools and added FMKT elements.

4) **Focus group discussion tool**. The FGD tool was designed to be administered during a school visit with all Grade 1–3 teachers who had been observed and interviewed, as well as the head teacher. It takes approximately 45 minutes to administer. As with the teacher interview tool, the FGD tool contained items focusing on teachers' use of AG and FMKT strategies, patterns observed during classroom observations, and viewpoints on issues of self-efficacy, outcome expectancy, and student-centered teaching. While the tool contained several "set" items, space was left for the enumerators to add questions or observations based on what had surfaced in classroom observations and teacher interviews.

5) **Head teacher interview tool**. The head teacher interview tool was designed to provide demographic data, such as the school's size. It also captured patterns observed by the head teacher, including observations about training that the head teacher attended, the role the head teacher plays in supporting teaching strategies, and successes and challenges implementing AG and FMKT across teachers. It takes approximately 30 minutes to administer and was administered to the head teacher at each school.

District Collaboration and Institutional Review Board Approval

STS presented the Research Project design and tools to the Arusha DEO. STS was also required to obtain approval from officials in Dodoma, Tanzania—including the Office of the President (known as *TAMISEMI*) and the Regional Education Secretary. Once approval was received, the Arusha DEO selected eight schools to participate in the Research Project, which they identified as schools with the highest need.

School Selection

In selecting schools for study, STS proposed to include both urban and rural schools, a mix of large and small class sizes, and a range of higher- and lower-performing schools. The goal was to ensure that the study's sample reflected the variety of educational environments found across the district. However, instead of following these criteria, the DEO identified eight schools based solely on their designation as high needs. This selection process ultimately shaped outcomes. Most notably, it limited STS research team's ability to compare different contexts and may have concentrated findings on challenges unique to higher-need schools.

Data Collection

STS hired two independent enumerators who were not involved in the Research Project's intervention to collect data in the eight schools.[4] STS programmed the data collection tools into SurveyCTO software, which the enumerators loaded on tablets for electronic data collection. On May 14–15, 2025, the Research Project's co-principal investigator Dr. Parnika Bhatia remotely trained two Arusha-based enumerators on five topics:

- Introduction to the Tanzania Numeracy Research Project
- Overview of sample and data collection schedule
- Overview of data collection tools
- Ethics and best practices
- Overview and administration practice of each tool

From May 16–27, approximately two months after teacher-training, the enumerators visited each of the eight schools to administer the data collection tools (Table 1).

Both enumerators conducted classroom observations simultaneously, each utilizing the same tool and recording their observations independently. After observing the classes, the enumerators discussed the patterns they observed and the questions that arose, which they added to the pre-set questions on the teacher interview and FGD tools.

All tools were administered using a tablet, except the teacher attitude survey, which was paper-based. Enumerators took notes on both tablet and paper. Each evening, each enumerator checked and entered any remaining data from their paper forms into SurveyCTO. One member of the team entered the FGD data, then the other member reviewed it for completeness and accuracy, discussing or revising where necessary. Once entered, all data were uploaded to a secure server for analysts to review. Each day of data collection, the principal investigator in-

[4] To ensure that implementers do not also play the role of evaluator, this project employed two principal investigators. The first co-principal investigator, Amani Nicolas, served as the lead trainer/facilitator. In this role, he worked to strengthen the capacity of district officers who served as co-facilitators of training and follow-up support visits to schools. For the data collection phase, co-principal investigator Dr. Parnika Bhatia led the recruitment and training of the enumerators. These enumerators did not play a role in training or follow-up support. One of the enumerators, Esther Lally, had been an implementer of STS interventions focusing on other populations, including parents and girls, so the threat of implementer bias with her was judged to be minimal.

country monitored data collection by phone or WhatsApp, then completed a data tracking form and sent it to STS's headquarters for review. The data monitoring role was strictly limited to tracking participant counts and promptly contacting enumerators if any data or participant information was missing

Table 1: Final Sample Size and Response Rates

Tool	Target	Sample	Response Rate
Classroom Observation	24	20 (average class size 72)	83%
Teacher Interview	24	21 (F:19; M:2)	87%
Teacher FGD	8	8	100%
Teacher Attitude Survey	24	20 (F:17; M:3)	83%
Head Teacher Interview	8	5 (F: 3; M:2)	63%

Analysis

STS analyzed the data using frequency tables for quantitative items on all tools, and crosstabs to present data by teacher, sex, class, and school. STS examined qualitative data for patterns and themes, then triangulated findings from quantitative data as well as findings from other tools. Additionally, classroom observation data were analyzed to identify rates of interrater reliability on quantitative items and fidelity of implementation for AG. STS interpreted these results in light of findings from the other tools.

Limitations

Limited scope. The Tanzania Numeracy Research Project was implemented over a limited, seven-month timeframe—January to July 2025—which places limitations on the ability of teachers and schools to achieve a high level of uptake of the new strategies they were trained on. In addition, due to the Research Project's small sample, it was not possible to conduct inferential analyses—correlations, prediction, and causation. Hence, data from this study should be viewed as exploratory, using descriptive analysis strategies only.

Sample size. The sample size was relatively small and purposively selected. Findings are not representative of the broader population and cannot be generalized beyond this group of participants.

No comparison group. The analysis focused exclusively on the treatment group, with no comparison or control group included. As a result, it was not possible to determine how student outcomes compare to those who did not participate in the program or to typical educational experiences.

Experimental design. There was no external baseline available against which to measure change, making it difficult to attribute observed improvements solely to the intervention rather than to other external or maturational factors.

These limitations should be taken into consideration when interpreting the results of this study.

FINDINGS

Ability Grouping

To address the first research question—"Ability Grouping: How well can teachers implement flexible AG during numeracy instruction?"—the study examined several key aspects of teacher practice. Specifically, the project team explored the following:

- **Fidelity of Implementation**: Do teachers implement AG as intended, following recommended practices for flexible grouping during numeracy instruction?
- **Instructional Adjustment**: Do teachers adjust their instructional approaches based on the AG results?
- **Group Mobility**: Do teachers move students between groups according to students' ongoing progress and learning?

By analyzing these components, the project team aimed to understand both the fidelity and consistency with which flexible AG is integrated into classroom instruction. Overall, the implementation of AG practices across classrooms showed moderate success. Both observational and interview data highlight a modest but promising uptake among teachers.

Fidelity was measured across four key steps outlined in the observation tool:

1. Whole-class instruction,
2. Formation of student groups,
3. Teacher monitoring and feedback during group work, and
4. Assessment and regrouping based on performance.

Most teachers demonstrated some level of implementation, but few applied all components consistently. While nearly all teachers provided whole-class instruction and formed groups, the more advanced steps—structured group-based assessment and using trackers for regrouping—were less commonly implemented. About 30 percent of teachers demonstrated strong fidelity, consistently following all the components. An additional 60 percent of teachers implemented all steps but did not regularly use the trackers, indicating moderate fidelity (Table 2). Classroom observations showed that most classrooms featured heterogeneous groups,[5] and teachers often circulated to provide individual or small-group support. Only one classroom had no visible

[5] The use of heterogeneous versus homogeneous (ability-based) grouping signals where the teacher is in the AG cycle. If they are still using heterogeneous grouping, it might mean: they are introducing a new concept (so no prior assessment data available), or They have not yet used assessment to reorganize groups. In contrast, homogeneous grouping (ability-based) suggests teachers assessed students previously and regrouped based on performance—a stronger indicator of fidelity to the AG model.

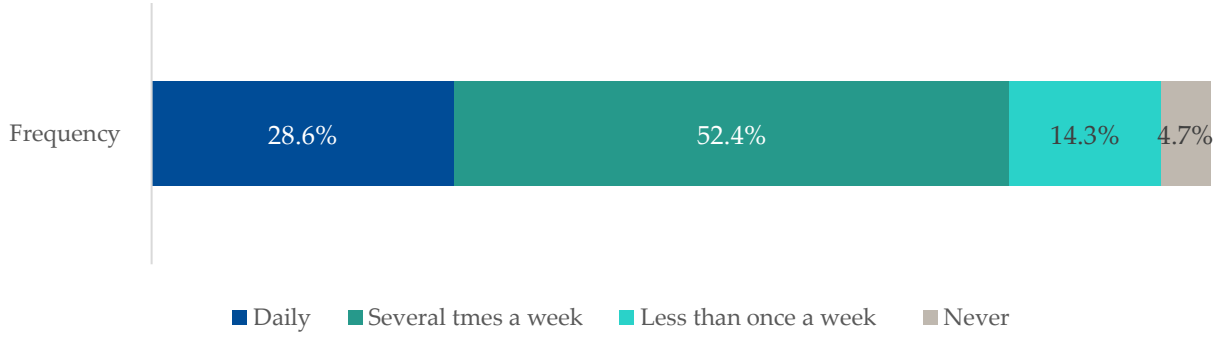
grouping; in this case, the teacher reported a lack of space and classroom management difficulties.

Table 2: Fidelity of Ability Grouping Implementation

Tool	Components	Percentage
Classroom Observation	Step 1: Conducted whole-class instruction	100%
	Step 2: Formed groups after instruction	93%
	– Used heterogeneous grouping	98%
	– Used two-level grouping	2%
	Step 3: The Teacher circulated during group work	98%
	– Provided feedback to individuals	45%
	– Provided feedback to groups	53%
Teacher Interview	Step 4: Implemented group-based assessment	80%
	Used the tracker to regroup students	29% (6 of 21)

Most of the teachers (17 out of 21) reported regrouping students based on assessment results, either several times a week or daily. However, four teachers reported that grouping did not occur regularly after assessment, indicating partial adoption of AG practices. Teachers who did not engage in grouping cited challenges, including difficulty using the trackers, technological barriers, absenteeism, low student engagement, classroom space constraints, and the limited ability of some students to engage with math learning due to reading struggles.

Figure 3: Teacher Interview Results: How Often Do You Group Students Based on Assessment Results?



Several strategies were reported to support student needs among teachers using trackers regularly (6 of 21). Using the tracker helps provide a systematic way for teachers to follow student progress over time. As a result, these teachers could speak specifically about which students were struggling, the nature of their mathematical challenges, and how they were responding. The strategies included pairing weaker students with stronger peers, providing

immediate feedback while circulating the room, and arranging seating to foster peer support. For instance, one teacher reported, "For those who are struggling, I usually bring together the higher-ability students with those of lower ability and stay close to them to guide them more." Another teacher noted, "I move around the group, and when I meet students who are struggling, I choose their peers to help them."

Findings from FGDs reinforced the challenges of the AG process that teachers noted in their interviews. These were difficulties using the AG tracker, infrastructure limitations, and inconsistent student collaboration. Nonetheless, teachers emphasized positive outcomes like increased student engagement, confidence, and attendance due to group-based learning. It is essential to recognize that, while group work offers valuable opportunities for peer interaction, shared problem-solving, and collaborative learning, its benefits are distinct from those provided by assessment-based ability-grouping. Without ongoing assessment and intentional regrouping based on student progress, the specific learning needs of individual students may not be sufficiently addressed. Therefore, the use of trackers to guide AG and targeted support plays a crucial role in ensuring that all students receive appropriate challenges and interventions. Further, teachers also expressed a continued need for support and training on using the tracker to log assessment data effectively for regrouping, along with improvements in classroom resources and space.

Head teachers reported a positive trend in the uptake of grouping practices, especially in lower grades. They observed improved classroom management and more structured lessons due to grouping. All five head teachers noted increased teacher effort, praising staff for adapting and trying new strategies despite large class sizes and limited resources. They also reported challenges to AG similar to those mentioned by teachers, such as difficulty managing the behavior in groups, lack of space and infrastructure, and overcrowded classrooms—average class size 72.

FMKT: Use of Concrete Materials and Addressing Student Errors Effectively

To explore the research focus on FMKT, the study examined the extent to which teachers used selected elements of FMKT in their numeracy instruction. Specifically, the project team looked at if teachers:

- Used concrete materials appropriately to support student understanding and engagement,
- Actively sought to understand the reasons behind students' mathematical errors, and
- Effectively addressed students' errors through targeted instructional responses.

Overall, classroom observations, teacher interviews, and head teacher interviews indicate that the use of concrete materials is a widespread and valued component of math instruction across classrooms. Additionally, findings from the classroom observation and teacher interviews note

that the majority of teachers identify student errors and demonstrate key strategies to help students with problem-solving.

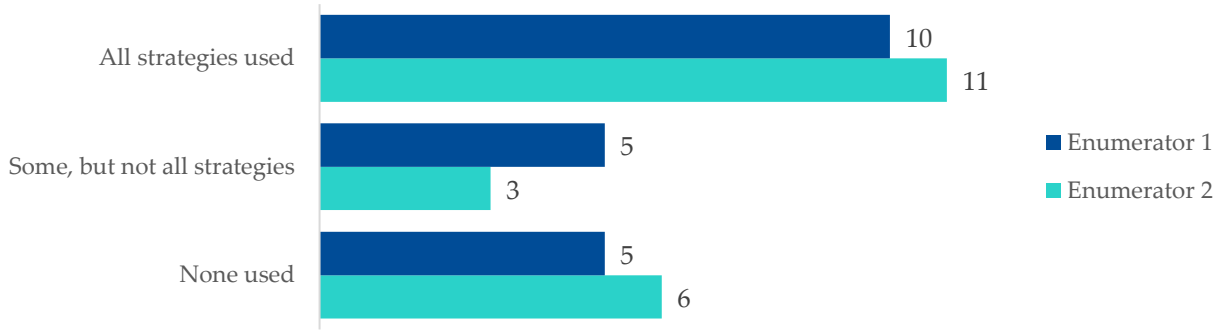
Concrete materials—such as bottle caps, sticks, corn lefts, grains, lemons, and stones–were used in all observed math lessons. In most classrooms, teachers introduced these materials by modeling their use before allowing students to explore them independently, a technique that was prioritized in the training. In three classrooms, however, teachers opted for diagrams or pictorial methods rather than hands-on tools to demonstrate the concepts. Across all classrooms, enumerators noted that students used the materials appropriately.

Teachers identified student errors or misconceptions during activities, with only two classrooms showing limited evidence of this practice. In 72 percent of the 24 observed classrooms, teachers demonstrated three of the five key instructional strategies presented in the training: asking students to explain their reasoning, describe their problem-solving steps, and explore alternative solutions.

Table 3: FMKT Strategies Implementation

Tool	Strategy	**Percentage**
Classroom Observation	Usage of concrete materials in math lessons	100%
	The teacher demonstrates the usage of concrete materials	85%
	The teacher asks open-ended questions when using concrete materials	100%
	The teacher provides students with opportunities for hands-on exploration	100%
	The teacher uses students' mistakes as a learning opportunity for the whole class	85%
Teacher Interviews	How often do you use concrete materials in math lessons?	86% (Every Lesson)
	Ask students to explain the reasoning behind their answers	95%
	Use a student's mistake as a learning opportunity for the whole class	91%
	Encourage students to think of multiple ways to solve a problem	81%

Figure 4: Reasoning Strategies Used by Teachers During Classroom Observation



In Figure 4, three items were combined to construct the reasoning strategies scale, which demonstrated high internal consistency ($\alpha = 0.93$). The items were: (1) The teacher asks students to explain their reasoning when providing an answer; (2) The teacher prompts students to describe the steps they used to reach their answers; and (3) The teacher encourages students to explore alternative solutions or different approaches.

Self-reported teacher interview data further confirmed high alignment of teacher practices with FMKT principles. All teachers said they use concrete materials and ask students to explain their reasoning in every lesson. Among those using concrete materials consistently, the majority (15 out of 18) also encouraged students to identify multiple strategies for solving problems. This suggests that teachers understood the importance of hands-on learning and instructional strategies that promote flexible, exploratory thinking. Only a small number of teachers (3 out of 18) did not report the usage of alternative methods for solving problems.

FGDs highlighted several challenges teachers face when using concrete materials. These include overcrowded desks that limit space, language barriers with early students not fluent in Kiswahili, and practical issues like materials being too heavy, dirty, or damaging to students' belongings. One teacher noted that "some tools, like stones and sticks, damage notebooks," while another pointed to overcrowded desks and behavior management challenges, saying, "Students often mix up numbers when counting because of the small desktop and fights over tools." Despite these constraints, teachers strongly preferred beginning lessons with hands-on materials followed by class discussions, believing this approach better engages students and supports peer learning and conceptual understanding. As teachers shared, "Both giving tools and discussing are important, but starting with hands-on practice generates more interest," "Provide materials and let them discuss before moving on—it helps students grasp the concept easily," and "It's better to give them tools first than jump straight to abstract explanation."

All five head teachers noted a growing incorporation of FMKT strategies, specifically the usage of concrete materials among teachers. While the implementation varies, most headteachers observed that usage of concrete materials is being applied "to some extent" or "with clear effort." In two schools , teachers were reportedly integrating real-world objects into their

instruction and adapting lessons accordingly. At another school, one school leader highlighted that students now bring materials from home, while another head teacher observed that the FMKT approach, related to the usage of concrete materials, was particularly effective with younger students. The most cited barriers to full implementation were limited access to clean concrete materials and time constraints. Head teachers noted that hands-on tools are not always available, limiting consistent practice. Additionally, some teachers struggle with preparation time and adapting strategies to students who are still adjusting to a new style of instruction. Despite these obstacles, head teachers consistently recognized that their teachers were motivated to use the concrete materials and the effort in applying these math strategies in their classrooms.

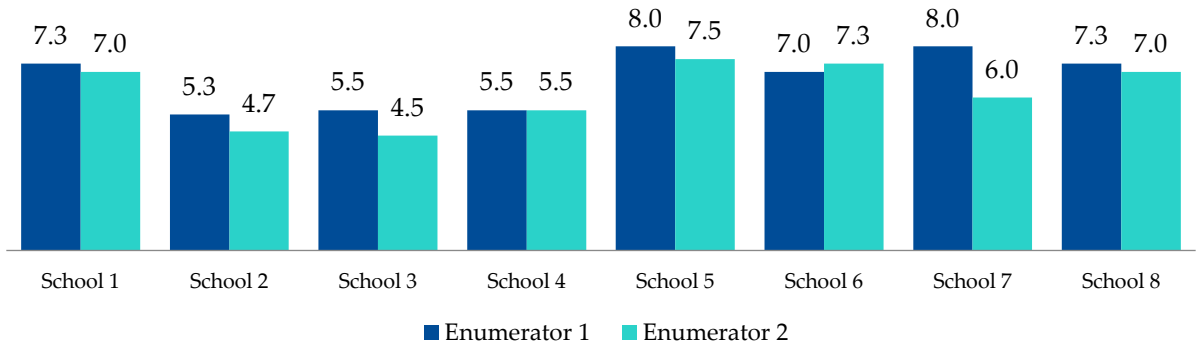
Attitudes: Teachers' Views of Their Role

This section presents the findings organized around the key research questions: How do teachers view their role? Do teachers see their role as transmitters of knowledge or facilitators of learning? How does the use of AG and FMKT principles influence teachers' views of their role?

Classroom observations, teacher attitude survey, and teacher interviews were analyzed to address the research question. At the end of classroom observation, enumerators were asked to rate teachers' approach on a scale of 1–10, 10 being most student-centered and 1 being most teacher-centered instruction. These ratings were based on several key factors: how teachers delivered instruction to whole classroom and smaller groups, actively monitored student learning through ongoing assessments, used student errors as learning opportunities, and encouraged students to engage in reasoning strategies—such as exploring alternative solutions and explaining their thinking when answering questions. On average, both enumerators rated teachers an average of 7 out of 10, indicating that many teachers lead their classrooms more as facilitators than as only transmitters of knowledge.

Ratings across classrooms ranged from 4 to 9, suggesting variation in practice. Several classrooms scored particularly high (above 8), reflecting clear examples of student-centered instruction. One school stood out for having multiple high-scoring classrooms (3 of 3), suggesting a school-wide emphasis on facilitation over direct instruction. Importantly, enumerator ratings on the classroom observation tool showed high consistency for all items, signaling strong inter-rater reliability (Figure 5). This reinforces the credibility of the observed trends in teacher practice.

Figure 5: Mean Facilitator Scores per School



The data show a wide range of teacher self-efficacy scores for teaching mathematics, as measured by the Mathematics Teaching Efficacy Beliefs Instrument (MTEBI) scale. This study used the validated scale to measure teachers' level of self-efficacy in teaching mathematics and an understanding of how their students learn math, which may influence their approach to teaching math. Most teachers expressed moderate to high confidence in their instructional abilities, with Grade 3 teachers reporting the highest overall self-efficacy in teaching mathematics (scores ranging from 78 to 97 out of 100). Teachers felt most confident in some FMKT-aligned practices, such as preparing students to understand mathematical units in context, responding to student questions meaningfully, supporting students from disadvantaged backgrounds, helping students work independently, and justifying their reasoning. Teachers shared lower confidence in more advanced pedagogical tasks, such as allowing students to struggle with challenging problems, analyzing and synthesizing different student strategies, and supporting students who do not speak Kiswahili as a first language. These gaps indicate areas where continued professional development would be beneficial.

Teacher beliefs about how students learn math were mixed. Understanding teachers' beliefs about how students learn math is crucial because these beliefs shape their instructional choices, influence how they respond to student needs, and ultimately affect whether their instruction will be teacher-centered or student-centered. Scores on the "feelings about how students learn math" scale ranged between 27 and 50 out of 60, reflecting variation in their beliefs across grades and schools. Despite this spread, most teachers strongly agreed with statements such as: "students from impoverished homes can succeed with proper support," "low-motivation students can be engaged through good teaching," and "it is the teacher's responsibility to ensure student understanding regardless of background." These responses demonstrate that teachers believe all students can learn and that teachers are responsible for helping them learn—a central student-centered pedagogical principle. This finding also reinforces observers' ranking of teachers as mostly at the facilitator end of the continuum.

Teachers' views on their professional role varied. Of the 21 teachers surveyed, six identified themselves as "facilitators of learning," while 12 saw themselves more as "transmitters of

knowledge." Three described their roles as blending both identities. FGDs echoed this divide, revealing mixed perceptions among teachers about their instructional roles.

One potential explanation for the differences observed between teachers' self-perceptions and enumerator reflections is language-related challenges. Specifically, teacher training was conducted in Kiswahili, whereas interviews and evaluation tools were administered in English. This discrepancy may have influenced teachers' understanding of terms like "transmitter" and "facilitator of learning," contributing to variation in responses and observer ratings. Despite this variation, all teachers agreed that both the AG and FMKT approaches have made their teaching "easier" and "more effective," highlighting the perceived value of these methods.

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

The Tanzania Numeracy Pilot Research Project demonstrates that combining AG and FMKT has strong potential to strengthen early-grade mathematics instruction. The findings suggest that while foundational steps of AG—such as forming groups—are being adopted in selected schools in Arusha, the more advanced practices of dynamic regrouping based on assessment and systematic use of trackers remain limited. In contrast, FMKT practices—particularly the use of concrete materials and the promotion of reasoning—showed higher uptake, pointing to teachers' readiness to implement student-centered strategies when appropriate resources and guidance are provided. Teacher attitudes further reinforce this potential. Teachers expressed openness to facilitation-oriented practices. However, their confidence in managing more complex instructional shifts, such as flexible regrouping or multiple solution strategies, was mixed.

The findings of this pilot project indicate modest adoption of AG (roughly 30 percent) and FMKT strategies (roughly 80 percent) in high-needs primary classrooms in Tanzania. Teachers demonstrated effort and intent, as reported by the head teachers, with about 90 percent actively adjusting their instruction based on student learning, utilizing group-based activities and peer support. Head teachers noted that grouping practices contributed to structured lessons and supported student engagement.

Despite these positive beliefs and intentions, there were some gaps in implementation and monitoring. The fidelity of AG implementation drops significantly when it comes to assessment-driven regrouping and structured use of trackers. While heterogeneous grouping aligns with the initial phase of AG, the low incidence of two-level grouping suggests that many teachers are not yet fully applying differentiation based on assessment data. These patterns highlight the need for continued training on using assessment results to regroup students, simplifying tracking tools to strengthen implementation of the complete AG cycle, and providing teachers with a list of ideas that can help struggling students in their groups.

The evidence of student regrouping based on ongoing assessment was limited; only six of 21 teachers reported using the tracker. This limitation is partly because only one classroom observation was conducted per class, preventing observation of if students were regrouped differently in subsequent lessons. Additionally, enumerators did not directly verify the use of trackers reported by six of the 21 teachers during interviews.

Persistent barriers to fidelity included overcrowded classrooms and inadequate furniture, which constrained the organization of multiple groups. Additionally, many teachers struggled to manage student behavior in groups effectively, indicating a need for further training and support.

Teacher self-efficacy was generally moderate to high for foundational practices, especially among lower-grade teachers. However, support is still needed for advanced strategies, like allowing students to struggle with challenging problems, analyzing and synthesizing different student strategies, and supporting students who do not speak Kiswahili. Additional supports should include classroom adaptation and serving linguistically diverse students. Both school leaders and teachers acknowledge that AG and FMKT approaches hold significant promise for improved teaching and learning, but emphasize that the continued investment in teacher preparation, material resources, and infrastructure remains essential.

Recommendations

Maximize Use of Concrete, Locally Available Materials through Community Involvement: Teachers consistently reported that using local, hands-on materials made mathematics lessons more engaging and accessible for students. Head teachers observed increased engagement and deeper learning as a result of these practices. To further strengthen this approach, head teachers can actively involve the wider community and parents in sourcing and preparing everyday objects for classroom use. In two schools, students have already begun bringing materials from home, showcasing the potential for strong community participation. Involving families not only reinforces the value of education but can also help address specific challenges—such as overcrowded desks or the use of dirty materials—by promoting better organization, the development of complex materials, and cleaner resources. Encouraging teachers to use community-sourced materials can be especially effective in high-needs, low-resource settings.

Engage Head Teachers for Continuous Improvement and Peer Learning: During implementation, the project provided follow-up mentoring to head teachers to encourage classroom observations and coach teachers in the classrooms. Head teachers were provided with a list of suggested actions as well as a timeframe to conduct the monitoring and revert to the trainer with their observations. Positive head teacher involvement boosted teacher effort and implementation, especially in supporting flexible grouping and concrete material use. Monitoring and supporting head teachers through regular classroom walk-throughs focused on AG and FMKT strategies could help them develop skills in constructive feedback and facilitate

peer sharing among their teachers. Additionally, head teachers reported increased motivation, effort, and willingness of their teachers to try new techniques. Regularly sharing stories and evidence of impact at staff meetings and community events, or even a simple recognition system—such as an "innovative teacher" award or sharing lesson success stories—could help build momentum and collective pride.

Rigorous, Sustained Training and Support: Teachers reported positive changes in their classroom practices following AG and FMKT-focused training and expressed a desire for longer-term professional development and support. Implementers should provide ongoing, context-specific training that deepens teachers' understanding of both the "how" and the "why" of AG and FMKT. This training should include practical strategies for managing large classes, implementing assessment-based flexible grouping, and addressing behavior management. Further, establishing professional learning communities would allow teachers to share successful strategies, observe each other's lessons, and collaboratively address challenges, ensuring continuous improvement and peer support.

Design and Data Collection Tools: Before future data collection, researchers should pilot all study instruments within the specific local context to ensure their appropriateness, reliability, and cultural relevance. While pilot testing of the MTEBI and FMKT strategies allowed for effective adaptation to local contexts, time and budget limitations meant that other data collection tools could not be piloted. Conducting pilot testing allows researchers to identify and address potential issues related to clarity, language, and contextual suitability. Additionally, planning for multiple observation time points enables more robust triangulation of findings and accurate assessment of implementation fidelity, particularly for dynamic practices such as AG. This approach would strengthen the validity and reliability of future research findings.

Expansion to other schools. Given the high level of enthusiasm for AG and FMKT strategies resulting from a relatively modest level of intervention, this intervention could prove to be a low-cost way to transform classrooms in Tanzania. The model should be expanded to include a larger number of schools, with the intent to study its effect using pre/post-measures in contrast to comparison schools. In particular, future research should focus on strategies to aid teachers in tracking assessment results, regrouping students based on results, and providing appropriate support, especially to struggling students.

ANNEXES

Annex I: Works Cited

Adamu, Amina, Aisha Tsiga, and Stephanie Zuilkowski, "Teaching reading in northern Nigeria: the challenges of large class size," Pedagogy, Culture & Society (30). https://doi.org/10.1080/14681366.2020.1794948. (Original work published 2022).

Ciccarelli, T. (2024). 3 ways to use flexible grouping in real time to support student growth. https://www.nwea.org/blog/2024/3-ways-to-use-flexible-grouping-in-real-time-to-support-student-growth/

Creswell, John W. and Vicki L. Plano Clark (2017), Designing and Conducting Mixed Methods Research. 3rd ed. (Thousand Oaks, CA: SAGE Publications, 2017), https://books.google.com/books/about/Designing_and_Conducting_Mixed_Methods_R.html?id=eTwmDwAAQBAJ&redir_esc=y

Loewenberg Ball, Deborah, Mark Hoover Thames, and Geoffrey Phelps. (2008). "Content Knowledge for Teaching: What Makes It Special?" Journal of Teacher Education, 59(5), 389-407. https://doi.org/10.1177/0022487108324554 (Original work published 2008).

Lynd, Mark and Amani Nicolas. "Ability Grouping to Improve Teaching and Learning in Large Classes: An Experience in Six Schools in Tanzania." Unpublished manuscript, 2023.

Morin, A. (2024). *Flexible grouping: What you need to know*. https://www.understood.org/en/articles/what-is-flexible-grouping

Naudé, Mariana, Corinne Meier, and Linda Bosman (2019), Teaching Foundation Phase Mathematics: A Guide for South African Students and Teachers (Van Schaik Publishers, 2019), https://books.google.com/books?id=CPHOwQEACAAJ&redir_esc=y.

Riggs, Iris M., et al (2018), "Measuring Teachers' Beliefs in Relation to Teaching Mathematics with Mathematical Practices in Mind," School Science and Mathematics 118(8), http://dx.doi.org/10.1111/ssm.12303.

Annex II: Teacher Checklist

Extra Handout: Concrete Materials Selection Checklist

School-to-School International

Criteria	Key Questions / Considerations	Yes/No/Partial	Comments / Notes
1. Clear Learning Goal	- Have I identified the specific math concept (e.g., fractions, place value, geometry)? - Does the manipulative directly address that concept?		
2. Conceptual Fit	- Does the structure of the material mirror the math concept (e.g., fraction tiles to illustrate fractional parts; base-ten blocks for place value)? - Will it help students visualize or model the key idea?		
3. Age & Cognitive Appropriateness	- Is the manipulative developmentally appropriate for this grade/age? - Could it be too abstract for younger students or too simplistic for older students?		
4. Prior Knowledge / Familiarity	- Are students familiar with these materials already? - If not, do I have time to introduce how they work before using them for the concept? - Does this build on previous lessons or known experiences?		
5. Clarity & Ease of Use	- Is the manipulative straightforward (e.g., minimal pieces, simple design)? - Will students be confused by extra features (colors, shapes) that do not serve the math concept? - Do I need extensive setup?		
6. Transition to Abstract	- Can this tool easily support a Concrete → Representational → Abstract progression (CRA)? - Will I be able to show students how the physical model translates into symbols or equations?		

Criteria	Key Questions / Considerations	Yes/No/Partial	Comments / Notes
7. Realistic & Relevant	- Does the material reflect real-life or authentic contexts (e.g., coins for money, measuring cups for fractions) where possible? - Is it something that resonates with students' everyday experiences?		
8. Potential for Engagement & Depth	- Will the manipulative facilitate hands-on exploration and discussion? - Can I pose open-ended tasks that go beyond just rote procedures? - Can students connect it to deeper or related concepts?		
9. Feasibility & Availability	- Do I have enough sets for all students or small groups? - Is it cost-effective, or can I use inexpensive/no-cost alternatives (e.g., bottle caps for counters)?		
10. Teacher Comfort & Knowledge	- Am I confident in modeling and guiding students with this tool? - Do I understand common misconceptions that might arise when using it? - Do I have a plan for introducing and managing it in class?		
11. Research / Best Practices	- Does research or curriculum guides recommend this manipulative for the concept (e.g., recommended by NCTM, or included in textbook resources)? - Have colleagues found it effective in similar contexts?		
12. Assessment & Reflection	- How will I assess whether students are truly understanding the concept when using this material? - Will I gather formative feedback (e.g., exit slips, quick tasks) to see if the manipulative helps?		

Annex III: Research Tools

Teacher Attitudes Survey

Note: Administered in Kiswahili with English translation provided below.

Introduction

Thank you for participating in this survey. We are interested in your beliefs and confidence about teaching mathematics. Your answers will help us improve professional development programs and support for math teachers. There are no right or wrong answers; we want your honest opinions. Your responses are confidential.

School: _____

Date: _____

Sex: __ m __ f

Class you teach (circle):		1	2	3	4	5	6	7

Please indicate the degree to which you **agree** or **disagree** with each statement by circling the answer the most closely matches how you feel. Choose only one answer to each item using the following scale:

1 = Strongly Disagree | 2 = Disagree | 3 = Uncertain | 4 = Agree | 5 = Strongly Agree

Question	STRONGLY DISAGREE 1	DISAGREE 2	UNCERTAIN 3	AGREE 4	STRONGLY AGREE 5
Section A: Feelings about teaching mathematics.					
1. I know how to prepare students to consider the meanings of units used in different contexts.	1	2	3	4	5
2. When a student commits an error in math, I am able to diagnose his/her conceptual errors.	1	2	3	4	5
3. I understand math concepts well enough to be effective in teaching math.	1	2	3	4	5
4. I feel comfortable addressing students' questions about mathematical concepts and ideas.	1	2	3	4	5
5. I am comfortable allowing my students to make their own approximations or simplifications when approaching a real-life problem.	1	2	3	4	5

Question	STRONGLY DISAGREE 1	DISAGREE 2	UNCERTAIN 3	AGREE 4	STRONGLY AGREE 5
6. I can easily integrate students' strategies and ideas into my math lessons even if they are different from my lesson plan.	1	2	3	4	5
7. I know how to prepare students to plan their own approaches to solving problems.	1	2	3	4	5
8. I can develop students' ability to produce mathematics (e.g., a number sentence, expression, or equation) to model their own interpretation of a situation.	1	2	3	4	5
9. I am able to make sure my students can use materials to represent problems in multiple ways.	1	2	3	4	5
10. I can incorporate multiple representations into my lessons to improve student learning.	1	2	3	4	5
11. I can help students learn to work on their own to gather appropriate evidence to support their mathematical ideas.	1	2	3	4	5
12. I can teach students to determine on their own which situations require an exact answer and which require an estimate.	1	2	3	4	5
13. I can teach my students to decompose and recombine numbers and expressions in different ways depending on the context.	1	2	3	4	5
14. I feel comfortable teaching students to understand relationships between concepts of algebra and concepts of arithmetic.	1	2	3	4	5
15. I am comfortable helping my learners who do not speak Kiswahili as their first language gain conceptual understanding of mathematics.	1	2	3	4	5
16. I am comfortable letting my students struggle with a problem for which there is no immediately obvious method of solution.	1	2	3	4	5
17. I can help students learn to see relationships between quantities.	1	2	3	4	5
18. I can teach students to make a habit of asking themselves whether their work makes sense.	1	2	3	4	5
19. I am comfortable analyzing and synthesizing different student approaches	1	2	3	4	5

Question	STRONGLY DISAGREE 1	DISAGREE 2	UNCERTAIN 3	AGREE 4	STRONGLY AGREE 5
to a mathematics problem to bring closure to a mathematical discussion.					
20. I know how to develop students' ability to use the math they know to solve problems in everyday life.	1	2	3	4	5
21. I am able to help students from impoverished backgrounds excel in math.	1	2	3	4	5
Section B: Feelings about how learners learn math					
22. Regardless of the teacher's instruction, students won't use available tools to investigate problems on their own.	1	2	3	4	5
23. Increased effort in math teaching produces little change in some students' math achievement.	1	2	3	4	5
24. Even a very skilled teacher cannot expect learners who speak Massai as their first language to attempt to understand complex mathematics problems.	1	2	3	4	5
25. No matter how skilled the teacher, some students won't understand what quantities mean, even if they can compute them.	1	2	3	4	5
26. Even a teacher with good math teaching abilities may not help some students learn math.	1	2	3	4	5
27. A teacher can be expected to help a student learn math despite his or her impoverished home background.	1	2	3	4	5
28. Students who have low motivation for learning math can be turned on to learning by their math teachers.	1	2	3	4	5
29. No matter what the teacher does, students can't seem to determine when an approximate answer is appropriate.	1	2	3	4	5
30. Even with appropriate instruction, most students rarely consider whether their math work makes sense.	1	2	3	4	5
31. Students at my grade level think concretely, and teachers can't be expected to teach them to work with abstractions in mathematics.	1	2	3	4	5
32. When students are given the opportunity to make their own generalizations, they end up more confused than if the teacher teaches mathematics directly.	1	2	3	4	5

Question	STRONGLY DISAGREE 1	DISAGREE 2	UNCERTAIN 3	AGREE 4	STRONGLY AGREE 5
33. Seeing many different approaches to solve one problem confuses many students and hinders their learning.	1	2	3	4	5

Thank you!

Classroom Observation

Introduction & Consent: The purpose of this tool is to track whether teachers implement flexible AG and use selected elements of FMKT during numeracy instruction.

Before the observation, explain the purpose of the observation as follows:
We are observing you today so we can learn about how teachers are using AG and FMKT strategies in the classroom. Your participation is voluntary, and all responses will remain confidential. Do you agree to participate?

If yes, continue. If no, thank the teacher and leave the classroom.

Background Information

Name of the observer:

Date:

Name of School:

Name of Teacher:

Gender of Teacher:

Class:

Number of Students (male and female):

The objective of the lessons:

Time lesson starts:

Time lesson ends:

Total School Day Time by Hours and Minutes:

Instructions: Tick "yes" for each strategy observed during the lesson. Where possible, describe the strategy used by the teacher.

1. Teacher states or writes the objective of the lesson
    a. If yes, what is the objective of the lesson?

Concrete-Pictorial-Abstract (CPA) Approach: Focus on Concrete

2. The teacher uses concrete materials/manipulatives during the lesson.
   If no, go to question 9.
3. The teacher uses concrete materials/manipulatives during the lesson.
   If yes, go to question 3a.
    a. If yes, what is the material used?
    b. If yes, are resources local? Describe.

4. The teacher demonstrates the use of material before allowing students to explore them (explicit modeling)
    a. If yes: What was the model example?

5. The teacher asks open-ended questions when using concrete materials to engage students.
    a. If yes, give examples:

6. The teacher provides students with opportunities for hands-on exploration in pairs and groups.
    a. If yes, is the task a different example than the model example?
    b. If yes, what is the practice example?

7. The teacher moves around the class and works with groups to observe and assess student interaction with the material.

8. Students have the opportunity to use concrete materials during the lesson in groups/pairs or individually.
    a. If not applicable, explain:
    b. If yes, are they using in groups/pairs or individually?
    c. If yes, are students using the material correctly?

9. Visual aids or diagrams are displayed in the classroom to aid understanding of concepts.
    a. If yes, describe what is displayed:

10. The teacher identifies student errors and/or misconceptions while using concrete materials.
    a. If yes, the teacher addresses them effectively.

11. The teacher moves to pictorial/ abstract concepts without having the children practice with concrete materials

Learner practice and student errors

12. The teacher asks students to explain their reasoning when providing an answer.

13. The teacher prompts students to describe the steps they used to reach their answer.

14. The teacher encourages students to explore alternative solutions or different approaches to solving the problem.

15. The teacher uses student mistakes as a learning opportunity.
    - Yes, by guiding students to identify their misconceptions in front of the class
    - Yes, by discussing errors with the class
    - No
    - N/A

16. The teacher has students evaluate or respond to each other's solutions.

17. Which does the teacher do more?
    - Reinforce correct thinking or effort
    - Focus on mistakes
    - A balance of both
    - Neither
    - Not Applicable

18. When the teacher starts discussing the new topic, does he/she start with:
    - Yes/no questions
    - Open-ended questions or prompts
    - Both
    - Neither
    - Not Applicable

Ability Grouping: Learning in groups

19. The teacher conducts the lesson including whole group instruction.

20. The teacher forms students into groups to continue learning.
    a. If yes, how long do students work in their assigned groups? __ write grouping time in minutes. __ write entire lesson time in minutes
    b. If yes, does the teacher circulate to monitor learning?
        i. If yes, does the teacher provide feedback?
            - To individual students?
            - To the group
            - Don't know/not applicable
    c. If yes, are groups:
        - Heterogenous
        - 1 group only: Remediation
        - 2 levels: Remediation & enrichment

        NB Observer will know if it is 2 levels because teacher will assign 2 different tasks to different groups

        i. If Remediation, do the high-ability students explain concepts to the low-ability students?

21. Are students encouraged to discuss the topic or task with their peers?

Ability Grouping: Assessment

22. The teacher presents assessment task by:
    - Writing the task on chalkboard
    - Giving the assessment task verbally
    - Not applicable, explain

a. If yes, students do the assessment:
   - Individually without talking
   - Individually but with talking
   - As a group
   - Not applicable, explain

b. If yes, the teacher writes assessment answers on chalkboard or says them out loud (circle one)

c. If yes, the teacher distributes score summary sheet to group leaders.

d. If yes, the students correct each other's work.

e. If yes, the group leader records student responses.

f. If yes, the teacher collects assessment scoring sheet

g. If yes, the teacher asks students to share what they learned

Enumerator reflection: After the observation, note the following:

On a scale of 1-10, how would you rate the teacher's approach in this lesson?

This space is to take additional notes:

Teacher Interview

Administered in Kiswahili with English translation provided below.

Introduction & Consent: This interview is designed to gather insights on your experiences with AG and FMKT strategies in numeracy instruction. It will take 30-45 minutes and include both structured and open-ended questions to help us understand how these strategies are implemented beyond daily lessons.

Your participation is voluntary, and all responses will remain confidential. Do you agree to participate?

If yes, continue. If no, thank the teacher and leave the room.

Background Information:

1. Name of Teacher: _____
2. Gender: _____
3. Grade Level(s) Taught: _____
4. School Name: _____
5. Years of Teaching Experience: _____
6. Subjects Taught: _____
7. Number of learners in class: __ boys __ girls __ total

Flexible AG Implementation:

1. IF YOU OBSERVED THE TEACHER ORGANIZING LEARNERS INTO GROUPS DURING THE LESSON TODAY, ASK: In the lesson we observed today, you organized your learners into groups. Were these groups are based on previous assessment results?
2. How often do you use ability grouping in your numeracy lessons – that is, both grouping students for learning and conducting assessments?
    - Every day
    - Several times per week
    - Less than once per week
    - Never
3. How often do you calculate results from the assessments?
    - Every day
    - Several times per week
    - Less than once per week
    - Never
4. How often do you group students based on the results from the assessments?
    - Every day
    - Several times per week
    - Less than once per week
    - Never

5. **If we provided tracking sheets:** Have you been able to use the tracking sheet easily?
    a. If yes, have you had any difficulties using it? If so, what?
    b. Ask if you can see the tracking sheet. Do they have at least one example of having done it correctly?
    c. Comments:

6. **If we provided an app:** Have you been able to use the app easily?
    a. Have you had any difficulties using it? If so, what?
    b. Ask if you can see how they are tracking student progress on their app. Do they have at least one example of having done it correctly? Yes/no
    c. Comments:

7. With AG, after an assessment, there are 3 options: Proceed to the next lesson, organize learners in 2 groups (enrichment and remediation), or organize all learners in 1 group (remediation only). In the past 3 weeks, how many times have you proceeded to the next lesson?
    - Every day
    - Several times per week
    - Less than once per week
    - Never

8. In the past 3 weeks, how many times have you organized learners in 2 groups?
    - Every day
    - Several times per week
    - Less than once per week
    - Never

If never, proceed to question 10.

9. When organizing learners in 2 groups - remediation & enrichment, how do the remediation learners respond? How do the enrichment learners respond? Have you ever observed learners feeling uncomfortable with being assigned to a particular group?

10. In the past 3 weeks, how many times have you organized learners in 1 group (remediation only)?
    - Every day
    - Several times per week
    - Less than once per week
    - Never

11. What challenges do you face when implementing ability grouping?

12. How do you support students who are struggling within their assigned groups?

13. How do you foster collaboration and peer learning within ability groups?

Use of FMKT Strategies

14. How often do you use concrete materials in your numeracy lessons?

    - Every lesson
    - Once per day
    - Less than once per day
    - Never

15. Do you ask students to explain their reasoning when answering a question?
    - Every lesson
    - Once per day
    - Less than once per day
    - Never

16. How do you typically respond to student errors? (choose one)

    - Provide the correct answer immediately
    - Allow students to find their own answers
    - Guide students to identify their mistakes
    - Other (please specify): _____

17. Do you ever Use a mistake a learner has made as a learning opportunity for the class?
    a. If yes, please describe a recent example:
18. Do you encourage students to think about different ways to solve a problem?
    a. If yes, please describe a recent example:
19. Do you ask your learners to have peer discussions in numeracy lessons?
    a. If yes, please give an example:

Teacher Attitudes Towards Teaching & Learning

17. Do you see your role more as a:

    - Transmitter of knowledge or a
    - Facilitator of learning?
    - Or a mix of both. If a mix, explain when you play one role and when you play the other._____

18. How has implementing ability grouping influenced your teaching approach?

    - It has made teaching more effective
    - It has made teaching more difficult
    - No significant impact
    - Other, please describe_____

19. How has implementing FMKT strategies influenced your teaching approach?

    - It has made teaching more effective
    - It has made teaching more difficult
    - No significant impact
    - Other, please describe_____

19. Do you think students benefit from ability grouping? Why or why not?
20. Do you think students benefit from FMKT strategies? Why/why not?
21. What aspects of your teaching have changed the most since incorporating these strategies?

Scenario-based questions (Enumerator: feel free to change based on classroom observation)

Scenario 1: A student provides an incorrect answer. How do you handle this situation?

Scenario 2: You notice that students in one group are struggling significantly while the other group is excelling at their tasks easily and getting bored. What do you do next?

Scenario 3: A student finds an alternative way to solve a problem that is different from your method but still produces the correct answer. How do you respond?

Teacher FGD

For Grade 1-3 Teachers & Head Teacher

Administered in Kiswahili with English translation provided below

Introduction & Consent: This Focus Group Discussion (FGD) aims to gather insights on teachers' experiences, perceptions, and challenges in implementing AG and FMKT strategies in numeracy instruction. The discussion will be built on classroom observations and teacher interviews and provide an opportunity for you to share best practices and concerns in a collective setting. Your participation is voluntary, and all responses will remain confidential. Do you agree to participate?

If yes, continue. If no, thank the teacher and leave the classroom.

Teachers' Experiences with Ability Grouping (AG) (10 minutes)

Facilitator Prompt: Let's discuss your experiences with AG in your classrooms.

1. When we interviewed (or observed) you, some of you said you had (or we observed) the following challenges when working with ability groups: XXX What do you think can be done to address these challenges?
2. Have you noticed any improvement in student learning outcomes since implementing ability grouping? Describe.
3. Have you noticed any improvement in student behavior or their attitudes toward learning since implementing ability grouping? Describe.
4. Which parts of ability grouping would you like to get better at?

Teachers' Use of FMKT Strategies (10 minutes)

Facilitator Prompt: We also observed how teachers use FMKT strategies, including the use of concrete materials, addressing student errors, and encouraging reasoning.

5. When we interviewed (or observed) you, some of you said you had (or we observed) the following challenges when working with FMKT: XXX What do you think can be done to address these challenges?
6. Do you find it challenging to using concrete materials to teach math concepts? If so, why?
7. What are the most frequent types of errors you find your learners making?
8. Do you think it is best to show learners how to do learn numbers and functions correctly, or have them work with materials and discuss, then explain their reasoning? Which approach do you think is better for student engagement and understanding? Why?
9. Which aspect of FMKT (teaching with concrete materials, working with student errors) would you like to get better at?

Self-Efficacy and Outcome Expectancy (10 minutes)

Facilitator Prompt: Now, let's talk about how using these strategies influences your confidence and expectations about student learning.

10. Have any aspects of your teaching changed since being trained in ability grouping or FMKT? If so, which ones?
11. Are there any additional resources that could make it easier to implement these strategies?

Student-Centered Teaching & School Support (10 minutes)

Facilitator Prompt: Finally, let's discuss how these strategies influence your role as a teacher and the support you receive.

12. How do you see your role as more of a teacher? A transmitter of knowledge or a facilitator of learning? Why?
13. What recommendations do you have for improving the use of AG and FMKT in your school?

Conclusion & Wrap-Up (5 minutes)

14. Summarize key discussion points and highlight common themes. Ask if anyone has final thoughts or additional comments.

Thank the teachers for their time and valuable insights.

Head Teacher Interview

Introduction & Consent: This Head Teacher interview aims to learn about your school, your experiences, and your perspectives concerning AG and FMKT strategies introduced by this project. Your participation is voluntary, and all responses will remain confidential.

Do you agree to participate?

If yes, continue. If no, thank the Head Teacher and leave the room.

Background Information

Name of Head Teacher:

Gender:

Grade Level(s) Taught (if any):

School Name:

Years of Teaching Experience:

Subjects Taught (currently, if any):

Number of learners in school: __ boys __ girls __ total

Teachers' Experiences with Ability Grouping (AG) (10 minutes)

Facilitator Prompt: Let's discuss your experiences with AG in your classrooms.

1. Did you attend the training of teachers on AG & FMKT?
    a. If yes, which ideas did you find to be the most important?
    b. If yes, which parts of the training were most challenging for your teachers?
2. Since the training, have your teachers been able to implement ability grouping successfully? Can you give any examples?
3. Which aspects of ability grouping are your teachers struggling with most?
4. As a Head Teacher, are you able to provide support to them on AG?
5. What kind of support would you like in order to be able to help your teachers more with AG?
6. Since the training, have your teachers been able to implement FMKT successfully? Can you give any examples?
7. Which aspects of FMKT are your teachers struggling with most?
8. As a Head Teacher, are you able to provide support to them on FMKT?
9. What kind of support would you like in order to be able to help your teachers more with FMKT?

10. In general, have you noticed any changes in your school since the training? Which ones?

11. Do you have any other comments for us?

Thank you!

School-to-School International

500 Victoria Avenue
Suite 203 #190
Millbrae, CA 94030

STS-International.org