USAID | JORDAN
FROM THE AMERICAN PEOPLE

EdData II

Education Data for Decision Making (EdData II): National Early Grade Literacy and Numeracy Survey–Jordan

Intervention Impact Analysis Report

Final Report

EdData II Technical and Managerial Assistance, Task Order Number 16
Contract Number: AID-278-BC-00019
August 2014

This publication was produced for review by the United States Agency for International Development. It was prepared by RTI International.

Education Data for Decision Making (EdData II): National Early Grade Literacy and Numeracy Survey–Jordan

Intervention Impact Analysis Report

Final Report

EdData II
Task Order No. 16
Period Ending November 30, 2015

Prepared for
USAID/Jordan
Angie Haddad, COR

Prepared by Aarnout Brombacher, Jonathan Stern, Lee Nordstrum, Chris Cummiskey, and Amy Mulcahy-Dunn
RTI International
3040 Cornwallis Road
Post Office Box 12194
Research Triangle Park, NC 27709-2194

RTI International is a trade name of Research Triangle Institute.

The authors' views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

Table of Contents

Page

List of Figures

List of Tables

Abbreviations

C	Control group
COR	Contracting Officer's Representative
cwpm	correct words per minute
DID	Difference-in-differences
Diff	Difference
EdData	Education Data for Decision Making
EGMA	Early Grades Mathematics Assessment
EGRA	Early Grades Reading Assessment
EGRP	Early Grade Reading Program [Egypt]
EMIS	Education Management Information System
ETC	Education Training Center
L1	Level 1
L2	Level 2
MoE	Ministry of Education
MSA	Modern Standard Arabic
NC	North Carolina
NEAT	Nonequivalent Anchor Test [design]
NGO	nongovernmental organization
NI	Number identification
ORF	Oral reading fluency
RTI	RTI International (a trade name of Research Triangle Institute)
SMS	Short message service
SSME	Snapshot of School Management Effectiveness
T	Treatment group
TE	Treatment effect (or impact)
USAID	United States Agency for International Development

Acknowledgements

The authors want to acknowledge the important contributions of many people who made this study and report possible, in particular:

- The Ministry of Education (MoE) and, in particular, Eng. Firyal Aqel (Director of the Donor Coordinating Unit), our most direct point of contact, for her continuous support and ongoing cooperation. Also, Dr. Khawla Abu Al- Haija (Education Training Center) for her invaluable support to the training component, Dr. Wafa Al_Abdallat (Curriculum) and her team for their enthusiastic engagement from the inception of the activities, and Dr. Ahmad Mousa Al-Ajarme for the coordinating role that he has played from the outset. Without their involvement and commitment, this study would not have been possible. From the conception of the intervention project, through the materials development, training of trainers (supervisors), training and coaching of teachers, and the endline study, the MoE has “owned” this study, and the results are all the more authentic and worthwhile for these reasons.
- The many MoE staff (too many to mention here—see Annex 3 for a comprehensive list), who have contributed to the various phases of the intervention. Without their care, commitment, and enthusiasm, this intervention would not have been possible.
- Angie Haddad, Allyson Wainer, and Lee Cohen (U.S. Agency for International Development [USAID], Jordan) and Susan Ayari (formerly USAID, Jordan), who made the study possible and have shown interest in and commitment to the process throughout the study.
- Wael Salah and Fathi El-Ashry, the Arabic language specialists, who provided technical support to all phases of the intervention. The leadership, technical expertise, and professionalism they exhibited were critical to the success of the intervention.
- Khalid Dajani and Samer Ghannam and the team at Dajani Consulting, who managed the logistics of the project, from organizing the various workshops to the production of the materials and managing the critical survey processes of the study.
- Chris Cummiskey, Jonathan Stern, Jennifer Pressley, Tracy Kline, and Lee Nordstrum from RTI for the sampling, data cleaning, data processing, analysis, and psychometrics needed to create and manage the vast data set associated with this study.
- Jonathan Stern and Lee Nordstrom, for their contributions to report writing.
- Amy Mulcahy-Dunn, Abhijit Nimbalkar, David Harbin, and Kate Batchelder from RTI, for the general guidance, oversight, and management of the study.
- Taghreed Sulayman in the RTI Jordan office for her local attention and support and, in particular, for her tireless efforts in maintaining the supervisor visits database.

- Ellen Lohr-Hinkel and Gail Hayes from RTI, for the editing and layout of this report.
- Finally, this work could not have succeeded without the cooperation and contributions of the school principals, teachers, students, and communities included in the study, who, for obvious reasons, must remain anonymous.

Executive Summary

Background

To gain insight into student facility with foundational skills and to better understand characteristics among Jordanian schools that are associated with student performance, USAID/Jordan, in partnership with the Jordan Ministry of Education (MoE), contracted with RTI International in 2011 under the Education Data for Decision Making (EdData II) project, Task Order 16, to conduct the Snapshot of School Management Effectiveness (SSME), including the Early Grade Reading Assessment (EGRA) and Early Grade Mathematics Assessment (EGMA), in a sample of primary schools in Jordan at the end of the 2011/2012 school year.

It is generally accepted that a child reads with comprehension when they can correctly answer 80% or more of the reading comprehension questions associated with a grade level text. The 2012 National Early Grade Literacy and Numeracy Survey (2012 National Survey) in Jordan revealed that the 17% of students who were able to answer 80% or more of the comprehension questions correctly were reading at an average oral reading fluency rate of 41.5 correct words per minute (cwpm). The *average* reading speeds recorded were, however, well below this rate (grade 2 students read 15.2 cwpm, while grade 3 students read 23.7 cwpm)—too slow to permit students to be reading with comprehension.

In terms of mathematics, students answered the more procedural addition and subtraction level 1 (L1) items correctly and with confidence—83.6% for addition and 79.4% for subtraction in grade 2, and 81.6% for addition and 75.9% for subtraction in grade 3—student performance dropped by 31% (in grade 2) and 27% (in grade 3) from L1 addition to level 2 (L2) addition, and by more than 47% (in grade 2) and 41% (in grade 3) from L1 subtraction to L2 subtraction. These results in Jordan suggested that memorization plays a large role in the way that children know and learn mathematics.

Intervention Description

The 2012 National Survey very clearly revealed that Jordanian children in the early grades were not reading with comprehension or doing mathematics with understanding. The opportunity to conduct a one-year intervention pilot presented a unique challenge: How could an intervention make the greatest possible impact on the students' performance in reading and mathematics in a short time period?

The 2012 National Survey had revealed that students were not receiving sufficient instruction in foundational reading and mathematics skills, with little hope of having this insufficiency addressed by their teachers or the curriculum that was in use in 2012. Against this background, it was decided, after discussions with the MoE Curriculum Team and the Senior Reading and Mathematics Supervisors, to develop an intervention program that would support teachers in providing deliberate, structured, and developmentally appropriate daily practice in foundational skills for reading and mathematics.

Teachers would be asked to spend the first 15 minutes of every reading and mathematics lesson to revisit and reinforce foundational skills. They would do so every day, so that the students experienced this activity as part of the classroom program—as a routine “warm-up” activity to the curriculum’s lesson for the day. In addition to addressing the foundational skills that the 2012 National Survey had identified as being underdeveloped in grade 2 and grade 3 students, the different activities for each of the skills that would be addressed by the daily routine would also be designed to introduce teachers to more research-based pedagogical practices.

The research questions of the intervention were to establish the following:

- Does daily practice of foundational skills through deliberate, structured, and developmentally appropriate activities support children to be able to read with comprehension and do mathematics with understanding?
- What are the conditions that help teachers to implement the daily routine and the associated activities with fidelity and confidence?

The materials for the intervention were developed by two teams of writers, one for reading and one for mathematics. The teams included members of the subject curriculum committees, supervisors, and teachers teaching in grades 1, 2, and 3. RTI appointed one technical expert per subject to provide leadership and guidance during the materials development process. Three sets of materials were developed for each subject and grade: (1) a teacher’s guide, (2) daily lesson notes, and (3) a students’ workbook. The teacher’s guide for each subject was developed as a resource to provide teachers with a pedagogical rationale for the teaching approach of the intervention and with guidance on how to conduct the activities associated with the different skills. A set of daily lesson notes was developed for the teacher. The lesson notes identified the skills to be included in each 15-minute routine of the year, as well as the activities to be used for each skill. The workbooks provided a resource for daily independent work for students.

The training and ongoing support of teachers was carried out exclusively by MoE supervisors, with technical support from RTI. Training was conducted in two stages: (1) training of trainers and (2) training of teachers. RTI technical experts provided the training of trainers (MoE supervisors), while MoE supervisors provided the training of teachers. A total of 20 MoE supervisors were assigned to the intervention. In addition to the training that supervisors provided to teachers, they also provided in-class coaching and support and coordinated a monthly reflection session among the participating teachers in their district.

Implementation

The intervention was implemented during the 2013/2014 school year by more than 400 teachers in 347 classrooms across 43 schools, reaching approximately 12,000 students.

Although the intervention implementation proceeded largely according to plan, it should be noted that it was not without challenges that involved two main categories—logistical challenges and challenges related to transferring the vision of the intervention into practice. The logistical challenges were related to the bureaucratic and centralized nature of the decision-making process within the MoE, which resulted in teachers not starting training as planned and teachers being assigned to schools after the training had been completed. The greater challenge that the intervention faced involved successfully sharing the vision of the intervention with the participating teachers. The key difficulty that teachers had in assimilating new pedagogies resulted from mismatches between the intervention vision and teachers' predominant teaching styles and approaches.

2014 Endline Survey

To measure the impact of the intervention pilot, an endline survey was conducted in May 2014. Because the intervention was solely concerned with improving reading and mathematics, the endline survey included the EGRA and EGMA and not the SSME tools of the 2012 Survey. Data was gathered from a wide range of sources, to gain a fuller understanding of the impact of the intervention and the variables that influence the chances of the intervention's success. These sources included (1) an **oral teacher questionnaire,** completed by the grade 2 and 3 teachers of the assessed pupils in the treatment schools; (2) a **written teacher questionnaire,** completed by the grade 1, 2, and 3 teachers who implemented the intervention in treatment schools; (3) an **oral student questionnaire,** completed by all students in treatment schools; (4) **supervisor visit reports,** submitted by SMS, and completed by supervisors each time they visited a teacher to provide a good source of data about the fidelity with which the intervention was implemented; (5) a record of **teacher participation** kept in a database that was developed and maintained to include information for the intervention about teacher status, number of days of training that teachers had completed, and the number of supervisor visits the teacher had received; and (6) **project team field notes**.

The final sample for the 2014 endline survey consisted of 151 schools (110 control schools and 41 treatment schools). Data were collected for a total of 5,911 students across the two years (2012–2014), with 2,976 students in 2012 (2,159 control; 817 treatment) and 2,935 students in 2014 (2,129 control; 806 treatment). To make the sample representative of the national population, weights were calculated. All scores reported for this study were calculated using these student weights. Test equating, to ensure that differences in scores are the result of differences in ability and not differences in test difficulty, was conducted to calculate comparable scores on the different forms of a test (the baseline and endline assessments).

To limit the effect of selection bias by comparing the average change over time in the outcome variable for the treatment group with the average change over time in the outcome for the control group, the intervention impact was determined using difference-in-differences (DID) methodology.

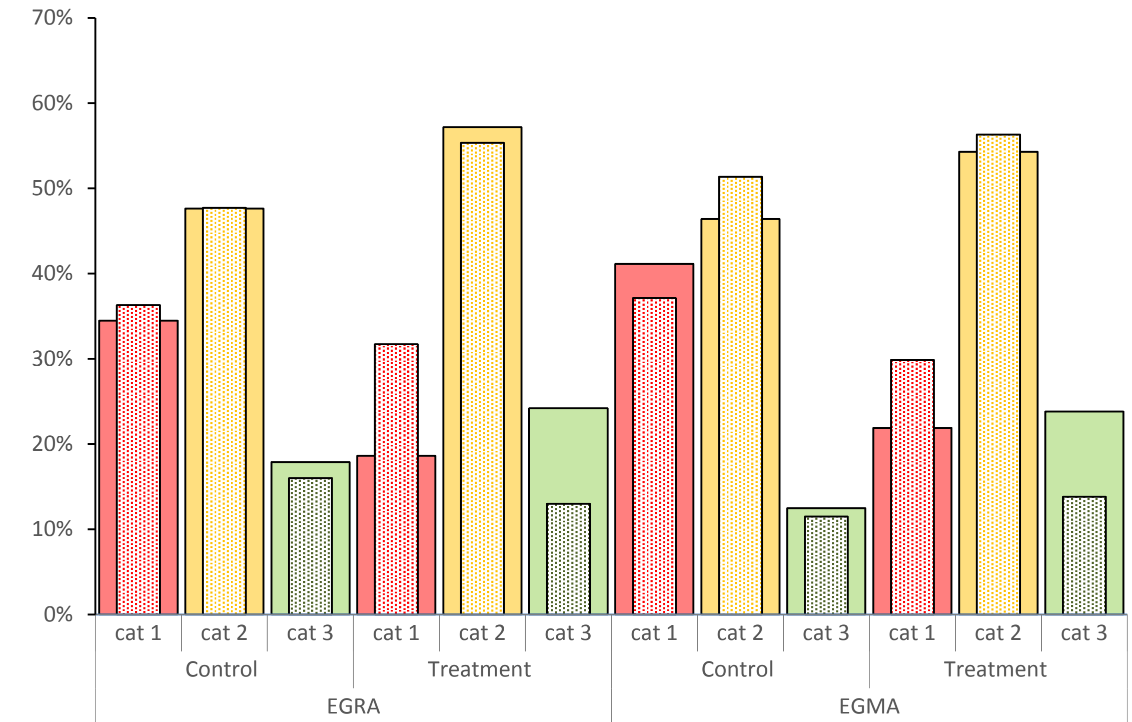
To ensure that the assessment instruments were measuring their intended constructs, both the EGRA and EGMA were tested for reliability. Cronbach's alpha values for both EGRA and EGMA indicated that the instruments showed good internal consistency on average ($\alpha = 0.86$ for EGRA and 0.90 for EGMA for the baseline assessment; and $\alpha = 0.86$ for EGRA and 0.85 for EGMA for the endline assessment).

Findings

To gain an overall impression, composite reading and mathematics scores were created to determine the aggregate effect of the intervention. Specifically, oral reading fluency and reading comprehension scores were used to calculate an overall reading score; missing number scores and the addition and subtraction L2 scores were used to create an overall mathematics score. These scores were used to classify the students as non-readers or beginning readers and non-mathematicians or early mathematicians, emergent readers and mathematicians, and readers and mathematicians.

Figure ES1 displays the changes in these categories from 2012 to 2014, both for treatment and for control schools.

Figure ES1. Overall treatment effect for EGRA and EGMA categories



	EGRA						EGMA					
	Control			Treatment			Control			Treatment		
	cat 1	cat 2	cat 3	cat 1	cat 2	cat 3	cat 1	cat 2	cat 3	cat 1	cat 2	cat 3
Endline	34.5%	47.7%	17.9%	18.6%	57.2%	24.2%	41.1%	46.4%	12.5%	21.9%	54.3%	23.8%
Baseline	36.3%	47.7%	16.0%	31.7%	55.3%	13.0%	37.1%	51.4%	11.5%	29.9%	56.3%	13.8%

Key: narrow bar: baseline; wide bar: endline

Figure ES1 provides direct evidence of the overall effectiveness of the intervention. While the percentage of non-readers or beginning readers and non-mathematicians or early mathematicians remains relatively consistent across years for the control group, there are large reductions in the proportion of non-readers or beginning readers and non-mathematicians or early mathematicians in treatment schools (from 32% to 19% in reading and 30% to 22% in mathematics). Additionally, while the proportion of readers and mathematicians remains constant for control schools, both proportions increase significantly in treatment schools (13% to 24% in reading and 14% to 24% in mathematics). The intervention did exactly what it was intended to do. While there were virtually no gains in control schools from 2012 to 2014, there were significant gains across treatment schools in terms of reducing the proportion of the lowest performers and increasing the proportion of the highest performers. These results are extremely promising, particularly because the intervention was implemented for only one school year.

The DID analysis for the EGRA subtasks shows that the intervention had a significant impact on treatment school students for every EGRA measure. The smallest impact was for the invented word subtask, which showed a 4.1% increase as a result of the intervention. The largest effect was found for letter sounds, for which the intervention was determined to provide a nearly 14 point increase in the percentage of correct responses. The DID analysis for EGMA subtasks shows that although there is no significant impact on number identification, scores on all five of the other subtasks were significantly increased by the intervention. The L2 addition and subtraction subtask realized the largest gains, with an increase of 15.5%.

An examination of the impacts of this intervention on male and female students reveals that, overall, girls outperformed boys in the sample, which raises the question if there were different effects of the intervention based on gender. The answer is yes. The intervention did not provide significant gains in achievement for male students on even a single EGRA subtask. Conversely, significant gains were seen by female students across every subtask. Similar results were found for EGMA subtasks. The intervention produced significant gains in achievement for female students across all subtasks, while providing no significant increases for male students. Furthermore, students in all-girls schools performed better than students in mixed schools, who performed better than students in all-boys schools. These differences were statistically significant on almost all EGRA and EGMA subtasks.

Key factors and their influence on results

To establish the factors that are associated with the top performing classrooms and districts, top performing classrooms are defined as those with at least a 10% increase in readers (or mathematicians) in grade 2 or at least a 20% increase in grade 3 readers (or mathematicians). Top performing districts are defined as the four out of the 12 intervention districts with the largest increases in readers or mathematicians from 2012 to 2014.

Analysis of the variables revealed that:

- 93% of teachers with frequent supervisor visits were in top performing classrooms for reading (i.e., largest increase in readers), as compared to only 41% of those teachers in classrooms who were visited by supervisors fewer than 16 times.

- 63% of teachers who attended both training sessions were in top performing classrooms for mathematics, as compared to only 11% of those teachers in classrooms who did not attend both trainings.
- 65% of the reading and 89% of the mathematics classes in all-girls schools were in top performing classrooms.
- 84% of the classes in which teachers marked all of the work in the student workbooks sessions were in top performing classrooms for mathematics.
- 80% of the classes in which teachers monitored student understanding by asking for further explanations were in top performing classrooms for mathematics.
- 69% of the classes in which teachers followed the notes and routines of the intervention with fidelity were in top performing classrooms for mathematics.

Teachers were interviewed about their experience with the intervention. On balance, teacher respondents had more to say that was positive than negative in terms of the overall aspects of the project: a majority of teachers felt positively about six elements of the project, while a majority felt negatively about only two elements. On the positive side, teachers asserted that students enjoyed the project, it developed thinking skills and student skills generally, had a positive impact on learning, and led to an improvement in teachers' skills. Other positive aspects cited by nearly one-half of respondents were project training and the support given teachers by their supervisors. On the negative side, teachers pointed out that the project increased teachers' workloads and required too much time or effort to keep up with instructional demands (e.g., marking student workbooks). Despite their overwhelmingly positive response to the intervention, when teachers were asked if the intervention should be or should not be continued in their schools, they responded overwhelmingly not in favor of continuing, by a margin of almost five to one. It is not completely clear why a majority of teachers made this judgment; however, the teachers' experience of the project as an add-on and hence an additional burden, as reported in discussions and during training, may provide some explanation.

Lessons Learned

The intervention has demonstrated unequivocally that it is possible to increase the number of readers and mathematicians in early grade classrooms by providing deliberate, structured, and developmentally appropriate practice in foundational skills for reading and mathematics. The implication may well be that there is much to be gained by an intervention that systematically addresses only those key elements of a teaching and learning program that has been shown to be deficient, instead of replacing the entire program.

Encouraged by the positive results, it is nonetheless critical to examine the different components of the intervention to see what lessons can be learned—lessons that will inform future interventions and improve their chances of success.

Gender. Although neither the survey nor the intervention set out to explore how the role of teacher and student gender affects student performance, the results nonetheless revealed that boys are, in general, not benefiting from early grade instruction in the same way that girls are. Furthermore, the results also indicate that there is a statistically significant relationship between the school type (all-girls, all-boys, and mixed), the gender of the teacher and the performance, by gender, of the students. It is very clear that future intervention projects will need to better understand the gender dynamics of Jordanian schools and to make conscious design decisions to ensure that boys benefit as much from the intervention as girls do.

Classroom support. Classroom visits by the supervisors contributed to the impact of the intervention on the proportion of readers and mathematicians in a teacher's class. More frequent support resulted in more effective intervention implementation. In the case of this intervention study, however, it was also clear that not all supervisors were able to visit classes as often as the intervention hoped that they would (once every two weeks). Future intervention projects will need to establish mechanisms for maximizing the ability of teachers' coaches (supervisors) to attend to this work.

Teacher training. The proportion of the training that teachers attended was a variable that had a significant impact on the success of the implementation. Teachers who attended more of the training had a greater proportion of readers and mathematicians in their classes than teachers who attended less training. Future intervention projects will need to examine the factors that prevent teachers from attending the training as expected and to find ways of dealing with these.

Translating the vision of the intervention into practice. Teachers experienced difficulties in assimilating new pedagogies into their practice. Some of teachers' criticisms about the intervention reflected not so much fundamental problems with the intervention, but rather mismatches between the intervention vision(s) and teachers' predominant teaching styles and approaches. Future intervention projects will need to explore different ways of introducing teachers to the pedagogies, including the role of web-based communities of practice.

Conclusion and Recommendations

This intervention study set out to explore whether daily practice of foundational skills through deliberate, structured, and developmentally appropriate activities can support children to be able to read with comprehension and do mathematics with understanding. And if so, what the conditions are that help teachers to implement the daily routine and the associated activities with fidelity and confidence.

The results show quite clearly that the intervention did exactly what it was intended to do. While there were virtually no gains in control schools from 2012 to 2014, there were significant gains across treatment schools in reducing the proportion of the lowest performers and increasing the proportion of the highest performers. These results are extremely promising, particularly because the intervention was implemented for only one school year.

At the presentation of the intervention results, key recommendations emerging from a meeting with ministry departments, representatives from the donor community, and nongovernmental organizations (NGOs) working in the field of early grade education include that:

- A qualitative study should be conducted to assist the MoE and other stakeholders to better understand why it is that boys do not benefit from schooling in the early years to the same extent as girls.
- In the short term, the MoE could explore the feasibility of using only female teachers and of having only mixed and all-girls schools in the early grades.
- The number of teachers for which each supervisor is responsible should be reduced, to allow supervisors to be more effective in providing teacher mentoring and support.
- Teachers need to be rewarded for participating in in-service training activities. This reward could be either direct, in the form of financial reward, or indirect, in credits earned as part of a continuing professional development program that impacts teacher employment, promotion, and tenure.
- Intervention implementers need to take care to ensure that teachers experience:
  - Intervention activities not as add-ons to the work that they do, but instead as supportive of and integral to what they do.
  - Intervention activities as activities of the MoE and directly linked to the curriculum.
- Teachers be encouraged to commit to exploring new methodologies; success stories about program impact be shared with teachers through a range of media, including social media web-based communities of practice. And, that video vignettes be developed. These videos should demonstrate the desired methodologies being successfully implemented by teachers in typical classrooms.

Benchmarks and targets

On conclusion of the 2014 Endline Survey, participants representing the various ministry departments set benchmarks and targets (reported in ***Table 21*** of the report) for EGRA and EGMA. The benchmarks were based on the results of grade 2 and grade 3 Jordanian students in the 2014 National Survey and informed by a range of international benchmarks, the participants' experience with and knowledge of the Jordanian context, and technical support provided by the researchers who led the RTI research team. Five-year targets were based on the evidence of the intervention's potential impact, as noted in this report.

1 Background

1.1 2012 National Survey

To gain insight into student facility with foundational skills and to better understand characteristics among Jordanian schools that are associated with student performance, USAID/Jordan, in partnership with the Jordan Ministry of Education (MoE), contracted with RTI International in 2011 under the Education Data for Decision Making (EdData II) project to conduct the Snapshot of School Management Effectiveness (SSME), including the Early Grade Reading Assessment (EGRA) and Early Grade Mathematics Assessment (EGMA), in a sample of primary schools in Jordan at the end of the 2011/2012 school year. The hope was that evidence-based information resulting from the survey could inform future education policy decisions, as needed.

The instruments used in the project—the National Early Grade Literacy and Numeracy Survey in Jordan[1]—were adapted specifically for the Jordanian context during an adaptation workshop with the MoE. Abbreviated versions of the EGRA and EGMA were developed, using curriculum materials for grades 2 and 3. In addition to administering individual oral assessments of students, school principals and teachers were interviewed, inventories of school and classroom resources were conducted, and reading and mathematics lessons observed as part of the SSME survey.

A nationally representative sample of 156 public primary schools across Jordan was involved in the study. In each school, a grade 2 teacher and a grade 3 teacher were randomly selected, and 10 students from each of these classes were randomly selected to take the EGRA and EGMA and to be interviewed about their experience with school. A total of 3,120 students were selected for participation in the assessments and interview. The selected teachers were interviewed, as was the school principal, and a researcher observed the selected grade 2 teacher teach a reading lesson and a math lesson. Researchers also took inventory of the school grounds and the selected classrooms. Data collection was completed at the end of May 2012.

The EGRA, which was administered orally in Modern Standard Arabic (MSA), consisted of five subtasks: (1) letter-sound knowledge, (2) invented word decoding, (3) connected text oral reading fluency, (4) reading comprehension, and (5) listening comprehension. Letter-sound knowledge and the ability to read unfamiliar single-syllable words are foundational skills needed for fluent reading and comprehension.

As part of the EGRA, to determine their oral reading fluency (ORF), students were asked to read as much of a short narrative as they were able in one minute. The results of this task were used to estimate ORF. On average, grade 2 students read 15.2 correct word per minute (cwpm), while grade 3 students read 23.7 cwpm, indicating progression in performance from grade 2 to grade 3. Research has shown that readers

[1] Although "early grades" is used generally in this report to refer to the first three school years (grades 1–3), in the case of the 2012 National Survey, and later the 2014 National Survey, only grade 2 and grade 3 students were assessed. In the intervention activity, all three grades (1–3) were involved.

must read at a certain minimum speed to understand what they have read. In the 2012 study, students who were unable to answer even a single comprehension question correctly were reading at a speed of fewer than 2 cwpm, and those who were able to answer all questions correctly were reading at a speed of 49.3 cwpm. It is generally accepted that a child reads with comprehension when they can correctly answer 80% or more of the reading comprehension questions associated with a grade level text. Students who were able to answer 80% or more of the comprehension questions correctly were reading at an average fluency rate of 41.5 cwpm. The *average* reading speeds recorded were well below this rate—too slow to permit students to be reading with comprehension. As a result, student performance on the comprehension questions was not as strong as curricular guidelines required.

The reported reading speeds and comprehension scores were not surprising given students' performance on the more foundational reading skill subtasks. A total of 24.1% of students were unable to respond correctly to a single item on the letter-sound subtask, and 47.1% were unable to respond correctly to a single item on the invented words subtask. Yet, strong ability with these foundational skills is essential for strong readers. The relationship that exists between students' foundational reading skills and reading fluency indicated that students' knowledge of letter sounds and decoding skills should be strengthened to improve their oral reading fluency and comprehension.

The EGMA, which was administered orally, consisted of six subtasks: (1) number identification, (2) quantity discrimination, (3) missing number (number patterns), (4) addition and subtraction (L1), (5) addition and subtraction (L2), and (6) word problems. The L1 addition and subtraction items were procedural in nature and involved single- and double-digit problems with sums/differences below 20, for which students were asked to solve the problems without using paper and pencil and then give their answer. The L2 addition and subtraction items were more difficult and required students to grasp mathematical concepts such as the bridging of tens. For these problems, students were permitted to use a pencil and paper to work out the solution.

Although students answered the more procedural addition and subtraction L1 items correctly and with confidence—83.6% for addition and 79.4% for subtraction in grade 2, and 81.6% for addition and 75.9% for subtraction in grade 3—student performance dropped by 31% (in grade 2) and 27% (in grade 3) from L1 addition to L2 addition, and by more than 47% (in grade 2) and 41% (in grade 3) from L1 subtraction to L2 subtraction. The curriculum for Jordan stipulates that grade 2 students should be able to perform addition and subtraction involving three- and four-digit numbers. And, in grade 3, students should be able to perform addition and subtraction problems involving five-digit numbers. However, the 2-digit addition and subtraction problems of the EGMA proved challenging to the sampled students, with grade 2 students correctly answering only 52.7% of the L2 addition problems and 32% of the L2 subtraction problems. Similarly, grade 3 students correctly answered only 55% of the L2 addition problems and 35% of the L2 subtraction problems.

These 2012 EGMA results in Jordan suggested that memorization plays a large role in the way that children know and learn mathematics. This suggestion was supported by the clear trend in the results showing that students were doing well on the items that rely on procedural knowledge—knowledge that can also be memorized—and doing markedly less well on the tasks and items that require both the understanding and the application of what should be procedural (rather than memorized) knowledge.

These assessments indicate that although students are quite comfortable with some of the procedural mathematics skills, their conceptual understanding needs to be strengthened by well-trained teachers. Similarly, although some students are reading with a high level of fluency and understanding and achieving 80% or more on their comprehension scores, the majority of students are not reading with fluency and lack strength in the foundational literacy skills normally taught in grade 1.

1.2 Intervention Description and Components

1.2.1 Rationale and Description

The 2012 National Early Grade Literacy and Numeracy Survey (2012 National Survey) very clearly revealed that Jordanian children in the early grades were not reading with comprehension or doing mathematics with understanding. The opportunity to conduct a one-year intervention pilot presented a unique challenge: How could an intervention make the greatest possible impact on the students' performance in reading and mathematics in a short time period?

In addition to the EGRA and EGMA, the 2012 National Survey included several assessments from the SSME suite of instruments. The classroom observations, in particular, revealed that daily lesson content was informed by the page in the textbook for the day. That is, teachers were teaching according to a schedule that determined what would be taught on each day with little regard for whether or not the children were developmentally ready for the lesson content. When asked what they do with assessment results, only 22% of teachers in 2012 responded that they used the assessment results to plan teaching activities or adapt their teaching to meet their students' needs.

An analysis of the curriculum in use in 2012, involving MoE officials from the Department of Curriculum and conducted as a follow-up to the 2012 National Survey, revealed a range of anomalies. The results of the 2012 National Survey revealed that grade 3 children were not performing well on the letter-sounds reading task. Moreover, this skill was not being addressed in the curriculum after grade 1. Likewise in mathematics, the curriculum in grade 2 required students to add and subtract three- and four-digit numbers, using the vertical column method, and yet one-half of the grade 2 students in the survey could not correctly subtract 3 from 19.

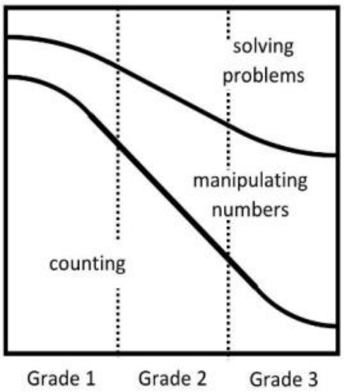
In short, the 2012 National Survey revealed that students were not getting sufficient instruction in foundational reading and mathematics skills—in foundational skills that research indicates are predictive of future success in reading and mathematics—with little hope of having this insufficiency addressed by their teachers or the curriculum that was in use in 2012.

Against this background, it was decided after discussions with the MoE Curriculum Team and Senior Reading and Mathematics Supervisors, to develop an intervention program that would support teachers in providing deliberate, structured, and developmentally appropriate daily practice in foundational skills for reading and mathematics.

Teachers would be asked to spend the first 15 minutes of every reading and mathematics lesson to revisit and reinforce foundational skills. They would do so every day, so that the students experienced this activity as part of the classroom program—as a routine "warm-up" activity to the curriculum's lesson for the day. The 15-minute activity's key feature was for it to become part of the daily routine using the same structure every day, with the rationale that as students (and teachers) became familiar with the routine, it would go quickly and not require a large amount of explanation; it would provide both the needed exposure to and the practice with key foundational skills.

For improving reading, the daily routine would address up to three of five different foundational skills each day. For mathematics, the daily routine would address three different foundational skills with the amount of time to be spent on each, varying from grade 1 to grade 3 (see ***Figure 1***). In grade 1 more time would be dedicated to counting than in grade 3, and the amount of time allocated each day to manipulating numbers would increase from grade 1 to grade 3.

Figure 1. Allocation of time to the three mathematics skills of the daily routine, by grade.



For each of the skills to be addressed in the reading and mathematics routines, a finite number of different classroom activities would be used (see ***Table 1a*** and ***Table 1b***). On the one hand, this is to ensure variety and that all aspects of the skill would be addressed; on the other hand, it is to reduce the number of different classroom activities for which teachers would need to be trained and to which students would need to become accustomed.

In addition to addressing the foundational skills that the 2012 National Survey had identified as being underdeveloped in grade 2 and grade 3 students, the different activities for each of the skills that would be addressed by the daily routine would also be designed to introduce teachers to more research-based pedagogical practices. Such pedagogical practices are known from research to support the development of foundational reading and mathematics skills. In this sense, it was hoped that the intervention would introduce teachers to more effective pedagogical practices than those that the 2012 National Survey had seen in use in early grade classrooms. However, rather than introducing these practices through direct instruction, the program would seek to do so through immersion. By doing the different activities as part of a daily routine, teachers would actually be implementing more effective pedagogies. As the teachers gained confidence in conducting the activities, it was hoped that they would reflect on what they were doing and would recognize the value of the pedagogies.

Table 1a. Reading skills addressed and activities used in the daily routines

Skill	Activities
Phonemic awareness:	• Distinguishing sounds • Blending sounds • Manipulating sounds
Letter sounds:	• Letter sounds with a short diacritic • Letter sounds with a long diacritic • Distinguishing between short and long diacritics • Blended words
Vocabulary:	• Contextualized words • Word families • Synonyms • Elaborating adjectives • Vocabulary networking
Comprehension:	• Predicting the title of a story based on the illustration of the story • Predicting the title of a story based on the text of the story • Summarizing • Self-regulation • Responding both to recall and to inferential questions
Writing:	• Writing letters • Writing words • Writing sentences • Functional writing • Creative writing

Table 1b. Mathematics skills addressed and activities used in the daily routines

Skill	Activities
Rote counting:	• Counting in ones • Counting rhymes and songs • Counting in steps
Rational counting:	• Counting small sets of counters in ones • Counting out small groups of counters • Estimating and counting larger sets of counters in ones • Counting in groups • Counting large sets of counters in groups
Manipulating numbers:	• Single digit arithmetic • Arithmetic with multiples of ten, hundreds, and thousands • Completing tens (hundreds and thousands), including adding to and subtracting from multiples of ten • Bridging tens (hundreds and thousand) • Doubling and halving
Solving problems:	Problems that support the development of: • Addition and subtraction (change, combine, and compare problems) • Division (sharing and grouping) • Multiplication (repeated addition and situations with a grid- or array-type structure) • Fractions, ratio, rate, and proportion, including sharing in a ratio

In some countries, interventions involve the development of lesson scripts for teachers to follow as the way to introduce teachers to new pedagogies. For this 2012 Jordan intervention design, it was decided that, for a variety of reasons, scripted lessons would not be used, most significantly because the education system in Jordan is functional and teachers generally know how to manage their classrooms. The intention of this intervention was not to teach classroom management, but rather to support teachers in developing pedagogical approaches that would support the development of foundational reading and mathematics skills in ways that are supported by research. To achieve this, the intervention relied on a limited number of activities that were designed to develop certain skills. These activities were to be conducted in the same way with grade 1 or grade 3 students, only with a different content.

The research questions of the intervention were to establish the following:

- Does daily practice of foundational skills through deliberate, structured, and developmentally appropriate activities support children to be able to read with comprehension and do mathematics with understanding?
- What are the conditions that help teachers to implement the daily routine and the associated activities with fidelity and confidence?

1.2.2 Materials

The materials for the intervention were developed by two teams of writers, one for reading and one for mathematics. For each subject, the writing team consisted of 10 members of the Jordanian MoE: a leader and three teams of three people—one team for each of grades 1, 2, and 3. Each grade team consisted of a member of the subject curriculum committee, a supervisor, and a teacher teaching in that grade. The team members were nominated by the MoE. The materials were developed in two stages: during stage one, the first semester materials were developed and during stage two, the second semester materials. Each of these development stages required a total of approximately four weeks.

RTI appointed one technical expert per subject to provide leadership and guidance during the materials development process. At the beginning, the technical experts spent most of the first week with the writing teams discussing the components of a research-based approach to teaching early grade reading and early grade mathematics.

For reading, the first week of the design phase was informed by the experiences of RTI on the Early Grade Reading Program (EGRP) in Egypt. EGRP was piloted in 60 schools for grade 1 before being rolled out to grades 2 and 3. In less than one year, grade 1 students in the 60 pilot schools were able to identify, on average, 19 more letter sounds per minute at the end of the school year than they had in the baseline at the beginning of the year, which was an increase of 194%. Meanwhile, students in the control group gained just two letter-sounds per minute, which was an increase of only 21% over the baseline.

After the discussions about the key elements of a research-based approach to teaching early grade reading and mathematics, the materials design teams worked together to develop the structure of the materials to be used in the intervention. It was decided to develop three sets of materials: (1) a teacher's guide (called "teacher notes" in the intervention to avoid confusion for teachers, because teachers already had a teacher guide), (2) daily lesson notes, and (3) a students' workbook.

Teacher's guide (teacher notes)

A teacher's guide (teacher notes) for each subject was developed as a resource to provide teachers both with a pedagogical rationale for the teaching approach of the intervention and with guidance for how to conduct the activities associated with the different skills (***Table 1***). These notes were developed by the technical experts and refined by each of the writing teams.

Pedagogical rationale. The teacher guide for each subject includes detailed notes explaining why each of the skills targeted by the daily routine is critical (foundational) for students' development of reading and mathematics skills. Although the notes are not as comprehensive as they would be if they were part of a university course or a textbook chapter on the importance of the skills, they do provide sufficient background logic to encourage the teachers to include the approach in their daily school activities.

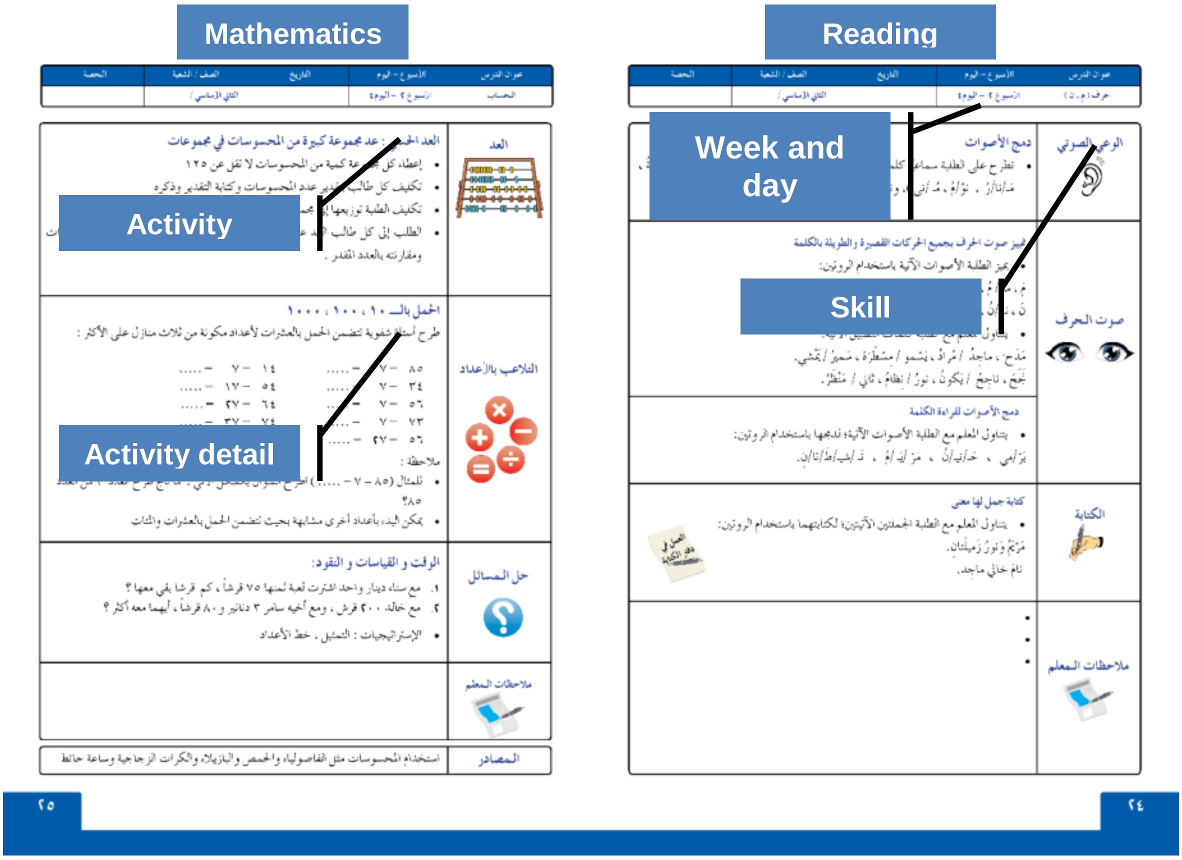
Conducting the activities. As previously discussed above, it was decided in designing this intervention that teachers would not be provided with scripted lessons. This design decision was deliberate, and although it was because the state of schools and teachers in Jordan did not warrant such an approach (teachers know how to conduct lessons), it was also guided by the pragmatics of a time frame that did not allow for the development of such scripted lessons. However, if the daily activities during the first 15 minutes of every lesson were to be conducted with fidelity, then the teachers also needed to have clarity about what was expected of them to conduct these activities. Therefore, the teacher notes were designed to provide a comprehensive guide about how to conduct and manage each of the activities—a series of mini-scripts, one for each activity. Although the lessons were not scripted to provide the exact words that the teacher should say, the lesson notes and teacher notes together created a highly structured environment.

With hindsight, these teacher notes did not have enough of a description of the rationale for conducting the daily 15-minute activity at the start of each lesson. Implementing the intervention revealed that teachers did not understand well enough the rationale for the daily practice of foundational skills, and this lack of understanding, in turn, led to confusions and frustrations. It is not possible to know if more comprehensive descriptive notes could have prevented these confusions and frustrations, but at least a more explicit description for teachers (and trainers) would have been available as a reference.

Daily lesson notes

For each grade, a set of daily lesson notes was developed for the teacher. These were published in a single book, with the reading and mathematics notes for a particular day printed on facing pages. Thus, when the teacher opened the book of lesson notes to the correct day, she would see the reading notes and the mathematics notes for the day next to each other. The book pages were numbered from "week 1–day 1" to "week 16–day 5" for each subject and each semester.

Figure 2. A typical pair of facing pages from the daily lesson notes



On each page in the book, the skills to be included in the 15-minute routine are listed, as well as the activities to be used for each skill. Furthermore, details are provided for each activity with the letters or words, numbers, and problems, etc. to be used during the activity. ***Figure 2*** shows a typical pair of facing pages from the lesson notes for the second semester in grade 2.

Teachers were expected to use the daily lesson notes to guide the contents and structure of the daily 15-minute routine. It was expected that teachers would plan for the daily 15-minute routine by referring to the lesson notes and identifying the skills to be addressed in the 15 minutes and, in particular, the activities to be used to develop those skills. They would then refer to the teacher guide to remind themselves of how to conduct the activity and would think about how to do that for the particular letters/words/numbers/ problems targeted on that day. Because there were a finite number of activities for each skill, it was hoped that over time, the teacher would need to refer to the teacher guide less frequently for how to conduct the activity. In a similar way, it was also anticipated that as the students in the class were exposed to the same activities repeatedly over the semester, students would need less and less explanation about what to do each time, and the 15 minutes would indeed evolve into a daily "warm-up" routine. In large part, this is exactly what happened across the classes where the intervention was implemented. Of course, this evolution did not happen naturally in all classes, and there were teachers who expressed frustrations, the most common of which was about the time taken to complete the routines. Teachers felt that the expectations for each day were such that they could not complete what was expected within 15 minutes. This issue of "not enough time" is further discussed under lessons learned (see section 4).

In developing the lesson notes, the writing teams first developed scope and sequence maps for each grade and each semester to ensure that the range of foundational skills to be practiced each day were both appropriate and revisited frequently during the semester. The scope and sequence were more of a developmental trajectory for the foundational skills.

The reading team faced a particular challenge in developing the scope and sequence, and specifically so for grade 3. The 2012 National Survey had revealed that grade 3 children were struggling with letter sounds and other foundational skills, and thus the daily routine needed to address this skill. In a situation where letter sounds (and the other skills) are practiced more throughout grades 1 and 2, letter sounds would only need occasional attention in grade 3. However, because this intervention was implemented in grades 1, 2, and 3 simultaneously, the grade 3 students needed more work on letter sounds (especially in the first semester) than would have been the case if the grade 3 students came to grade 3 having done the activities in grade 2. For this reason, in grade 3, the only skills that received attention in the first semester were phonemic awareness, letter sounds, and writing (limited to writing letters, words, and sentences). The other skills listed in ***Table 1*—**vocabulary, comprehension, and writing (functional and creative writing)—were introduced only in the second semester. The implication of this approach is that if the materials are to be used for a second and third year in schools—that is in schools where grade 3 students have been following the approach in grade 2, and grade 2 students have been following the approach in grade 1, then the materials will need to be revised.

The mathematics team did not feel that the problem described above for reading applied as much to mathematics as it did to reading, and instead they developed materials for the different years based on normal developmental expectations of students. The advantage of this approach is that, should the materials be implemented on an ongoing basis, the mathematics materials will not need the same revision as the reading materials will. However, and with hindsight, this decision may have contributed to the general feeling expressed by teachers implementing the materials that the mathematics materials were too difficult, especially so in the first semester.

Workbooks

In addition to the resources already described (teacher guide and daily lesson notes) that enable the teacher to conduct the daily 15 minutes of foundational skills practice, it was decided to also develop a series of student workbooks. It was expected that the workbooks would provide a resource for daily independent work for students. That is, in addition to the 15 minutes of whole-class foundational skills practice, students would have a resource that would engage them in independent practice of the skills that the teacher had worked on with the class during the 15 minutes at the start of the lesson. The workbook provided an “additional resource” for teachers and students. It was not expected that the workbook should be used as part of the 15 minutes, but instead that the workbook could be used at another time in the lesson when the teacher wanted to assign work for the students, or as homework, etc. Nevertheless, it was expected that the students should work through the workbook at a pace of one page per day. It was also expected that teachers would monitor students' work in the workbooks and provide feedback to students.

Because the 15 minutes of daily practice of foundational skills in reading are focused exclusively on phonemic awareness and letter sounds in the first semester (see discussion in the previous "Daily lesson notes" section), the workbook for reading in the first semester are focused exclusively on developing writing skills (letters, words, and sentences) only. The second semester workbooks for reading also include short stories to be read, comprehension activities related to those stories to be completed, and a range of vocabulary and creative and functional writing activities.

1.2.3 Training and Support

Similar to how the materials were developed by MoE personnel with technical support from RTI, the training and ongoing support of teachers was also carried out exclusively by MoE supervisors with technical support from RTI.

Training was conducted in two stages: (1) training of trainers and (2) training of teachers. The training of trainers (MoE supervisors) was provided by the RTI technical experts, while the training of teachers was provided by the MoE supervisors. There were two sets of training. The first set was held before the first semester of the 2013/2014 school year (from June to August of 2013), and the second set was held before the second semester (in February 2014).

Schools for the intervention pilot were selected from the 2012 National Survey sample such that there were at least two, and preferably four, or more schools in a school district (field directorate) with at least one supervisor available to provide training and support to two schools. The details of the treatment school selection process are provided in section 2.2.2 of this report.

A total of 20 MoE supervisors were assigned to the intervention. They were responsible for training more than 300 teachers in the 43 schools across 12 education districts. ***Figure 3*** highlights the education districts in the intervention study, where districts colored with shades of green each had two or more supervisors, while districts colored with shades of blue had only one supervisor associated with them. ***Table 2*** indicates the number of supervisors, schools, and teachers in each education district in the intervention. The intervention district numbers in the table correspond to the numbers on the map.

First semester training

Supervisor (trainer) training for the first semester of the 2013/2014 school year took place over two sessions. The first 10-day session was conducted at the end of June 2013 to coincide with the end of the 2012/2013 school year and to be completed before Ramadan. The second 5-day training took place after Ramadan and was both a refresher course to ensure that the supervisors were still confident with the material that had been covered a month earlier and used to make the many logistical arrangements for the teacher training. Teacher training took place over two weeks in August 2013 to coincide with the start of the school year. The training during the first week coincided with the week that teachers come to school before the students do, and the second week coincided with the first week that students were at school.

Training for the first semester of the 2013/2014 school year focused on the structure of the intervention and involved extensive modelling and practicing different activities associated with each of the reading and mathematics skills.

Table 2. Intervention districts and supervisor, school, and teacher allocation

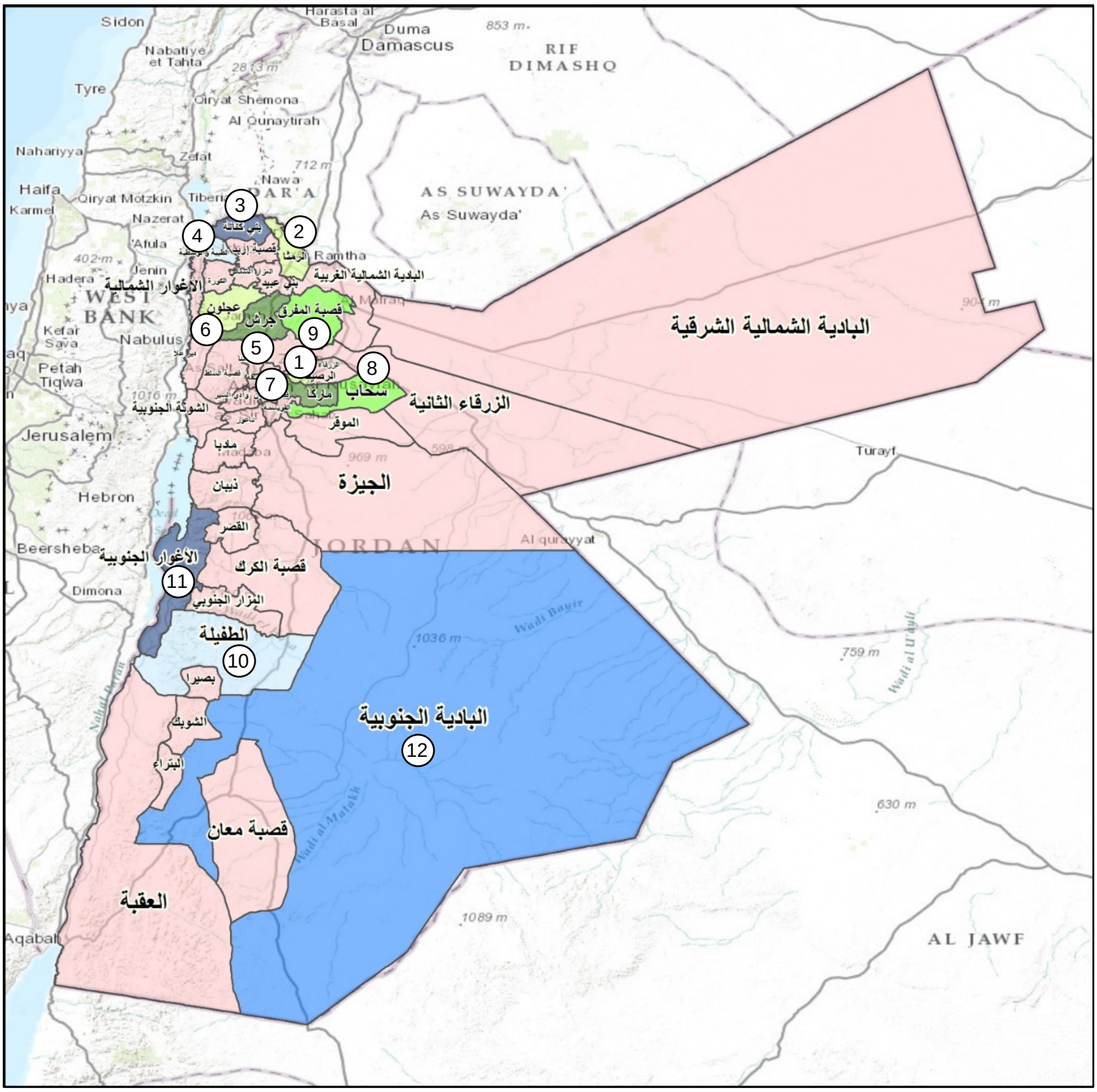
Intervention district	Number of supervisors	Number of schools	Number of teachers
1	2	4	38
2	2	4	21
3	1	2	13
4	1	4	17
5	2	4	20
6	3	3	34
7	2	6	59
8	2	6	64
9	2	4	23
10	1	2	8
11	1	2	16
12	1	2	34
Total	**20**	**43**	**347**

In designing the intervention, deliberate consideration was given to selecting districts with schools and supervisors such that at least two supervisors were working together in the same district, with each one supervising two schools. The rationale for this was two-fold. Firstly, having two supervisors working together meant that they would be able to discuss issues and support each other through the duration of the intervention. Secondly, having two supervisors to conduct the training together would help mitigate losses (typically associated with the cascade training model) in the message of the intervention as it is conveyed from one level to the next. Because of a special request by the MoE, three districts in the South were deliberately selected, despite having only one supervisor for each district. Schools in the South traditionally struggle to achieve on the Tawjihi[2] examinations, and including these schools in the intervention

[2] Tawjihi is the term for the Jordanian general secondary examination that Jordanian students take at the end of their high school education. Students who pass this examination are awarded the Tawjihi—the Jordanian General Secondary School Certificate. The Tawjihi, or its equivalent, is required for students who plan to attend university in Jordan for their undergraduate and graduate degrees.

provided an interesting opportunity to see if the intervention could be successful in this environment.

Figure 3. Training districts for the intervention



Legend: Green districts each have two or more supervisors, and blue districts have only one supervisor associated with them.

To compensate for some districts not having two supervisors, it was thought to combine the training for teachers from schools in neighboring districts (e.g., combining the training of teachers in districts 2 and 3, etc.). Sadly, that strategy was simply not possible because of two key obstacles to this plan. The first was because female teachers are not allowed to travel the distances that this implied. The second involved supervisor resistance to the plan. Supervisors who train teachers for the MoE are paid a daily training allowance. If two trainers share the training, the allowance is halved and each trainer gets one half of the allowance that they would have received had they conducted the training by themselves.

Even in districts where there were two or more supervisors allocated, they were unhappy to do the training together for the reason already mentioned. In the end, they proceeded to conduct the training together, albeit reluctantly.

During the two weeks of teacher training in August 2013, the RTI technical experts visited each of the training centers to monitor training, offer support to the supervisors doing the training, and answer questions from the participating teachers.

Second semester training

A second round of training was held before the start of the second semester. Firstly, this second round served to introduce supervisors and teachers to the newly developed second semester materials, and secondly, it served to address issues and concerns that had arisen during the first semester.

The training of supervisors (trainers) for the second semester of the 2013/2014 school year was conducted at the end of January 2014. Teacher training took place during the first week of February, to coincide with the start of the second semester.

Once again, the RTI technical experts visited each of the training centers to monitor training and offer support to the supervisors conducting the training. On this occasion, however, the technical experts also held a mini–focus-group type of discussion with the participating teachers, to encourage them to discuss their experiences from the first semester. These discussions were important and highlighted a number of misconceptions. Despite teachers expressing the feeling that the intervention was a good program, that students enjoyed the activities, and that students were benefiting from them, two recurring sets of concerns emerged:

- Firstly, the teachers did not see a relationship between the content of the daily 15-minute routine and the curriculum (text book page)[3] for the day. This was of great concern to them.
- Secondly, the teachers noted that the daily routine often took more than 15 minutes to complete, and they were concerned about this as they felt that it impacted their ability to "complete" the curriculum.

These teacher concerns, although very real, are also regarded as misconceptions about the program and will be discussed under lessons learned (see section 4)

Ongoing classroom support

In addition to the training that supervisors provided to teachers, they also provided in-class coaching and support and coordinated a monthly reflection session among the participating teachers in their district.

Supervisors visited participating teachers at their schools to observe them conducting the daily routine and to provide general support to them in implementing the intervention

[3] With "curriculum," teachers in Jordan mean the MoE textbook used in all schools, not a curriculum such as a syllabus. In Jordan, the textbook (curriculum) is followed in a systematic way, with all teachers in the same grade being on the same page on the same day. Thus, the textbook page of a specific day is seen as the "curriculum" for that specific day.

activities. Supervisors were encouraged to visit each teacher at least once every two weeks. For each visit, the supervisors completed an observation form and submitted an encoded summary of that form to the project staff via short message service (SMS). These SMS and the data they contained provided the project staff with a monitoring mechanism that not only allowed them to know which teachers were being visited how often and by which supervisor, but also allowed them to monitor the quality of implementation. Reports indicating supervisor visiting frequency, as well as how well each teacher was implementing the routines, were submitted to the ETC every quarter.

The frequency, with which supervisors actually visited schools, varied significantly from one supervisor to the next. The impact of the frequency of these visits is an important focus of the analysis in section 3 of this report.

Finally, in addition to the school visits, supervisors also arranged a monthly meeting with the participating teachers from the schools that were in their district, to discuss implementation issues and lessons learned. The frequency of these monthly sessions, as well as teacher attendance, varied significantly from one supervisor to the next.

1.3 Implementation

The intervention was implemented during the 2013/2014 school year by more than 400 teachers in 347 classrooms across 43 schools, reaching approximately 12,000 students. The impact of this intervention project will be analyzed in section 3 of this report.

1.4 Implementation Challenges

While on the one hand the implementation of the intervention went largely according to plan, on the other hand, it would be naïve not to acknowledge that the intervention experienced a range of challenges. The challenges faced by the intervention fall into two main categories—logistical challenges and challenges related to transferring the vision of the intervention into practice.

The logistical challenges were related to the bureaucratic and centralized nature of the decision making within the MoE. Although the intervention was, from the outset, an activity of the MoE, with MoE personnel involved in the design of the intervention, developing the materials, conducting the training and coaching of the teachers, and participating in the monitoring and evaluation activities of the project, the MoE nonetheless also contributed to and created some of the logistical challenges in implementing the intervention. To illustrate, the dates for the February 2014 training of teachers were agreed with the MoE in December 2013. However, on the Thursday preceding the Sunday on which training was to begin, the letter required to authorize that the training could take place had not yet been signed by the Secretary General of Education. This delay compromised the ability of training to start as planned on Sunday, because by the time the letter was signed, it was no longer possible to inform all the teachers involved about the starting time and training venue. The result was that many teachers could only join the training from the second day of the 5-day course.

Handling the reality of teacher changes proved to be another logistical challenge encountered in implementing the intervention. Training during the first week of the 2013/2014 school year had the consequence that many teachers assigned to teach in the intervention schools did not know that they were teaching in those schools until the second or third day of the week, which meant that they only started training on the third or fourth day of the 10-day training. In addition, a large number of substitute teachers are appointed to posts, and despite the encouragement from the project leadership that these teachers should attend training, there was reluctance from the supervisors to include them. In part, this was because the substitute teachers did not know how long they would be at their post, and in part it was because the supervisors seemed to be unwilling to train teachers who would not be at their posts for long. Over the course of the year, nearly 400 teachers taught in the approximately 300 classrooms involved in the intervention. This level of turnover made it difficult to ensure implementation continuity across the classrooms that were participating in the intervention.

Although the logistical challenges were very real and had their impact on the project, the greater challenge that the project faced involved successfully sharing the vision of the intervention with the participating teachers. The vision of the intervention was two-fold. First, the intervention hoped to improve students' performance on foundational reading and mathematics skills so that more students would read with comprehension and do mathematics with understanding than was the case in the 2012 National Survey. Second, the intervention hoped to expose teachers to alternate, research-based, approaches to teaching early grade reading and mathematics, in the hope that these approaches, practiced in the daily 15-minute routines, would impact the teacher's approaches to teaching reading and mathematics in general.

In all meetings that the project staff had around implementing the intervention, there were always three recurring themes. Whether it was in meetings with the MoE head office personnel, who had heard from teachers and supervisors; or in the debriefing meetings held with supervisors, in which they spoke about their experiences with teachers; or in the comments from teachers to the project staff, when they visited the training centers, the message was the same: (1) the intervention did not match the curriculum, (2) the intervention activities took too long, and (3) the students experienced the activities as either being too hard (which was more often the case for mathematics) or too easy (which was more often the case for reading).

When teachers in Jordan speak about the curriculum, they are referring to the MoE textbook used in all schools. They are not referring to the curriculum in the sense of a syllabus. Teachers in Jordan follow the textbook (curriculum) in a deliberate and systematic way, with all teachers in the same grade being on the same page on the same day. So, when teachers said that the intervention activities did not match the curriculum, they meant that the content (concepts and skills) of the intervention materials for a particular day did not align with the content of the *textbook page* for that day. There was, however, from the intervention design perspective, never any expectation that the two would align. The 15 minutes of intervention activity were intended to be a general reinforcement of foundational skills in reading and mathematics—a daily "warm-up" activity before the curriculum-based lesson starts. As discussed earlier in this report, the intervention was also intended to address the gaps that had been identified in the curriculum. In grade 3 reading, the 15 minutes "warm-up" involved activities related to letter-sound production, a skill that the curriculum does not spend time on in grade 3, but a skill that the 2012 National Survey had found to be lacking in grade 3 students. Although, according to the mathematics curriculum, grade 2 students are expected to be adding and subtracting three- and four-digit numbers, in the intervention students were developing the skills to manipulate numbers fluently and flexibly using a wide range of different calculation strategies with two-digit numbers only. While these examples could be considered (and were by teachers) as illustrations of how the intervention was not well aligned to the curriculum, the point is that they were deliberately not aligned. The activities of the intervention were different expressly because the intervention was hoping to address identified gaps in students' foundational knowledge—gaps that the curriculum had not and was not addressing. Teachers did not find this concept easy to embrace.

It was to be expected that teachers would struggle to complete the intervention activities for the day in 15 minutes—at least at first. At first, as teachers were inexperienced with the activities of the intervention, and as the activities were new to the students in their classes, it was reasonable to expect that they would struggle to complete them in 15 minutes. To be clear, 15 minutes was always considered a metaphor for a "short period" of time, ranging from 10 to 20 minutes. Thus, it was expected that as teachers gained confidence with the different activities, and as students were exposed to the same activities again and again, students would learn to do them with greater confidence, automaticity, and fluency, and the time to complete the daily activities would be reduced. Of course, it was also hoped that as teachers began to see the benefits of the activities, they would also start to incorporate activities they had used in the intervention into the delivery of the curriculum component of the lesson, as well—that is, they would adopt the research-based pedagogical approaches of the intervention in their general teaching.

It is worth noting that teachers were encouraged to not exceed 15 minutes for the intervention activities each day. They were encouraged to stop after 15 minutes, even if they had not yet completed the activities for the day. The rationale for stopping after 15 minutes was that teachers should not be worried about "completing the curriculum" because they felt that they were losing too much time in carrying out the intervention.

The above noted, the real problem that prevented teachers from limiting themselves to 15 minutes per activity was twofold. First, classroom observations revealed that teachers explained in great detail how each activity should be conducted each and every time that they did the activity with their class—no matter how many times the class had already done the activities. This was not the intention of the intervention. By limiting the number of activities used to develop each skill, it was hoped that teachers and students alike would become so familiar with the activities that they would simply do them without having to spend time discussing how the activity would be done. Second, teachers also struggled to stop an activity unless all of the students in the class were confidently engaging with the concept or skill. The expectation with the activities of the daily routines was not that each child in each class would master each concept or skill before continuing to the next one. Rather, the rationale of the intervention was that students would practice concepts and skills each day and repeat the practice day after day. Instead of mastering a concept or skill in a single day students would gain confidence with the concepts and skills over time and eventually use them fluently and flexibly in a range of different settings.

The idea that a skill or concept is developed over time and not all at once is not one with which teachers are necessarily comfortable; for this reason, among others, they struggled to contain the daily session to 15 minutes.

Finally, teachers complained that students experienced the activities as too difficult. The daily word-based problems in mathematics provide a good illustration of this issue. The role of word-based problems in the mathematics activities was to present students with a problem situation and to challenge them to understand the problem, to make an age and developmentally appropriate plan, and to solve the problem in a way that makes sense to them. Word-based problems used in this way provoke students to "do the mathematics" that teachers want them to learn, without explicitly teaching them about mathematics yet. By way of illustration, if a teacher asks a grade 1 child: "If mother has 12 falafel that she shares equally between three children, how many falafel will each child get?" then the grade 1 child is easily able to solve the problem. Either they will model the situation with counters (objects), or they may draw a picture. A grade 2 child, by contrast, might draw a representation of the sharing by using numbers and conclude that each child will get four falafel. If students are sincerely allowed to make a plan and solve the problem, then this problem does not seem difficult to a grade 1 student or even a kindergarten student. However, in a more traditional "word problem" situation, teachers teach children to read a problem, to set up an equation to represent the problem, and then to solve the equation. In the case of a mother with the 12 falafel being shared by three children, the equation would be: $12 \div 3 = \square$. Of course if the expectation is that children should write an equation to represent the problem and then to solve that equation, then this problem would be beyond grade 1 capabilities because children do not yet do division in grade 1. If the expectation is that children will make a plan and solve it, then the problem is accessible and teaches important mathematical habits of mind; if by contrast, the expectation is that children will set up and solve an equation, then the problem is inaccessible.

Clearly, the expectation of the intervention for word-based problems was that children should be allowed to make age and developmentally appropriate plans to try to solve these problems. However, because the intervention method was very different from the predominant teaching approach that teachers used, the teachers found it hard to embrace the philosophy of the intervention, and in turn, complained that the materials were too difficult.

The implementation challenge represented by these recurring themes is not unusual in contexts where attempts are made to introduce teachers to alternate pedagogies. Indeed, challenges are to be expected. The discussion here serves to foreground the difficulties that teachers had in assimilating new pedagogies and some of their criticisms about the intervention, as these were not primarily problems with the intervention, but rather were mismatches between intervention vision and teachers' predominant teaching styles and approaches.

1.5 2014 Endline Assessment Tools and Other Sources of Data

To measure the impact of the intervention pilot, an endline study was conducted in May 2014. Because the intervention was only concerned with improving reading and mathematics, the endline survey included only the EGRA and EGMA.

The report for the 2012 survey had discussed in detail the rationale and background for EGRA and EGMA,[4] thus it was felt that the rationale would not need to be addressed again in this 2014 report. Although the EGRA and EGMA of 2012 were both abbreviated versions of the instruments because they formed part of a larger SSME study, in 2014, both the EGRA and the EGMA reverted to full versions to provide a richer data set.

1.5.1 Early Grade Reading Assessment

Table 3 compares the 2012 and 2014 versions of the EGRA. Two additional subtasks were included in the 2014 EGRA version, to assess skills that the intervention has directly addressed and which were not assessed in 2012: syllable names and dictation.

Table 3. EGRA instrument subtasks

EGRA Subtask	2012	2014	Skill	Description The child is asked to…
Letter-sound identification (*timed*)	✓	✓	Alphabetic principle—letter-sound correspondence	…say the sound each letter makes, while looking at a printed page of 100 letters of the alphabet in random order, upper and lower case.
Non-word reading (*timed*)	✓	✓	Alphabetic principle—letter-sound correspondence and fluency (automatic decoding)?	…read a list of 50 non-words printed on a page. Words were constructed from actual orthography, but were not real words.

[4] Brombacher, A., P. Collins, C. Cummiskey, E. Kochetkova, and A. Mulcahy-Dunn. 2012. *Student Performance in Reading and Mathematics, Pedagogic Practice, and School Management in Jordan*. Section 2.2 and 2.3. Prepared by RTI International for USAID. Available at https://www.eddataglobal.org/countries/index.cfm?fuseaction=pubDetail&ID=425 (accessed August 15, 2014).

EGRA Subtask	2012	2014	Skill	Description The child is asked to…
Oral reading (*timed*)	✓	✓	Fluency (automatic word reading in context)	…read out loud a grade-level appropriate short story printed on a page.
Reading comprehension (*untimed*)	✓	✓	Comprehension	…verbally respond to five questions that the assessor asks about the short story.
Listening comprehension (*untimed*)	✓	✓	Oral language comprehension and vocabulary	…listen to a story that the assessor reads out loud, and then verbally answer five questions about the story.
Syllable names (*timed*)		✓	Beginning decoding skills and identifying syllables from the language	…read a list of 50 syllables presented in random order.
Dictation (sentence) (*untimed*)		✓	Spelling, orthographic/ phonological knowledge, language knowledge, and grammar skills	…write, spell, and use grammar properly in a dictation exercise.

1.5.2 Early Grade Mathematics Assessment

Table 4 compares the 2012 and 2014 versions of the EGMA. Although the subtasks of the 2012 version were all included in the 2014 version, some of the subtasks were assessed differently in 2014. In particular, a number of subtasks that had been timed in 2012 were not timed in 2014.

Table 4. EGMA instrument subtasks

EGMA Subtask	2012	2014	Skill	Description The child is asked to…
Subtasks that assess procedural (recall) knowledge				
Number identification	Timed 30 sec	Timed 60 sec	The ability to identify written number symbols. If students cannot identify numbers, they cannot do mathematics.	…say the names of numbers presented on a page with 20 numbers. The numbers range from one- to two- and three-digit numbers.
Addition and subtraction L1 (basic facts)	Timed 30 sec	Timed 60 sec	Knowledge of and confidence with basic addition and subtraction. It is expected that students should develop some level of automaticity/ fluency with	…solve addition/subtraction problems, with sums/ differences below 20, without the aid of paper and pencil. The items range from problems with only single digits to problems that involve the bridging of the

EGMA Subtask	2012	2014	Skill	Description The child is asked to…
			mathematics facts such as these, since they are foundational mathematics skills.	ten.[5] (Ten items per addition and subtraction subtask)
Subtasks that assess conceptual (applied) knowledge				
Quantity discrimination (number comparison)	Timed 60 sec	Not timed	The ability to make judgments about differences by comparing quantities, represented by numbers.	…identify the larger of a pair of numbers. The number pairs used range from a pair of single-digit numbers to five pairs of double-digit numbers and four pairs of three-digit numbers. (Ten items)
Missing number (number patterns)	Timed 60 sec	Not timed	The ability to discern and complete number patterns.	…determine the missing number in pattern of four numbers, one of which is missing. Patterns used include counting forward and backward by ones, by fives, by tens, and by twos. (Ten items)
Addition and subtraction L2[6]	Timed 60 sec	Not timed	The ability to use and apply the procedural addition and subtraction knowledge assessed in the L1 subtask to solve more complicated addition and subtraction problems.	…solve addition/subtraction problems that involve the knowledge and application of the basic addition and subtraction facts assessed in the Level 1 subtask. Students were allowed to use any strategy that they wanted, including the use of paper and pencil supplied by the assessor. The problems extended to the addition and subtraction of two-digit numbers involving bridging. (Five items per addition and subtraction subtask)

[5] "Bridging the ten" refers to addition and subtraction situations where the addition and subtraction involves moving from one decade to the next. For example, 8 + 6 and 28 + 6 both involve "bridging the ten." A common strategy that may be adopted by children when bridging the ten mentally is first to "make" or "complete the ten"—e.g., 8 + 6 = 8 + 2 + 4 = 10 + 4 = 14, and 28 + 6 = 28 + 2 + 4 = 30 + 4 = 34.

[6] The addition and subtraction Level 2 subtasks were more conceptual than the addition and subtraction Level 1 subtasks because the pupils had to understand what they were doing and apply the Level 1 skills. In other words, while the Level 2 subtasks were not purely conceptual—because with time and practice, pupils will develop some automaticity with the types of items in these subtasks—they were more conceptual than the Level 1 subtasks, especially for grade 2 pupils.

EGMA Subtask	2012	2014	Skill	Description The child is asked to…
Word problems	Not timed 3 items	Not timed 6 items	The ability to interpret a situation (presented orally to the student), make a plan, and solve the problem.	…solve problems presented orally, using any strategy that they wanted, including the use of paper and pencil and/or counters supplied by the assessor. The numerical values involved in the problem were deliberately small to allow for the targeted skills to be assessed without confounding problems with calculation skills that might otherwise impede performance. The problem situations used were designed to evoke different mathematical situations and operations. (Six items)

The changes between the 2012 and 2014 EGMA versions were brought about because (1) there was less time pressure on the assessors during the 2014 testing cycle at schools because the wide range of SSME instruments that had been administered in 2012 were not administered in 2014, and (2) removing the timing on some of the subtasks allowed students a broader opportunity to demonstrate their mathematics skills.

1.5.3 Additional Sources of Data

To gain a fuller understanding of the impact of the intervention and the variables that influence the chances of the intervention's success, data was gathered from a wide range of sources. These sources include the following:

- **Oral teacher questionnaire**: Completed by the grade 2 and 3 teachers of the assessed pupils in the treatment schools. The oral questionnaires focused on three aspects of teachers in treatment schools: (1) the teachers' background characteristics (e.g., experience, training), (2) their implementation of the reading and mathematics components (e.g., how often they followed the routines, whether they used the lesson notes), and (3) their perceptions of the intervention.
- **Written teacher questionnaire**: Completed by the grade 1, 2, and 3 teachers who implemented the intervention in treatment schools. In addition to collecting some background information about the teachers, the written questionnaire asked teachers to discuss both the reading and the mathematics components of the intervention in terms of the materials, the implementation, their perception of student response to the project, and the role of the supervisors in training and supporting the teachers as they implemented the project. For each of these (materials, implementation, student response, and supervisor role), teachers were

asked to list what they had experienced as positive and as negative and to make recommendations for possible changes.

- **Oral student questionnaire (full version)**: Completed by all students in treatment schools, this questionnaire was used to gauge students' reaction to the intervention. It was difficult to ask students directly about the intervention, because they did not experience the intervention as peculiar—the intervention was simply part of their reading and mathematics lessons. For this reason, the questionnaire tried to find out if the students had or had not actually done the activities in the routine and if they had or had not enjoyed them. In addition, the questionnaire also tried to find out how the students worked with the workbooks and whether or not teachers were marking the workbooks. The students who participated in the oral student questionnaire were also asked questions related to typical wealth variables, to allow the development of a wealth indicator for students in the data set to serve as a control factor for any regressions calculated in the study.
- **Oral student questionnaire (only wealth variables version):** Completed by all students in control schools, this questionnaire consisted of only the questions related to typical wealth variables to allow the development of a wealth indicator for students in the data set to serve as a control factor for any regressions calculated in the study.
- **Supervisor visit reports**: The responses to the teacher observation sheet that were completed by supervisors each time they visited a teacher (submitted by SMS) also provide a good source of data about the fidelity with which the intervention was implemented.
- **Teacher participation data**: A teachers' database was developed and maintained that includes information for the intervention about:
  - Teachers' status—if they are still involved in the intervention or if they have left their school and are no longer implementing the intervention;
  - The number of days of training that teachers attended out of two training sessions (August 2013 and February 2014); and
  - The number of supervisor visits that teachers received.
- **Project team field notes**: Field notes maintained by the project staff throughout the project as they visited training sessions, schools, and classes, and includes, in particular, notes taken during observations of teachers implementing the intervention at five different schools toward the end of the second semester.

2 Methodology

2.1 Research Design

Table 5 represents the research design of the baseline and endline assessments, as well as the intervention activity. The research was designed to assess the efficacy of the intervention activities, in particular:

- Does daily practice of foundational skills through deliberate, structured, and developmentally appropriate activities support children to be able to read with comprehension and do mathematics with understanding?
- What are the conditions that help teachers to implement the daily routine and the associated activities with fidelity and confidence?

Table 5. Implementation of intervention, baseline, and endline studies

	2011/2012 School Year		2012/2013 School Year	2013/2014 School Year	
Intervention activity	–		Materials development	Implementation	
Student assessment		**May 2012**			**May 2014**
Treatment		2,159			2,129
Control		817			806
Total		2,976			2,935

2.2 Sample

2.2.1 2012 Sample

In 2012, the Jordan Education Management Information Systems (EMIS) unit provided a list of all public primary schools in the nation, totaling 2,227 schools. Of these, 162 schools were removed from the list because they did not have grade 2 enrolment, and 31 additional schools were removed because they did not have grade 3 enrolment. A total of 2,043 schools remained in the final population, from which a study sample was drawn. The 2,043 schools contained an estimated 175,571 grade 2 and grade 3 students.

Before drawing the random sample of schools to be included in the study, the 2,043 schools were stratified by region (North, Middle, and South) and school gender (all-boys, all-girls, and mixed schools) to form nine strata. For each region, the goal was to draw a sample of 15 all-boys schools, 15 all-girls schools, and 20 mixed schools, to allow for maximum statistical power within each stratum. However, because of the small number of all-girls schools and all-boys schools in the South, only 11 all-boys schools and 14 all-girls schools were selected in that region. Additional schools were added to different strata that resulted in a total of 156 randomly sampled schools and 3,063 students.

Within each stratum, schools were sorted by district and the combined enrolment of grades 2 and 3. Schools were then selected with equal probability proportional to grade 2 and grade 3 enrolment. For each selected school, two replacement schools were selected, to be used if the sampled school were not available to participate or were not eligible. A total of nine schools were replaced for the following reasons: six schools did not have grade 2 or grade 3 enrolment; two schools were assessed during the pilot study; one school was closed indefinitely.

The second stage of selection involved sampling class/teachers within each sampled school. After the research team arrived at each selected school, all of the grade 2 classes were listed and one grade 2 class was selected at random with equal probability. The selection process was repeated for the grade 3 class.

The third stage of selection involved random selection with equal probability of students from the randomly selected classes. After a grade 2 class was randomly selected, the assessor would go to the selected class and randomly select 10 students from that class. If 10 or fewer students were present, then the assessor would automatically select all of the students in that class. The same procedure was followed for the grade 3 class.

2.2.2 2014 Sample

After the baseline measures were obtained for the initial 2014 sample, purposive sampling was used to select treatment schools from within the sample for the intervention. First, education districts were selected that had at least three supervisors working in the district and had a minimum of four sample schools. Next, the Education Training Center (ETC) was asked to exclude from the sample any schools that were participating in existing projects or were receiving educational interventions. At this selection stage, 12 of the 38 districts in the initial sample met the selection criteria, and treatment schools were identified in the districts. In five of the 12 identified districts, all of the sample schools were selected for the intervention. In a further three districts, all but one of the schools were selected, and the excluded school was excluded on the basis of its geographic distance from the other schools. In the remaining four districts, the treatment schools were selected based on their proximity to each other and to facilitate the ease of support to the school that would be provided by the supervisors that the Ministry of Education (MoE) had assigned to the intervention. Ultimately, 45 schools were selected to receive the intervention. Two schools dropped out at the beginning of the program for reasons that are discussed elsewhere in this document. Thus, a total of 43 schools received the intervention. It was necessary to exclude the data for two treatment schools from the analysis because there was no baseline data available for them.

During data collection, classes and students in the 110 control schools were selected using the approach described above for the 2012 baseline.

The classes for the intervention schools were purposefully selected to represent, as far as possible, those classes in which the intervention conditions could be as ideal as possible. This was to ensure that the endline survey measured what could be achieved if the intervention were implemented under the best possible conditions. To achieve this selection, teachers and classes in the intervention schools were classified according to four criteria: (1) whether or not the teacher had attended all the training; (2) whether or not the teacher at the endline survey was the same teacher who had started the year as the class teacher; (3) the frequency with which the teacher had been visited by the supervisor; and (4) the ratings that the supervisors had given for the teachers. Based on these classifications, classes in the treatment schools were ranked from “most preferred” to “least preferred,” and the assessor team selected the class and teacher (for the oral interview) that were highest on the list (in each grade), taking into account whether or not the teacher was at the school on the day of the survey and or whether or not the teacher was willing to consent to being interviewed. Students in the selected classrooms were chosen at random, with equal probability in the same manner as had been applied in 2012.

In summary, the final sample for this study consists of 151 schools (110 control schools and 41 treatment schools). Data were collected for a total of 5,911 students across the two years (2012–2014), with 2,976 students in 2012 (2,159 control; 817 treatment) and 2,935 students in 2014 (2,129 control; 806 treatment) (see ***Table 5***). Additional data sources included:

- Oral teacher questionnaires: completed by 72 grade 2 and grade 3 teachers.
- Written teacher questionnaires: completed by 233 grade 1, 2, and 3 teachers.
- Teacher observation/coaching reports from 2,171 classroom visits.
- Field notes and reports from project staff who visited training sessions and schools to observe the training and implementation of the intervention.

2.3 Descriptive Statistics

Table 6 summarizes the student characteristics across grades and treatment status. The proportion of female students, proportion of schools by gender (i.e., all-boys, all-girls, and mixed gender), and the age of students is strongly similar across grades. The only salient difference noted in this table is the higher proportion of students from all-girls schools in the treatment group, compared with the control group.

Table 6. Descriptive statistics for final sample—student level

Category	Characteristic	Grade 2	Grade 3	All
All schools	Female	52.8%	54.4%	53.5%
	Male	47.2%	45.6%	46.5%
	All-boys school	15.6%	19.2%	17.4%
	All-girls school	12.2%	13.0%	12.6%
	Mixed school	72.2%	67.8%	70.1%
	Age	7.8	8.9	8.3
Treatment schools	Female	56.2%	53.2%	54.7%
	Male	43.8%	46.8%	45.3%
	All-boys school	16.4%	18.0%	17.2%
	All-girls school	19.1%	18.6%	18.9%

Category	Characteristic	Grade 2	Grade 3	All
	Mixed school	64.5%	63.4%	63.9%
	Age	7.8	8.9	8.3
Control schools	Female	51.2%	54.9%	53.0%
	Male	48.8%	45.1%	47.0%
	All-boys school	15.2%	19.8%	17.5%
	All-girls school	9.0%	10.4%	9.7%
	Mixed school	75.8%	69.8%	72.9%
	Age	7.8	8.8	8.3

2.4 Weighting

To make the sample representative of the national population, weights were calculated as the inverse of the selection probability for each student. Three stages of weighting were applied (stratum, school, and student) so that the sample of student scores could be representative of the overall national level of student performance. All scores reported for this study are calculated using the student weights as noted.

2.5 Equating Procedures

The purpose of test equating is to calculate comparable scores on different forms of a test (in the case of this study, the baseline and endline assessments). Equating is done to ensure that differences in scores are the result of differences in ability and not differences in test difficulty.

Based on the Jordan assessments' construction and piloting, two different equating designs were used in this study: (1) chained linear equating and (2) a linear prediction model. Although the linear prediction model is not technically an equating procedure, it does provide rescaled scores that are comparable across different forms of the assessment. This technique was used for the results of three subtasks: (1) oral reading fluency, (2) reading comprehension, and (3) listening comprehension. For these subtasks, the same pilot students within each of the two years (2012 and 2014) took both the 2012 and the 2014 test forms. This process represents two single-group designs, one in 2012 and another in 2014, either of which could be used to conduct equating. Therefore, it was possible to calculate how the same students scored on each of the three measures and to determine whether the 2014 test version was easier or more difficult than the 2012 test version, in order to adjust the 2014 scores accordingly. Ultimately, only one of these three subtasks was adjusted: oral reading fluency. Reading comprehension scores were not adjusted because the averages and distributions were so similar across versions; listening comprehension scores were not adjusted because the distributions differed in unexpected ways and ultimately pointed to evidence that the two versions may have been measuring slightly different constructs.

The remainder of the subtasks were not piloted jointly but instead involved non-equivalent groups (the 2012 and 2014 students) with an anchor test, in what is referred to as a Nonequivalent Anchor Test (NEAT) design. For the 2012 and 2014 forms, an anchor test, or subset of items common to each test form, was used to estimate and control for any ability differences over time. As a result, the difference in form difficulty, controlling for any differences in ability, could be estimated using chained linear equating and compensated for prior to reporting the results. This equating was carried out using students only from control schools, to mitigate concerns about inappropriately adjusting score differences that resulted from the intervention. It was not possible to equate three of the subtasks: (1) letter sounds (anchor items used multiple times across years), (2) quantity comparison (too few anchors), and (3) word problems (too few anchors).

With the exception of the letter sounds, listening comprehension, quantity comparison, and word problem scores that were not equated, all scores reported for this study are based on the equated values for the subtasks.

2.6 Difference-in-Differences (DID) Approach for Determining Intervention Impact

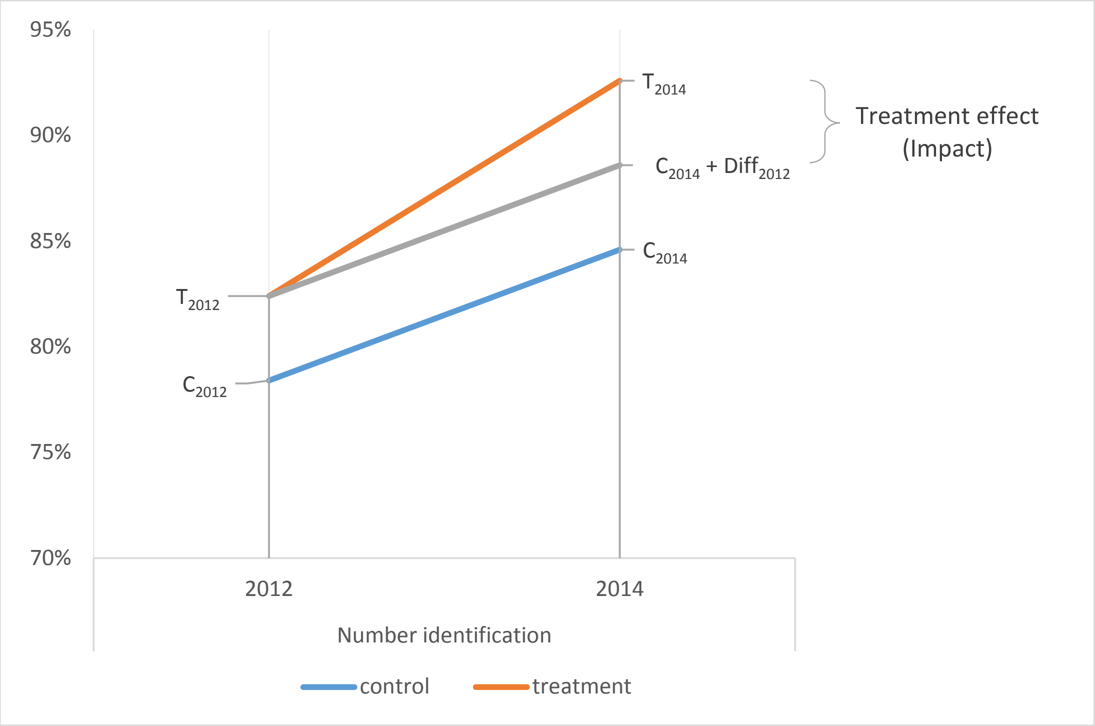
The difference-in-differences (DID) approach for determining the impact of an intervention limits the effect of selection bias by comparing the average change over time in the outcome variable for the treatment group with the average change over time in the outcome for the control group. This method estimates the baseline difference between treatment and control groups (prior to the intervention) and compares it with the post-treatment differences. The rationale behind this approach is illustrated in ***Figure 4***. The difference at baseline between the groups is $Diff_{2012} = T_{2012} – C_{2012}$ (Diff = difference; T = treatment group; C = control group). Without an intervention, the difference between the groups at endline would be expected to be the same as the difference between the groups at the baseline, that is $T_{2014} = C_{2014} + Diff_{2012}$. However, an effective treatment would raise scores more for the treatment group, and the treatment effect could be determined by calculating the difference between T_{2014} (the treatment group outcome) and $C_{2014} + Diff_{2012}$ (the treatment group expected outcome, based on the initial difference between the treatment and control groups).

In this report, the treatment effect (or impact) for each subtask is calculated as the difference between the average scores for students in treatment and control schools in 2014, minus the difference between the average scores for students in treatment and control schools in 2012. For example, the treatment effect (or impact) (TE) for number identification (NI) would be calculated as:

$$TE = (NI_{T2014} – NI_{C2014}) – (NI_{T2012} – NI_{C2012})$$

Where TE is the treatment effect, and the subscripts for the NI subtask contain identifiers for treatment versus control (T/C) and year (2012/2014). Treatment effects for all subtasks were calculated with the same formula. Ultimately, this treatment effect provides an unbiased estimate of the impact of the intervention for treatment students.

Figure 4. Illustration of the difference-in-differences approach to determining intervention impact



2.7 Reliability Estimates

To ensure that the assessment instruments were measuring their intended constructs, both the Early Grades Reading Assessment (EGRA) and Early Grades Mathematics Assessment (EGMA) were tested for reliability. Cronbach's alpha values for both EGRA and EGMA indicated that the instruments showed good internal consistency on average ($\alpha = 0.86$ for EGRA and 0.90 for EGMA for the baseline assessment; and $\alpha = 0.86$ for EGRA and 0.85 for EGMA for the endline assessment). Overall, these reliability measures provide evidence that the assessments were each measuring a single underlying construct: early grade reading ability for EGRA and early grade mathematics ability for EGMA. For the endline assessments, only word problems (EGMA) and listening comprehension (EGRA) had reliability estimates that fell below the conventional cutoff of 0.70, and the results on these subtasks should therefore be interpreted accordingly.

3 Findings

3.1 *Impact of the Intervention*

While there were virtually no gains in control schools from 2012 to 2014, there were significant gains across treatment schools in terms of reducing the proportion of the lowest performers and increasing the proportion of the highest performers.

Overall, the Early Grades Reading Assessment (EGRA) and Early Grades Mathematics Assessment (EGMA) results indicate that the intervention was successful in raising reading and mathematics achievement in treatment schools. The impact of the intervention on individual subtasks will be discussed in section 3.2. To gain an overall impression, composite reading and mathematics scores were created to determine the aggregate effect of the intervention. Specifically, oral reading fluency and reading comprehension scores were used to calculate an overall reading score, as defined by the following three categories:

1. Non-reader/Beginning reader: ORF greater than or equal to 0 cwpm[7] and reading comprehension equal to 0% (category 1)
2. Emergent reader: ORF greater than 0 cwpm and reading comprehension greater than 0 but less than 80% (category 2)
3. Reader: ORF greater than 0 cwpm and reading comprehension of at least 80% (category 3)

For the overall mathematics measure, missing number scores and the addition and subtraction level 2 scores were used to create the following three categories:

1. Non-mathematician/Early mathematician[8]: either missing number and/or addition and subtraction level 2 below 30% (category 1)
2. Emergent mathematician: missing number and addition and subtraction level 2 both above 30% (category 2)
3. Mathematician: missing number and addition and subtraction level 2 both above 80% (category 3)

These "reader" and "mathematician" categories are used extensively and remain consistent throughout the report.

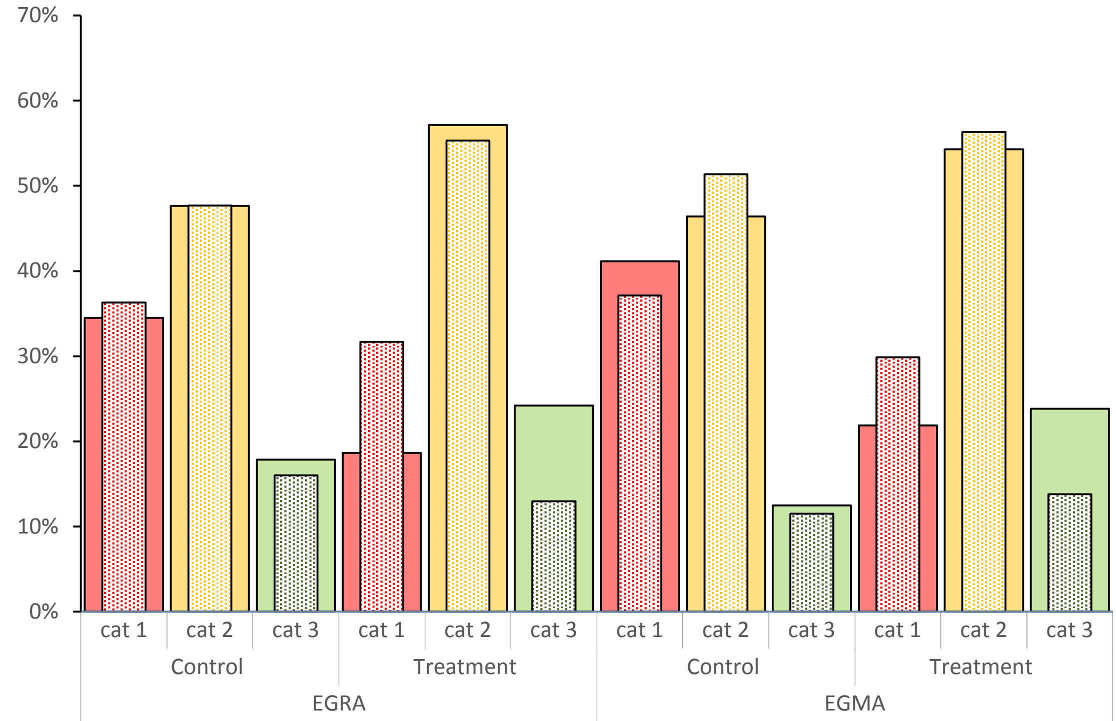
Figure 5 displays the changes in these categories from 2012 to 2014 both for treatment and for control schools. This figure provides direct evidence of the overall effectiveness of the intervention. While the percentage of non-readers or beginning readers and non-mathematicians or early mathematicians remains relatively consistent across years for the control group, there are large reductions in the proportion of non-readers or beginning readers and non-mathematicians or early mathematicians in treatment schools (from 32% to 19% in reading and 30% to 22% in mathematics).

---

[7] cwpm = correct words per minute.
[8] Mathematician is used here in the sense of doing mathematics at an age- and grade-appropriate level.

Additionally, while the proportion of readers and mathematicians remains constant for control schools, both proportions increase significantly in treatment schools (13% to 24% in reading and 14% to 24% in mathematics). In other words, the intervention did exactly what it was intended to do. While there were virtually no gains in control schools from 2012 to 2014, there were significant gains across treatment schools in terms of reducing the proportion of the lowest performers and increasing the proportion of the highest performers. These results are extremely promising, particularly because the intervention was implemented for only one school year. It should also be specifically noted that one of the districts in the South was among the top four performing intervention districts, both for reading and for mathematics, demonstrating that the intervention was also successful here.

Figure 5. Overall treatment effect for EGRA and EGMA categories



	EGRA						EGMA					
	Control			Treatment			Control			Treatment		
	cat 1	cat 2	cat 3	cat 1	cat 2	cat 3	cat 1	cat 2	cat 3	cat 1	cat 2	cat 3
Endline	34.5%	47.7%	17.9%	18.6%	57.2%	24.2%	41.1%	46.4%	12.5%	21.9%	54.3%	23.8%
Baseline	36.3%	47.7%	16.0%	31.7%	55.3%	13.0%	37.1%	51.4%	11.5%	29.9%	56.3%	13.8%

Key: narrow bar: baseline; wide bar: endline

3.2 Descriptive Analyses

Although it is useful to discuss overall effects, it is also important to examine the individual EGRA and EGMA subtasks to gain a more complete understanding of how consistently the intervention proved to be. The first step in this process is to determine the baseline (2012: pre-intervention) and endline (2014: post-intervention) scores for students both in treatment and in control schools, as these measures will be used to calculate the treatment effect (or impact) for each subtask. Accordingly, ***Table 7*** provides an overview of average scores across all EGRA subtasks for students in treatment and control schools for 2012 and for 2014. This process provides a comparative view of baseline scores and test score gains across subtasks for those receiving the intervention and those not receiving it. For example, the first row of invented words in ***Table 7*** shows that in 2012 students in control schools average about 5.9% correctly read invented words, while students in treatment schools averaged about 5.5% correctly read invented words. In 2014, however, students in control schools increased slightly to 7.9% correctly read, while the students in schools receiving the intervention more than doubled from their 2012 scores (up to 11.6%). Students began at approximately the same level in 2012, but the intervention appeared to have a large impact on increasing scores by 2014, as compared to control schools. This trend appears to hold true for the majority of subtasks listed in the table, but these effects will be tested directly in the section 3.3, using a difference-in-differences (DID) approach.

Although the majority of measures in ***Table 7*** are specific subtask scores, reading comprehension at 80% is a derived variable that shows the percentage of students who scored at least 80% correct on the reading comprehension section. Additionally, it should be noted that the syllable sounds and dictation subtasks were administered only in 2014, not in 2012.

Table 7. EGRA subtasks—percentage correct by year and control/treatment

EGRA Subtask	Year	Control	Treatment
Letter sound*	2012	54.4%	47.7%
	2014	64.9%	71.8%
Invented words	2012	5.9%	5.5%
	2014	7.9%	11.6%
ORF (correct words per minute)	2012	19.2 cwpm	20.0 cwpm
	2014	22.2 cwpm	28.0 cwpm
Oral reading	2012	34.0%	37.2%
	2014	31.7%	47.1%
Reading comprehension	2012	34.6%	33.6%
	2014	33.8%	45.0%
Reading comprehension 80%	2012	17.0%	13.8%
	2014	17.9%	24.2%
Listening comprehension	2012	42.4%	42.1%
	2014	67.2%	74.1%

EGRA Subtask	Year	Control	Treatment
Syllable sounds*	2014	63.8%	76.7%
Dictation	2014	40.0%	55.1%

*Signifies that the subtask score was calculated as percentage correct given attempted.

Table 8 displays the percentage of zero scores for each EGRA subtask in 2012 and 2014 for treatment and control schools. In this case, positive impacts come in the form of reductions. For example, while there is virtually no change in the percentage of zero scores for letter sounds in control schools, the proportion is nearly cut in half in treatment schools (decreasing from 28.4% in 2012 to 15.7% in 2014). In addition to calculating treatment effects for overall scores, it is also possible to calculate treatment effects for the reduction of zero scores. These effects will be explored in section 3.3.

Table 8. EGRA subtasks—zero scores by year and control/treatment

EGRA Subtask	Year	Control	Treatment
Letter sound*	2012	22.1%	28.4%
	2014	21.2%	15.7%
Invented words	2012	47.8%	45.3%
	2014	30.7%	16.3%
ORF (correct words per minute)	2012	22.2%	15.9%
	2014	9.1%	4.0%
Reading Comprehension	2012	32.2%	27.3%
	2014	34.1%	18.6%
Listening comprehension	2012	22.1%	28.4%
	2014	21.2%	15.7%
Syllable sounds*	2012	47.8%	45.3%
Dictation	2014	30.7%	16.3%

*Signifies that the subtask score was calculated as percentage correct given attempted.

Tables 9 and 10 provide the EGMA equivalents of the two EGRA tables above. For the majority of subtasks, the increase in scores for treatment schools here is greater than that of control schools (which is to be expected for a successful intervention).

Unlike the EGRA subtasks, however, the reduction in the proportion of zero scores for the EGMA subtasks does not appear to be universally larger for treatment schools. On the one hand, this difference may provide preliminary evidence of a possibility that, although the intervention may effectively increase EGMA scores, it may not be as effective at targeting low performing students. On the other hand, by contrast, the 2012 zero scores for many EGMA subtasks were already low, and thus large decreases are more difficult to achieve.

Table 11 provides the EGRA and EGMA scores by student gender and school type for each of the subtasks for the treatment schools. What is striking from this table is that girls outperform boys on all tasks. The difference is statistically significant on eight of the nine EGRA and three of the six EGMA measures. School type is also quite clearly a predictor of success with students in all-girls schools performing better than those in mixed schools and those in all-boys schools performing most poorly.

Table 9. EGMA subtasks—percentage correct by year and control/treatment

EGMA Subtask	Year	Control	Treatment
Number identification	2012	78.4%	82.4%
	2014	84.6%	92.6%
Quantity comparison	2012	69.9%	72.9%
	2014	78.9%	89.4%
Missing number	2012	49.0%	54.2%
	2014	58.3%	72.3%
Addition and subtraction L1	2012	31.3%	34.9%
	2014	52.9%	63.1%
Addition and subtraction L2	2012	39.2%	39.9%
	2014	43.5%	59.8%
Word problems	2012	40.6%	38.7%
	2014	57.6%	68.7%

Table 10. EGMA subtasks—zero scores by year and control/treatment

EGMA Subtask	Year	Control	Treatment
Number identification	2012	1.2%	0.0%
	2014	0.3%	0.0%
Quantity comparison	2012	3.8%	2.1%
	2014	2.0%	0.1%
Missing number	2012	5.1%	2.4%

EGMA Subtask	Year	Control	Treatment
	2014	3.1%	0.4%
Addition and Subtraction L1	2012	9.7%	4.4%
	2014	2.3%	0.4%
Addition and subtraction L2	2012	7.0%	5.4%
	2014	13.4%	5.4%
Word problems	2012	32.3%	31.1%
	2014	7.3%	3.0%

Table 11. EGRA and EGMA scores by student gender and school type

EGRA/EGMA (treatment scores)	Student (mean score)		School Type (mean score)		
	Male	Female	Boys	Girls	Mixed
Letter sound	40.0%	48.8% *	39.9%	53.1% ^{t}	43.4%
Invented words	8.8%	13.8% ***	8.2%	14.8% ***	11.4% *
ORF (words per minute)	24.4	30.8 ***	22.0	33.0 ***	27.8 **
Oral reading	39.1%	53.0% ***	34.1%	56.6% **	47.4% *
Reading comprehension	40.1%	49.7% **	32.7%	55.0% ***	45.5% **
Reading comprehension 80%	21.6%	26.3%	17.2%	31.2% *	23.6%
Listening comprehension	70.5%	76.9% *	63.4%	79.2% ***	75.2% ***
Syllable sound	29.8%	37.3% ***	26.5%	38.2% ***	34.5% **
Dictation	45.5%	62.6% ***	40.9%	66.5% **	54.9% *
Number identification	90.9%	93.9% *	88.0%	94.9% **	93.1% *
Quantity comparison	87.8%	90.6%	84.1%	93.5% **	89.3% ^{t}
Missing number	71.4%	73.1%	66.2%	75.3%	72.9%
Addition/subtraction L1	61.1%	64.7% ^{t}	57.9%	67.5% **	63.0% ^{t}
Addition/subtraction L2	54.1%	64.3% **	51.0%	70.4% ***	58.4%
Word problems	66.6%	70.3%	63.4%	74.5% *	68.0%

* ^{t} p<.10, p<.05, ** p<.01, *** p<.001.

3.3 Intervention Impact—Determining Treatment Effects by Means of Difference-in-Differences (DID)

As described in the methodology section, the most appropriate way to calculate treatment effects (or an intervention impact) in this study is by using a DID approach. Most simply, this approach determines how much better treatment school students performed than control students after the intervention, as compared to their relative starting points (i.e., baseline scores from 2012).

The results of the DID approach for EGRA subtasks are displayed in ***Table 12***. This table shows that the intervention had a significant impact on treatment school students for every EGRA measure. The smallest impact was for the invented word subtask, which showed a 4.1% increase as a result of the intervention. The largest effect was found for letter sounds, for which the intervention was determined to provide a nearly 14 point increase in the percentage of correct responses. (Note that the impacts for syllable sounds and dictation only show mean differences in 2014 and should therefore not be interpreted as impacts of the intervention.) The final column in ***Table 12*** provides the treatment effect for zero scores. Although the intervention had a significant impact on increasing scores across all subtasks, it was only effective in significantly reducing the proportion of zero scores for letter sounds and invented words.

Table 12. Intervention impact (treatment effects) for EGRA subtasks

EGRA Subtask	Impact (% Correct)	Impact (Zero Scores)
Letter sound	13.6% *	-11.9% *
Invented words	4.1% ***	-11.9% *
ORF (words per minute)	5.0 *	1.3
Oral reading	12.8% **	n/a
Reading comprehension	12.7% **	-10.6%
Reading comprehension 80%	9.5% *	n/a
Listening comprehension	7.1% *	-3.8%
Syllable sounds^	12.9% ***	-6.4% ***
Dictation^	15.3% ***	-12.4% ***

* p<.05; ** p<.01; *** p<.001
^ signifies 2014 data only. Value is the mean difference in place of intervention impact.

Figure 6 provides a visual representation of the intervention impact for the five major subtasks tested both in 2012 and 2014. The blue lines in this figure represent the change in scores for students in control schools. The orange lines depict the change in scores for students in treatment schools. The gray lines represent what would have happened to the treatment schools had they followed the same trajectory as the control schools, that is what would have been expected had there been no intervention. The impact of the intervention is indicated by the difference between the orange line and the gray line in 2014. These differences are exactly the same as the values that are presented in the Impact (% Correct) column in ***Table 12.***

Turning to the impact of the intervention on EGMA subtasks, results are presented in an identical format to those from the EGRA. Accordingly, the results of the DID approach for EGMA subtasks are displayed in ***Table 13***. Although there is no significant impact on number identification, scores on all five of the other subtasks were significantly increased by the intervention. The level 2 addition and subtraction L2 subtask realized the largest gains, with an increase of 15.5%. This is also the only EGMA subtask to have a significant reduction in the proportion of zero scores as a result of the intervention—reducing the zero scores by approximately 6%. It is important to note that when compared with EGRA and with the exception of the addition and subtraction L2 and word problems subtasks, the EGMA zero scores were already low at baseline.

Figure 6. Impact of intervention for EGRA subtasks—difference-in-differences

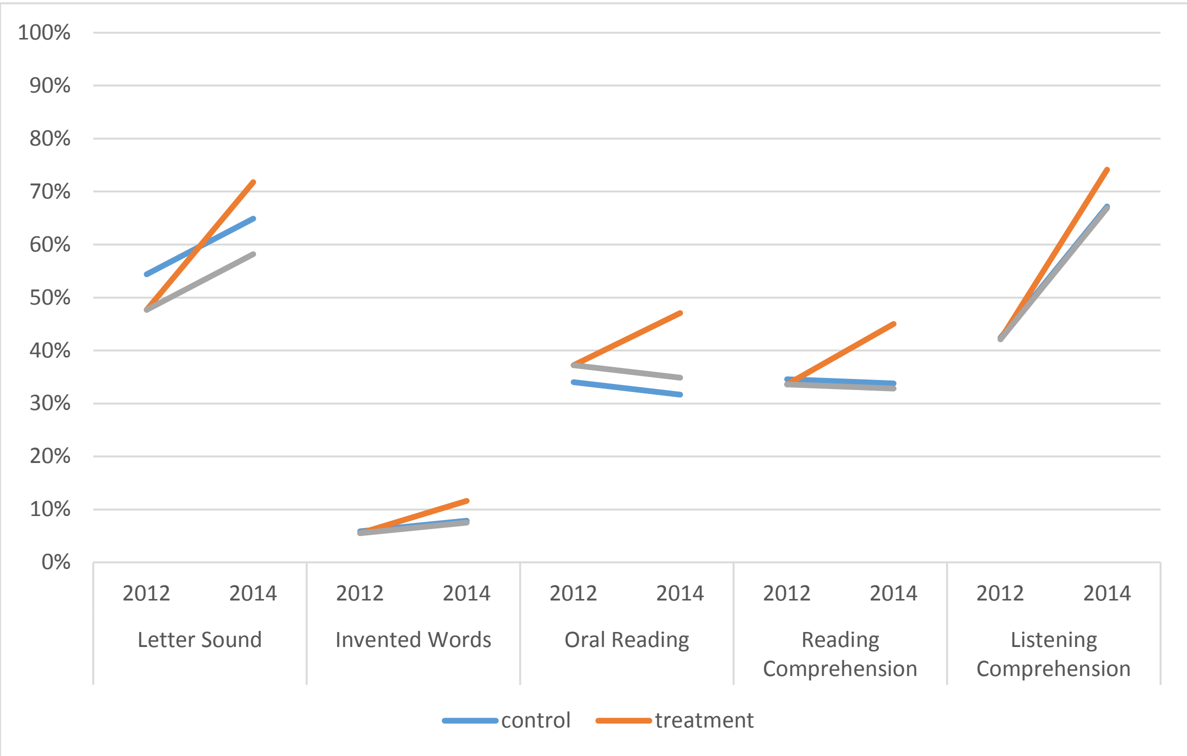


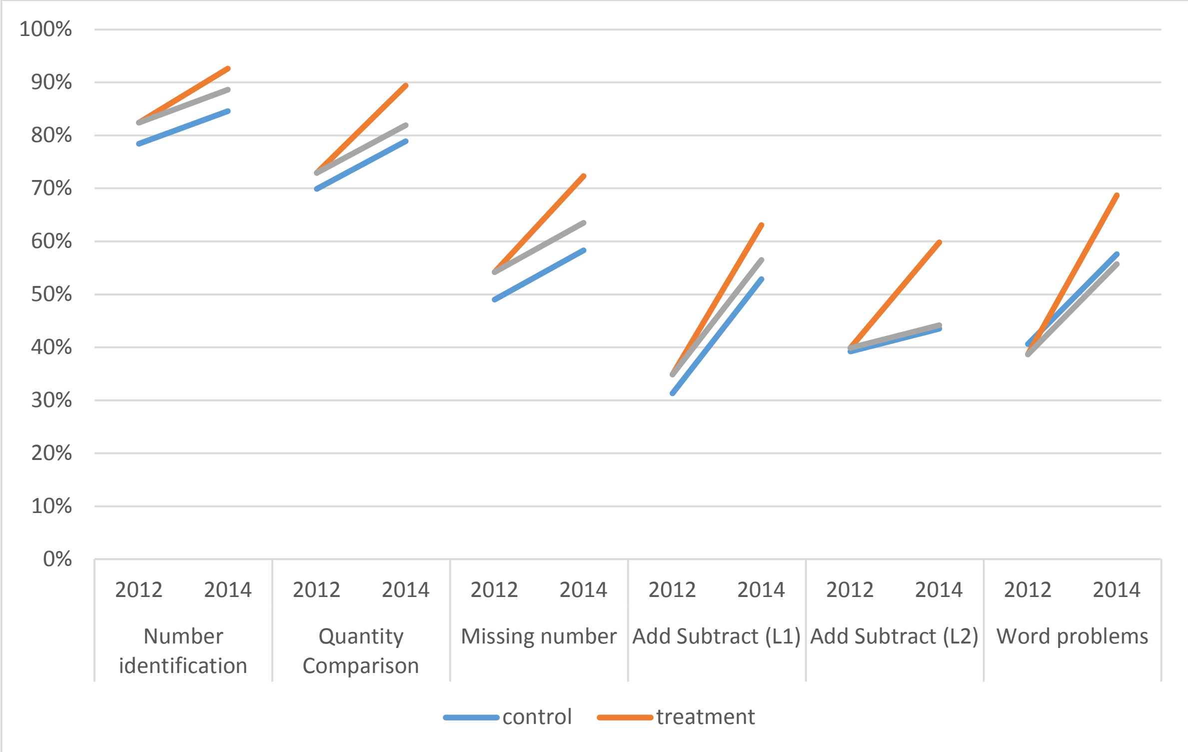
Table 13. Intervention impact (treatment effects) for EGMA subtasks

EGMA Subtask	Impact (% Correct)	Impact (Zero Scores)
Number identification	3.9%	0.9%
Quantity comparison	7.6% **	-0.2%
Missing number	9.3% **	0.0%
Addition/subtraction L1	6.7% **	3.5%
Addition/subtraction L2	15.5% ***	-5.9% *
Word problems	13.0% ***	-3.1%

* p<.05, ** p<.01, *** p<.001

The EGMA impacts are also represented visually, in ***Figure 7***. That scores appear to increase over time both for treatment and for control school students on a number of subtasks is a result of the timing criteria of these subtasks having changed from 2012 to 2014 (effectively allowing students more time to answer questions, thus increasing average scores). However, one important aspect of the difference-in-differences approach is that it is a calculation of how much better the treatment school students performed in 2014 (after taking into account any differences in scores in 2012). Therefore, the impact of the intervention is not reliant on consistent timing across assessments. Ultimately, ***Figure 7*** clearly shows that treatment and control school students often started with very similar scores in 2012 but that the intervention then had provided significantly larger gains for treatment school students by the time of the 2014 assessment.

Figure 7. Impact of intervention for EGMA subtasks—difference-in-differences



3.4 Intervention Impact by Gender

In general, because gender equity in education is an important topic, and particularly so in Jordan, it is useful to examine the impacts of this intervention on male and female students. Overall, girls outperformed boys in the sample, which raises the question if there were different effects of the intervention based on gender. According to the results presented in ***Table 14*** and ***Table 15***, the answer is yes. It can clearly be seen in ***Table 14*** that the intervention did not provide significant gains in achievement for male students on even a single EGRA subtask. Conversely, significant gains were seen by female students across every subtask. For zero scores, the intervention produced significant decreases for three subtasks for female students but provided no significant effect for male students.

Table 14. EGRA intervention impact (treatment effect) by gender—percentage correct and zero scores

EGRA Subtask	Impact (% Correct)		Impact (Zero Scores)	
	Male	Female	Male	Female
Letter sound	8.9%	17.1% **	-11.1%	-12.0% *
Invented words	1.7%	5.9% ***	-2.4%	-19.3% **
ORF (words per minute)	3.2	6.1 *	1.8	1.2
Oral reading	8.2%	14.9% **	n/a	n/a
Reading comprehension	8.9% *	15.4% ***	-7.6%	-12.4% *
Reading comprehension 80%	7.3%	11.0% *	n/a	n/a
Listening comprehension	5.2%	8.4% *	-2.9%	-4.4%

* p<.05, ** p<.01, *** p<.001

Similar results were found for EGMA subtasks. ***Table 15*** shows that the intervention produced significant gains in achievement for female students across all subtasks, while providing no significant increases for male students. Female students also saw a significant reduction in the proportion of zero scores for the level 2 addition and subtraction measure; the intervention did not significantly reduce the zero scores for male students on any subtask. Note that although it appears that female students had an increase in zero scores for number identification, this is actually just an artifact of the DID approach because of a reduction for control students, while treatment students were at 0% in both 2012 and 2014.

Table 15. EGMA intervention impact (treatment effect) by gender—percentage correct and zero scores

EGMA Subtask	Impact (% Correct)		Impact (Zero Scores)	
	Male	Female	Male	Female
Number identification	1.5%	5.8% *	0.7%	1.2% *
Quantity comparison	5.0%	9.7% **	-1.8%	1.1%
Missing number	6.0%	12.2% ***	-1.4%	1.4%
Addition/subtraction L1	3.5%	9.4% ***	2.5%	4.3%
Addition/subtraction L2	7.8%	21.7% ***	-4.2%	-7.3% *
Word problems	4.6%	20.2% ***	2.2%	-7.7%

* p<.05, ** p<.01, *** p<.001

Treatment effects were re-estimated for all subtasks using only mixed gender schools. The re-estimates were completed based on all-boys schools tending to be outperformed by mixed gender schools, thus it was important to disentangle the school gender from the student gender. However, the intervention remained ineffective for boys even within mixed gender schools, across all EGRA and EGMA subtasks. Ultimately, it is clear that this intervention was significantly more successful for female students on both the EGRA and EGMA, but it is not possible from these data to determine why this was the case. This gender difference is an issue that is worth exploring further in future studies.

3.5 Analysis of Key Factors and Their Influence on Results

3.5.1. Examination of Top Performing Classrooms and Districts

Before examining in detail the teacher, training, and supervisor characteristics that impact the proportion of readers and mathematicians (as defined earlier), it is important to examine the factors that are associated with the top performing classrooms and districts. For these analyses, top performing classrooms are defined as those with at least a 10% increase in readers (or mathematicians) in grade 2 or at least a 20% increase in grade 3 readers (or mathematicians). Top performing districts are defined as the four out of the 12 intervention districts with the largest increases in readers or mathematicians from 2012 to 2014 (that is, the top third of districts in terms of improved performance).

To determine the characteristics that are associated with top performing districts and classrooms, the mean of these characteristics, for the classrooms and districts by top performing status, were determined. The results of these analyses are displayed in ***Table 16***. The interpretation of the results involves simple means. For example, the first column under Top Performing Classrooms shows that 93% of teaches with frequent supervisor visits were in top performing classrooms (i.e., largest increase in readers), as compared to only 41% of those teachers in classrooms who were visited fewer than 16 times. For example, the teacher education variable shows that 91% of teachers with a diploma are in top performing classrooms (both for reading and for mathematics); however, significantly smaller proportions of bachelor's degree, higher diploma, and master's degree holders are in top performing classrooms.

Table 16. Characteristics associated with top performing classrooms and districts

Variable	Value label	n	Top Performing Classrooms: Reading	Top Performing Classrooms: Mathematics	Top Performing Districts: Reading	Top Performing Districts: Mathematics
Frequent visits by supervisors	Infrequent	62	41% *	-	17% *	-
	Frequent	**9**	**93% ***	-	**85% ***	-
Teachers attended both trainings	No	8	-	11% *	-	0% *
	Yes	63	-	**63% ***	-	**33% ***
School gender	All-boys (ref)	22	31% *	46% *	0% *	-
	All-girls	**18**	**65% ***	**89% ***	**38% t**	-
	Mixed	41	-	-	34% *	-
Student workbooks marked by teacher	None	3	-	-	-	0% *
	Less than half	7	-	24% *	-	0% *
	More than half	44	-	49% *	-	-
	All (ref)	17	-	**84% ***	-	**53% ***
Teacher monitoring of student understanding	No questions (ref)	10	-	34% *	-	0% *
	Asks questions	23	-	-	-	21% *
	Further explanation	38	-	**80% ***	-	**44% ***
Teacher education	Diploma (ref)	**9**	**91% ***	**91% ***	**60% ***	-
	Bachelor's degree	48	47% *	53% *	-	-
	Higher diploma	12	57% t	30% *	-	-
	Master's degree	4	33% *	33% *	5% *	-
Teacher experience	<5 years	16	58% *	-	-	-
	5–7 years	19	-	-	-	-
	8–14 years	18	33% t	-	-	-
	15 or more (ref)	**19**	**71% ***	-	-	-
Student participation in the lessons observed	Minimal (ref)	6	26% *	-	0%	0%
	Moderate	**30**	**54% t**	-	25% t	**43% ***
	Active	**35**	**53% ***	-	**42% ***	18% *
Teacher use of the lesson notes	Does not follow	1	-	-	-	-
	Follows notes	22	-	27% *	-	-
	Follows notes and routines (ref)	48	-	**69% ***	-	-

t p<.10, * p<.05, ** p<.01, *** p<.001

Although many characteristics tend to be associated with top performing classrooms and districts within subjects, the magnitude of the differences are universally smaller at the district level. This means that there is significantly more heterogeneity in districts (i.e., top performing teachers are spread across districts). However, it is important to note that there are several characteristics with a zero percentage of teachers in top performing districts. For example, of the teachers whose classrooms have minimal student participation, none of these teachers are in top performing districts either in reading or in mathematics. The same is true for poor monitoring of student understanding, lack of marks in student workbooks, and not attending both trainings, and how these issues are related to the largest increases in mathematicians and the all-boys schools for top improvement in readers. This latter finding bears repeating: while 38% of all-girls schools are located in top performing districts (for reading), no all-boys schools are found in these districts. Interestingly, there are no characteristics that are consistently associated with top performing classrooms and with top performing districts across both subjects.

3.5.2 The Role of the Supervisor

The assessment dataset includes supervisor visit observation data for 72 classrooms, over the 28 school weeks from September 2013 to April 2014. Based on these data, it is possible to explore how often teachers were visited and how the supervisors reported on the teachers for a variety of measures. It is furthermore interesting to explore the existence of relationships between these measures and the increases in readers and mathematicians.

Because all training and supervisor variables are available at the classroom level, these analyses use the proportion of readers and mathematicians in each classroom as dependent variables. Analyses were conducted separately for readers and mathematicians. Additionally, the proportion of students in each classroom in the reader and mathematician categories in 2012 was used as a control in each model. This was done to ensure that the findings are providing evidence of an impact on increases in readers and mathematicians, as opposed to simply signaling that some classrooms always tend to have higher achieving students. Thus, each coefficient in these regression models can be interpreted as the expected increase in readers or mathematicians given a one unit change in the independent variable (holding constant all else in the model, including baseline readers or mathematicians in 2012).

Before examining the training and supervisor measures directly, it was essential to determine which control variables were necessary to include in the final models. These are variables that impact the proportion of readers and mathematicians but are not necessary to explicitly report in the final models. Because the sample is limited to 72 classrooms, it is important to be judicious about the inclusion of controls. Initially, a range of potential classroom characteristics were tested by regressing the proportion of 2,014 readers (and separately 2,014 mathematicians) for grade, class size, student absenteeism, student workbooks, teacher gender, school gender, and wealth. The results of these preliminary analyses are displayed in ***Table 17***.

Table 17. Control variables for supervisor analyses

Variable	Readers	Mathematicians
Grade	27.79 ***	12.57 ***
Class size	-0.08	0.26
Student absenteeism	-19.88	-15.05
Student workbooks	0.10	0.11
Female teacher	-10.45 **	-8.43 **
All-boys school (*ref*)	-	-
All-girls school	22.56 *	20.84 ***
Mixed gender school	16.48 ***	9.92 **
Wealth	-6.41	-0.86

* p<.05, ** p<.01, *** p<.001

Regardless of the outcome (i.e., readers or mathematicians), it can be seen that only grade, teacher gender, and school gender are significantly predictive of changes in the proportion of top performers at the classroom level. For example, ***Table 17*** above shows that being in a grade 3 classroom is associated with an approximate 28 point increase in the percentage of readers, as compared to a grade 2 classroom (holding all else in the model constant). Although the negative coefficient for female teachers might initially lead one to believe that female teachers have a negative impact on reading and mathematics achievement, it is important to recognize that this is confounded with the school gender variable. Because all teachers in all-girls schools (and all but one teacher in mixed gender schools in this sample) are female, the following interpretation is more appropriate: All-girls schools and mixed gender schools are associated with higher proportions both of readers and of mathematicians than all-boys schools. Female teachers in all-boys schools have lower proportions of readers and mathematicians than their male counterparts in all-boys schools. Because there is such a large overlap between these two variables, it is only necessary to include one of them in the final model—and because school gender requires the estimation of an additional parameter, teacher gender was ultimately chosen as the most appropriate gender control. Additionally, grade was selected as a control due to its strong significance, and wealth was included in all models due to the overwhelming evidence of wealth effects across the education literature.

As with the models created for the determination of control variables, the impact of training and supervisor variables on the classroom proportion of readers and mathematicians were calculated separately. ***Table 18*** provides the regression coefficients (impact) for all variables that were significant in at least one of the models. The numbers in the impact column can be interpreted as the increase in the proportion of readers/mathematicians for a one-unit increase in each given training or supervisor variable. For example, the first variable in the table is the total number of supervisor visits for a given teacher. The 1.5% point impact on the proportion of readers comes from each additional supervisor visit. Likewise, attending both training sessions leads, on average, to a 9.4 percentage point increase in readers.

Table 18. Impact on readers and mathematicians—training and supervisor variables

Variable	Readers	Mathematicians
Total number of supervisor visits (teacher)	1.5% *	0.77% [t]
Frequent visits (16 or more)	27.2% *	14.6% *
Total number of supervisor visits (school)	n/a	0.14% *
First training—days attended	2.1% *	n/a
Second training—attended all	9.5% *	n/a
Attend both trainings	9.4% *	14.8% *
Total days of training	0.84% *	1.1% *
Teacher is on expected page of workbook during visit (proportion)	15.4% *	14.5% *
Following lesson notes	8.0% *	8.0% *
Monitoring student understanding	3.2% [t]	n/a

Controls: grade, teacher gender, wealth, 2012 scores
[t] p<.10, * p<.05, ** p<.01, *** p<.001

Table 18 shows that nine training and supervisor variables were found to significantly increase the proportion of readers within classrooms and seven that significantly increased the proportion of mathematicians. For both outcomes, the most important factor was frequent supervisor visits (which was calculated as the top 20% of the number of visits to a teacher). The frequency of supervisor visits was associated with a 27% increase in the percentage of readers and a 15% increase in the percentage of mathematicians. Additionally, the total number of supervisor visits per teacher was significant for both outcomes, but the total number of school-level visits was only significant for mathematicians. As for training, while all four training variables were significant for readers, only attending both trainings and increasing the number of total days attended resulted in an increase in the proportion of mathematicians. The bottom three rows of ***Table 18*** display the impact of supervisor ratings on readers and mathematicians. It is clear that being on the expected page of the teacher workbook more often and closely following the lesson notes are associated with gains in both outcomes.

Although the interpretations of the coefficients in ***Table 18*** are straightforward, it is important to understand that they are not directly comparable because of the different scales of the independent variables. In other words, just because the coefficient is larger, these results alone do not mean that attending both trainings is more important than supervisor visits. To determine the relative impact of each measure, it is necessary to standardize these effects (see ***Figure 8*** and ***Figure 9***).

Figure 8. Impact on readers—training and supervisor variables effect sizes

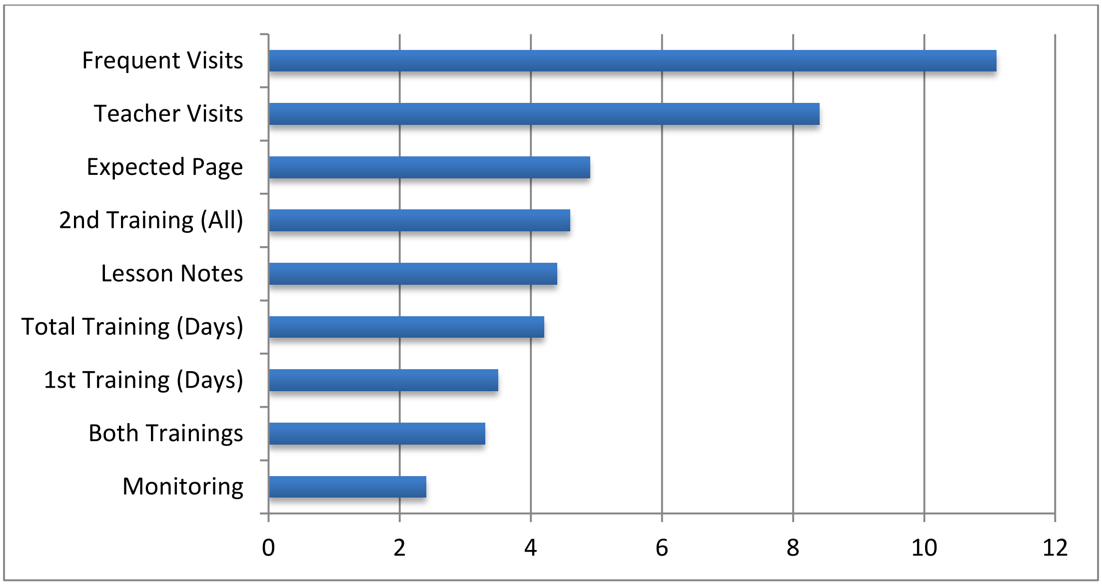
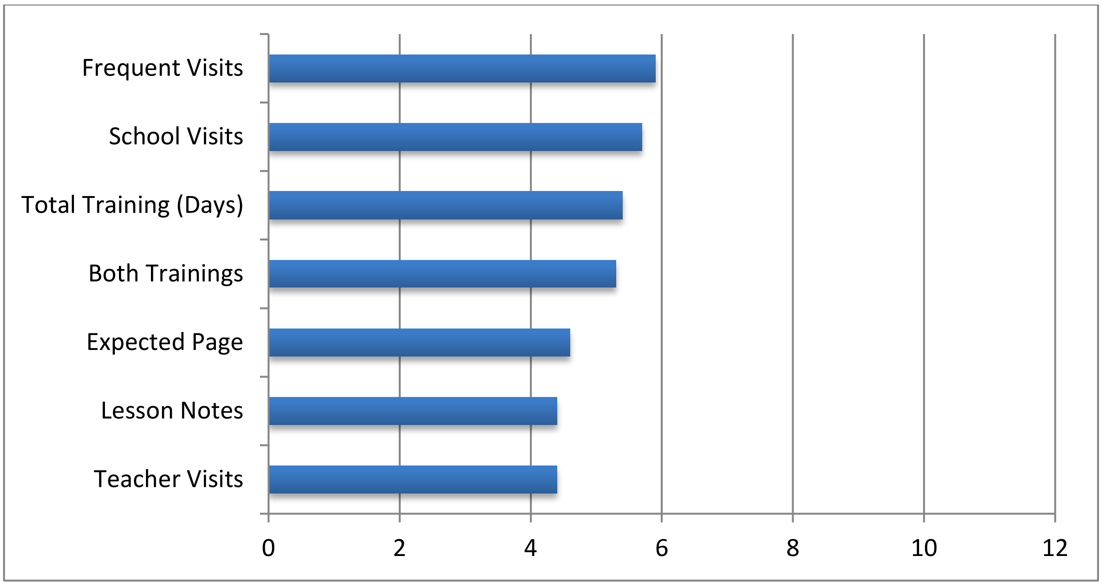


Figure 8 and ***Figure 9*** each provide standardized coefficients (or effect sizes), which are all scaled to be directly comparable. Therefore, while the specific interpretation of effect sizes is complicated, they provide a simple measure of the relative impact of each measure. For example, while the coefficient on supervisor visits to teachers was smaller than that of attending both trainings ***(Table 18)***, ***Figure 8*** show that teacher visits are actually almost twice as effective as attending both trainings, given the same relative change (i.e., one standard deviation) in each measure. These figures show a large drop-off in the impact of supervisor measures on the proportion of readers after frequent visits and teacher visits, but that the relative impacts are much more similar to one another for increases in mathematicians. Ultimately, the interpretation of the coefficients from ***Table 18*** is more straightforward, but the effect-size figures provide a measure of the relative impact for each variable. Therefore, it is important to look at both the table and the figures when interpreting these results.

Figure 9. Impact on mathematicians—training and supervisor variables effect sizes



3.5.3 The Role of the Teacher, Analysis of Teacher Data

Teacher questionnaires were completed in treatment schools only. There were two questionnaires: (1) an oral questionnaire that was administered only to the teachers of the classes from which the students were selected to complete the EGRA and EGMA assessments and (2) a written questionnaire that was completed by all of the grade 1, grade 2, and grade 3 teachers in the school. In all, oral questionnaires data was obtained for 72 teachers/classrooms, and written questionnaire responses were collected for 233 teachers.

The purpose of this section is to explore the teacher-level variables that explain some of the variation seen in the EGRA and EGMA outcomes in treatment schools, and specifically, the difference between pre- and post-treatment levels of student reading and mathematics achievement. The oral questionnaires focused on three aspects of teachers in treatment schools: (1) the teachers' background characteristics (e.g., experience, training), (2) their implementation of the reading and mathematics components (e.g., how often they followed the routines, whether they used the lesson notes), and (3) their perceptions of the intervention itself.

Teacher characteristics associated with student achievement (oral questionnaire)

The teacher data, obtained from the oral questionnaires in treatment schools, enable an analysis similar to that conducted with the training and supervisor data above. For the sake of continuity, a linear regression model akin to that described in ***Table 18*** was employed to discern the relationship, if any was present, between teacher-level variables and the proportion of "readers" and "mathematicians" in their classrooms in 2014, while controlling for grade, teacher gender, student wealth, and the proportion of readers and mathematicians in their classrooms during the baseline year (i.e., 2012). The regression coefficients are presented in ***Table 19*** for the variables that have not already been identified in ***Table 16*** and that were not discussed earlier.

Table 19. The predicted impact of teacher variables on readers and mathematicians

Variable	Value label	n	Readers	Mathematicians
Only teacher of this class	Yes (ref)	53		
	No	13	-5.53% **	-3.35% **
Permanent or substitute teacher	Permanent (ref)	63		
	Substitute	3	-10.79% **	-14.14% **
Attend 10-day training workshop	Yes (ref)	57		
	More than half	3	-20.52% **	-19.89%
	No	5	-9.77% *	-13.69% *
Attend 5-day training workshop	Yes (ref)	48		
	No	5	-8.48% t	-6.8%
	More than half	10	-8.7% t	-1.77%
	Less than half	3	-14.75% t	0.75%

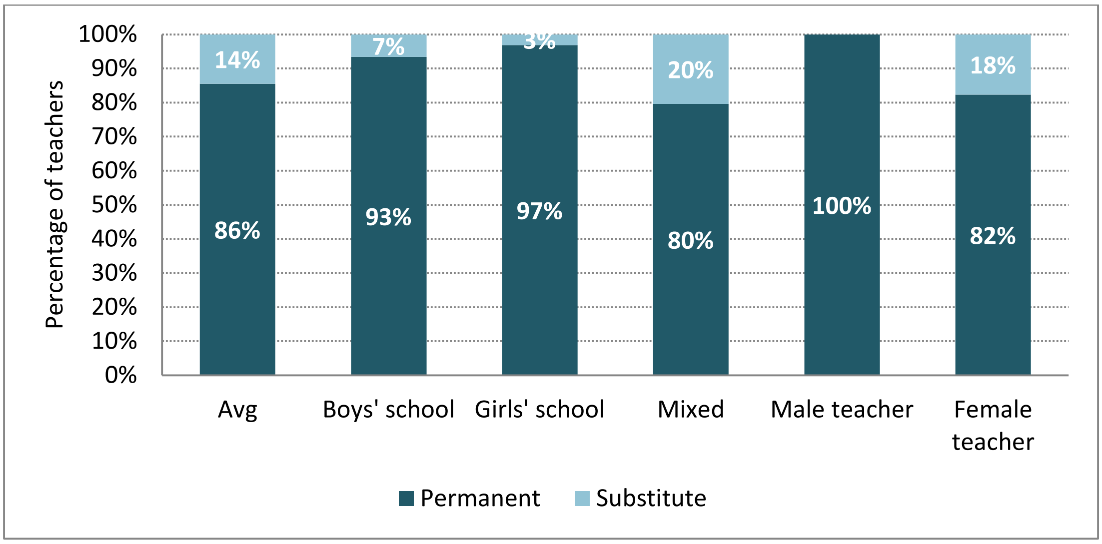
Controls: grade, teacher gender, wealth, 2012 readers and mathematicians
t p<.10, * p<.05, ** p<.01, *** p<.001

In addition to those teacher characteristics already reported in ***Table 16***, the analysis of the teacher questionnaire data found two additional teacher characteristics that are significantly associated with the proportion of readers or mathematicians in classrooms in 2014, as noted below.

First, having multiple teachers teach a single class over the course of a year was negatively associated with the proportion of readers and mathematicians, as compared to classrooms with a single teacher for each classroom. Specifically, classrooms with multiple teachers over the course of a year had fewer readers (5.5 percentage points fewer) and mathematicians (3.4 percentage points fewer) than classrooms for which a single teacher was responsible throughout the school year.

Second, being a substitute teacher was also found to be negatively associated with the proportion of both readers and mathematicians in the classroom. Substitute teachers tended to have fewer readers (10.8 percentage point difference) and fewer mathematicians (14.1 percentage point difference) in their classrooms than did permanent teachers. As seen in ***Figure 10***, the majority of treatment-school teachers, who completed the oral questionnaire were permanent (86% on average) rather than temporary teachers. All substitute teachers were female (18%), and substitutes tended to be more widely used in mixed schools (20%) as opposed to boys' or girls' schools.

Figure 10. Proportion of permanent and substitute teachers



Teacher participation in the 10-day and 5-day in-service workshops conducted in the implementation of the reading and mathematics project were both found to be predictive of the proportion of readers and mathematicians in teachers' classrooms. In the questionnaires, teachers were asked if they attended the 10-day and 5-day in-service training workshop for the reading and mathematics project. Specifically, teachers were asked if they attended all of the training, more than half, less than half, or none of the training. ***Figure 11*** and ***Figure 12*** depict the percentage of teachers that participated in the training workshops and the extent of their participation. In the regression model, attending less than the full duration of the 10-day workshop was negatively associated with the proportion of readers and mathematicians in a classroom. Teachers who attended more than five days, but less than the full 10, tended to have fewer readers (20.5 percentage points fewer) in their classrooms than teachers who attended all days of the training. Similarly, teachers who attended none of the training tended to have fewer readers (9.8 percentage points fewer) and mathematicians (13.7 percentage points fewer) than did teachers who completed the training. ***Figure 11*** shows that most teachers, who completed the oral questionnaire (82% on average), completed the entire 10-day training workshop. Participation rates were somewhat lower, however, in mixed schools (77% of teachers completed the 10-day workshop), and among female teachers (78% completed the workshop). All male teachers and approximately 90% of teachers in boys' and girls' schools completed the training, suggesting that, of the female teachers who did not attend any of the training, most of these taught in mixed schools. Partial participation was uncommon: only two teachers reported attending more than half of the training days, yet not completing the training.

Figure 11. Teacher participation in 10-day project training workshop

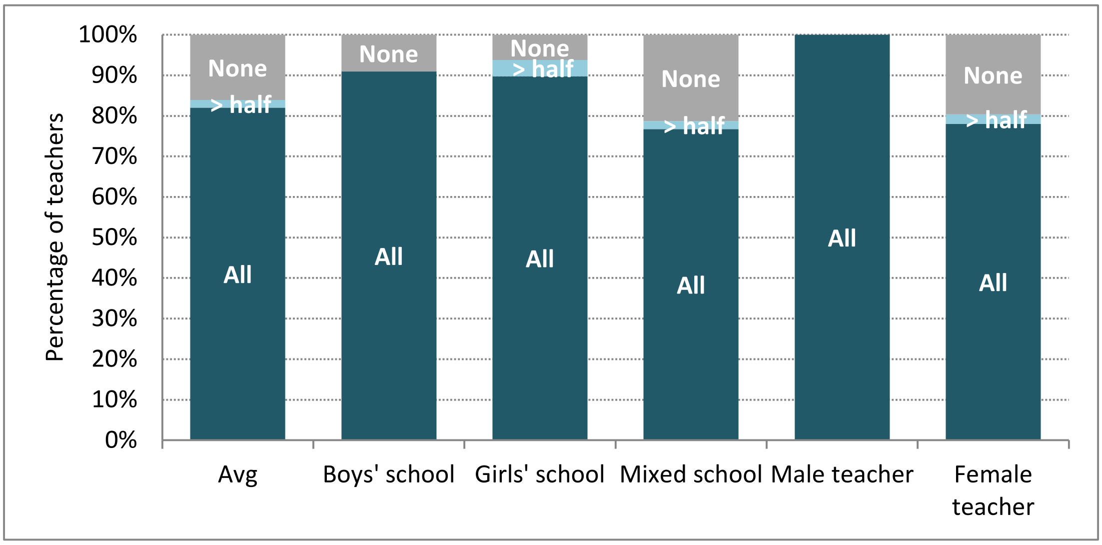
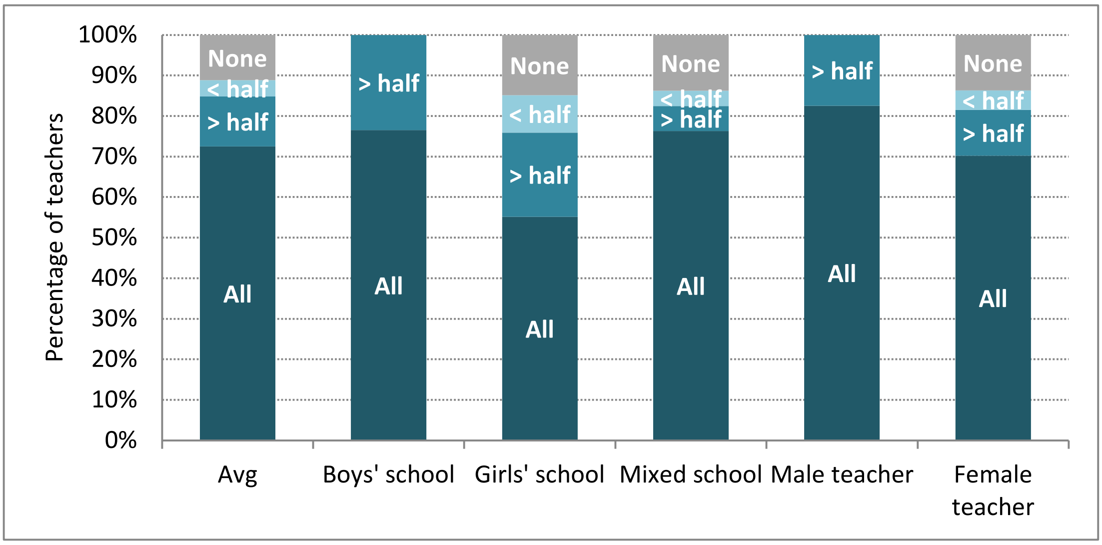


Figure 12 shows that participation rates for the 5-day training workshop were lower: on aggregate, 72% of teachers completed all of the training days while 12% completed more than half; 5% completed less than half; and 11% did not participate at all. Again, male teachers tended to participate in higher proportions than female teachers: nearly 20% of female teachers completed less than half of the training or none at all. These female teachers tended to teach in girls' and mixed schools, suggesting that access to training might not have been equal across school types. Access to project training could be further equalized going forward.

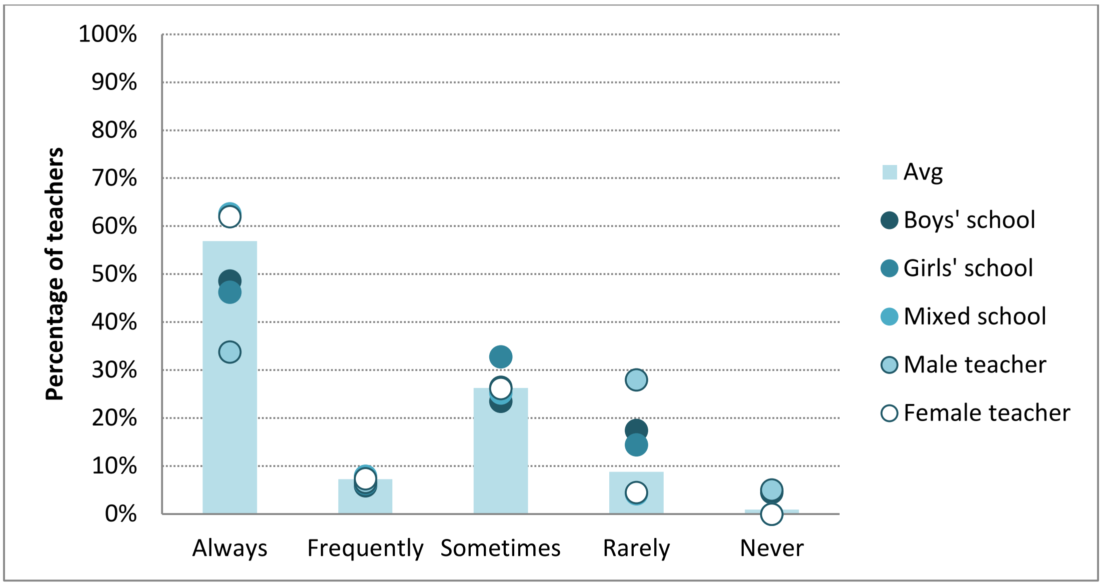
Figure 12. Teacher participation in 5-day project training workshop



Teachers' implementation of the project classroom routine (oral questionnaire)

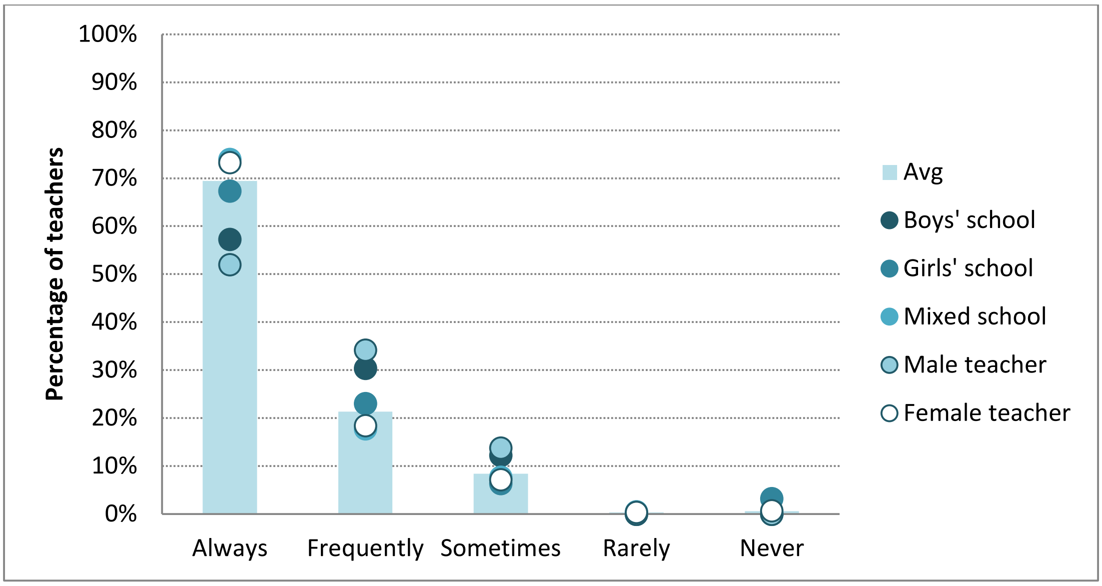
It has already been demonstrated in this report that supervisors' accounts of teachers' implementation of the reading and mathematics routines are associated with desirable program outcomes: teachers who were following the lesson notes, were on the expected workbook page, used monitoring techniques to check students' understanding, and attended both training sessions tended to have higher achievement in their classrooms. In the teacher questionnaire, teachers were also prompted to give a descriptive account of their implementation of the classroom routines: overall and for the reading and mathematics components**.** ***Figure 13*** shows the percentage of teachers who report that the reading and mathematics routines take longer than 15 minutes. Although the routine was designed to take 15 minutes, a vast majority of teachers report that it often takes longer: 90% of teachers overall indicated that implementing the classroom routine sometimes, frequently, or always takes more than the time allotted for it. This differs by teacher gender and school type: approximately two-thirds of female teachers and teachers in mixed schools (who are also female) reported that the routine always took more than 15 minutes, whereas only one-third of male teachers noted this. While these results, to an extent, might be expected given the relative newness of the classroom routines to these teachers and schools, these data do not indicated why the routine frequently took longer than was designed. These findings also foreshadow teachers' concerns that the activities were too numerous and that the time given to implement was too short. These are points that will be discussed below.

Figure 13. Teachers' response to: How often did the reading and mathematics routine take more than 15 minutes?



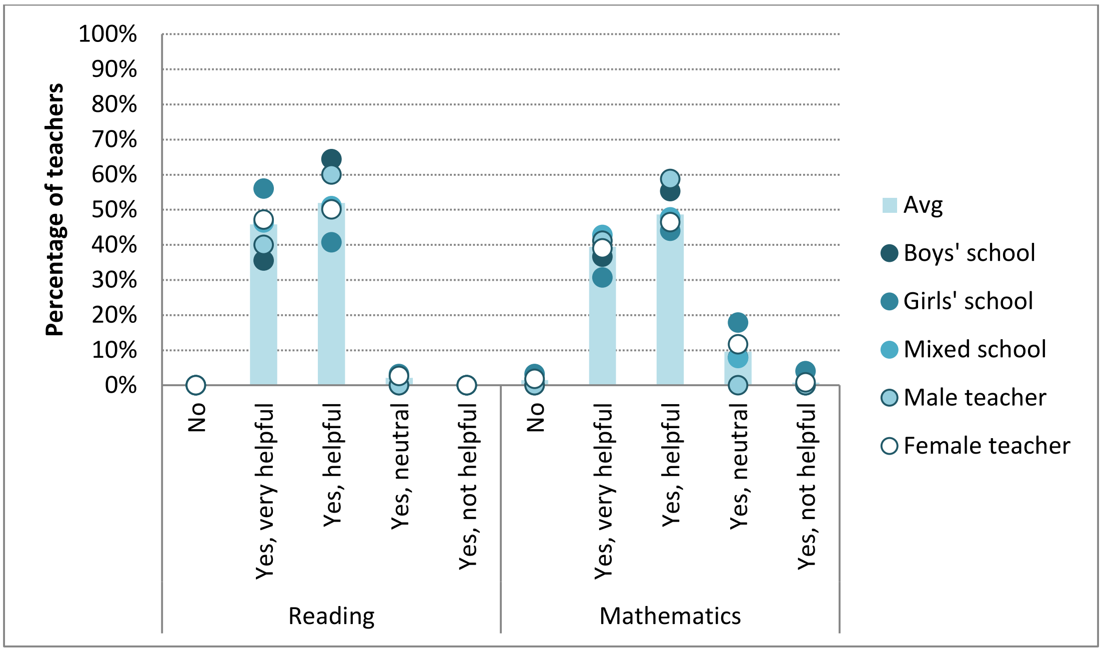
Teachers were also asked how frequently they followed exactly the lesson activities as specified by the teacher manual and lesson notes; their responses are displayed in ***Figure 14***. As seen in the figure, most teachers (90% on aggregate) followed the project activities exactly all of the time or frequently. Here again, there are some differences by teacher gender and school type. On the whole, female teachers and those in mixed schools were more likely to report that they always followed the project activities as specified in the manuals and notes (i.e., 73% and 74%, respectively), while teachers in boys' schools and male teachers were less likely to do so (57% and 52% of these teachers, respectively, always followed the routines as specified). On the one hand, although teachers in boys' schools and male teachers were still likely to follow the activities more often than not (at least according to their reports), these teachers were more likely to deviate from the lesson scripts at least some of the time. On the other hand, virtually none of the teachers reported that they rarely or never followed the scripts provided. However, it is worth pointing out that the same teachers who reported always following the project activities exactly (i.e., female teachers and those in girls' schools) are the same teachers who tend to report that the classroom routines always take longer than the time allotted for them.

Figure 14. Teachers' response to: How often did you follow the reading and mathematics project activities exactly as instructed in the lesson notes?



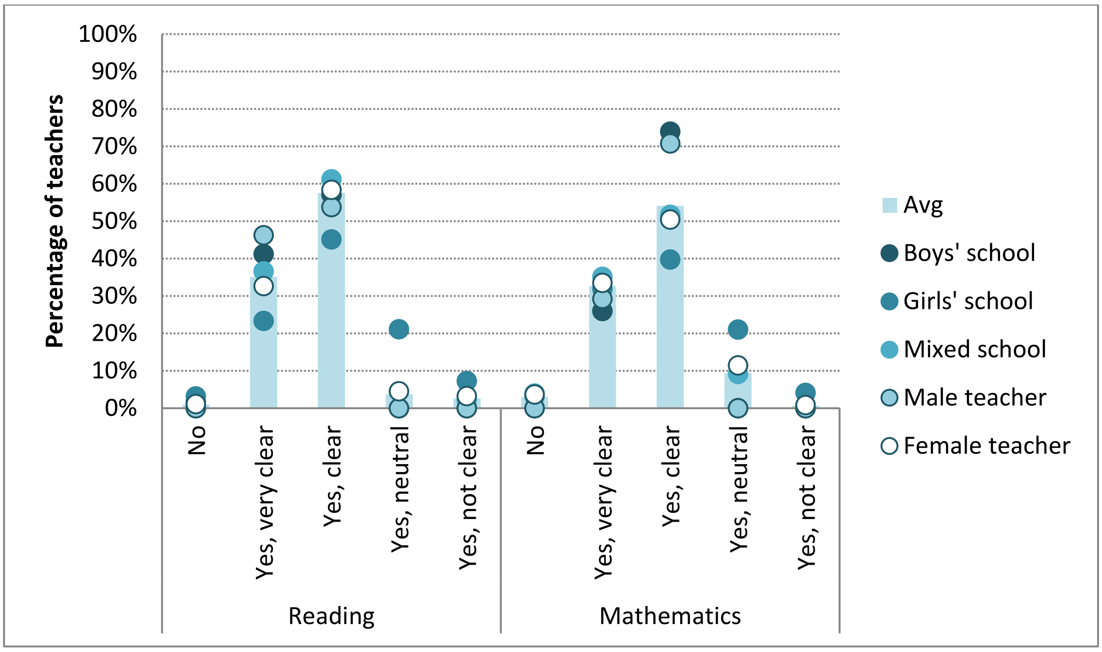
The following ***Figures 15–18*** attempt to illustrate in detail teachers' use and perceptions of specific resources and elements of the reading and mathematics project. ***Figure 15*** relates to the descriptions provided to teachers in the teacher's manual on how to conduct activities in the daily classroom routine. Teachers were asked whether they studied these notes and, if so, whether they found them useful. Evidence suggests that virtually all teachers used these notes to some extent and found them to be helpful or very helpful. It should be noted that teachers were slightly more ambivalent about these resources in the mathematics component as compared with the reading component; teachers (in particular, female teachers and those in girls' schools) were more likely to rate the teacher notes as unhelpful or neutral. Regardless, these resources appear on the whole to be well-used and well-received by the implementing teachers.

Figure 15. Teachers' response to: The teacher manual provided a description how to conduct the activities in the daily routine. Did you study these notes? If yes, were they very helpful, helpful, neutral, or not helpful?



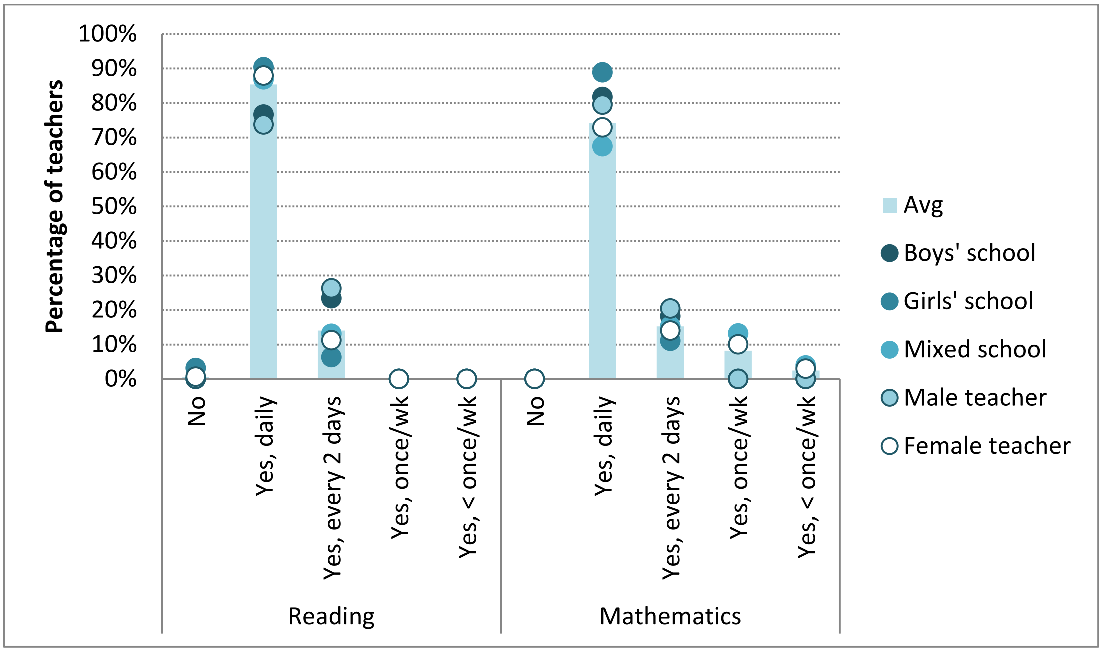
Teachers were also provided lesson notes that prescribed the lessons that should be taught during the classroom routines of the reading and mathematics components. ***Figure 16*** shows whether teachers used these resources and, if so, whether they perceived them to be clear and understandable. Responses indicate that teachers tended to study the lesson notes: only a very few teachers reported not using them. As with the teacher responses in ***Figure 15***, teachers that used the lesson notes tended to perceive them as clear or very clear (i.e., overall, 93% and 87% of teachers found the reading and mathematics lesson notes, respectively, to be clear or very clear). However, more variation also appears to exist across teacher gender and school types: while female teachers and teachers at girls' schools are equally likely to report using the lesson notes as are male teachers and those at boys' schools, the former (i.e., female teachers and girls' school teachers) are more likely to view these resources more critically.

Figure 16. Teachers' response to: The daily lesson notes described which activities you should do during the daily session with you class. Did you study these notes each day? If yes, were they very clear, clear, neutral, or not clear?



Student workbooks were supplied to all students and intended to be used on a daily basis when teachers go through the reading and mathematics routines with their classes. Teachers were asked whether their students worked in these workbooks and, if so, the frequency with which they did. ***Figure 17*** depicts the responses to this question. As shown, nearly all of the teachers reported that their classes worked in the student workbooks at least somewhat. Of these, the vast majority reported that students worked in the reading (99% of teachers) and mathematics workbooks (89% of teachers) daily or every other day.

Figure 17. Teachers' response to: The student workbook provided daily written activities for the students. Did your students work in these books? If yes, did they do so daily, at least every second day, once per week, or less than once per week?

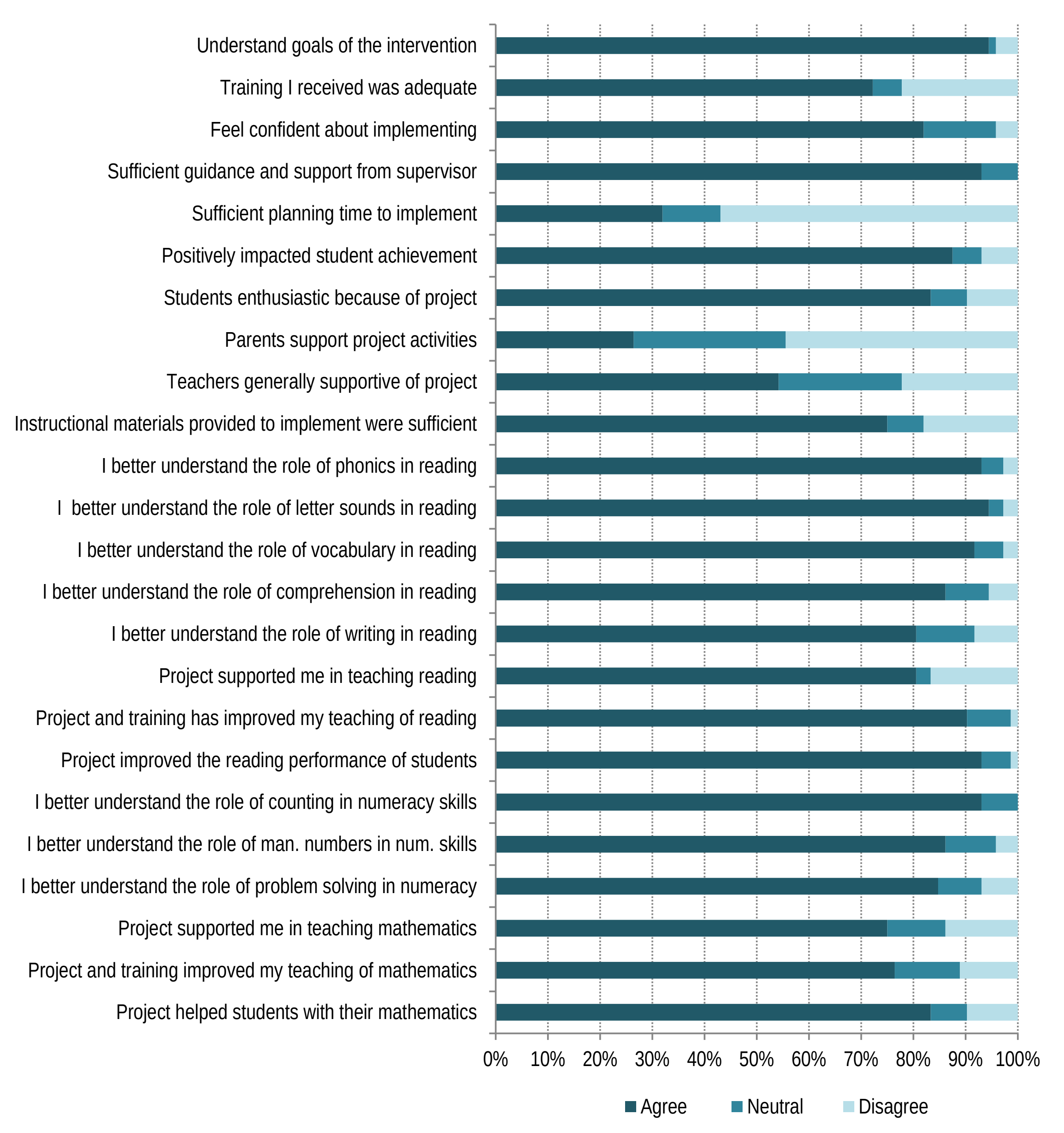


Overall, the picture portrayed in the above figures on teachers' implementation of the reading and mathematics components appears to be more or less favorable (most teachers reported using the provided resources and viewed them positively), the data also suggest that only a minority of teachers view the resources more critically. The questions about implementation behaviors by teachers in classrooms may indicate that some teachers are not carrying out the intervention with fidelity. For example, ***Figure 17*** showed that nearly one in every three teachers opted not to use the mathematics workbooks on a daily basis.

Teachers' perceptions of the reading and mathematics project in Jordan (oral questionnaire)

The teacher oral questionnaires also asked teachers to reflect on their experience in implementing the intervention project overall and the reading and mathematics components specifically. A number of questionnaire items consisted of statements about the intervention and asked teachers whether they agreed with the statements, disagreed with them, or neither agreed nor disagreed (neutral). ***Figure 18*** tabulates the results of these questionnaire items.

Figure 18. Teachers' overall perceptions of the reading and mathematics intervention



Note: man. = manipulating; num. = numeracy.

Most teachers agreed with the statements and, overall, few disagreed. This is an important finding—that the project was viewed positively—but it is equally important to highlight aspects of the project that were not held in such high regard by implementing teachers. Statements such as "Teachers were generally supportive of the project," "Parents support the project activities," "There was sufficient planning time to implement," and "The training I received was adequate" attracted some disagreement. Indeed, more than half of responding teachers felt there was inadequate planning time for the prescribed activities (only one in three felt there was sufficient time). Only a small proportion of teachers (approximately 23%) felt the training that they had received was inadequate. Despite otherwise general agreement that the project had positively impacted both reading and mathematics achievement, helped them with their teaching, and improved their understanding of various elements of literacy and numeracy, teachers tended to be more pessimistic when asked if parents and other teachers tended to support the project activities. Nearly 45% of teachers felt that parents did not support the project and 23% of teachers felt that other teachers did not support the project activities.

Table 20 summarizes data for the top 10 elements of the project that were viewed positively or negatively by teachers responding to the oral questionnaire. On balance, teacher respondents had more to say that was positive than negative in terms of the overall aspects of the project: a majority of teachers felt positively about six elements of the project, while a majority felt negatively about only two elements. On the positive side, teachers asserted that students enjoyed the project, it developed thinking skills and student skills generally, had a positive impact on learning, and led to an improvement in teachers' skills. Other positive aspects cited by nearly one-half of respondents were project training and the support given teachers by their supervisors. On the negative side, teachers pointed out that the project increased teachers' workloads and required too much time or effort to keep up with instructional demands (e.g., marking student workbooks). Other, less-cited, negative aspects were (1) the lack of encouragement, support, and incentives for teachers to undertake the work of implementation, and (2) the duration of training sessions.

Table 20. Overall positive and negative aspects of the reading and mathematics project—teachers' perceptions

Overall positive aspects of the project	Number of teachers (n = 73)	Overall negative aspects of the project	Number of teachers (n = 73)
Students enjoyed project	70	Increased teacher workload	64
Develops thinking skills	58	Too much time, effort to mark	54
Positive impact on learning	55	Insufficient encouragement, support, no reward	25
Activities support learning	53	Training (time, duration)	21
Improvement in student skills	52	Teacher turnover	18
Improvement in teaching skills	52	Training (content, presentation)	9
Training	47	Unclear objectives, materials	5
Supervisor support	46	No positive impact on learning	5
Encouragement to school/district	20	Lack of supervisor support	4
Parents enjoyed project	13	Supervisor creating confusion	1

Teachers were also asked to reflect on project aspects specific to the reading and mathematics components. ***Figure 19*** displays the reading and mathematics aspects of the project that teachers nominated as positive. Intervention materials (e.g., teacher notes, lesson notes), the activities' support of the formal curriculum, and students' enjoyment of the project activities were all cited as positive aspects both of the reading and the mathematics components. It is notable, though, that responding teachers tended to feel somewhat less positive about these aspects in the context of the mathematics components as compared with the reading components.

Figure 19. Positive aspects of the program specific to reading and mathematics components—Teachers' perceptions

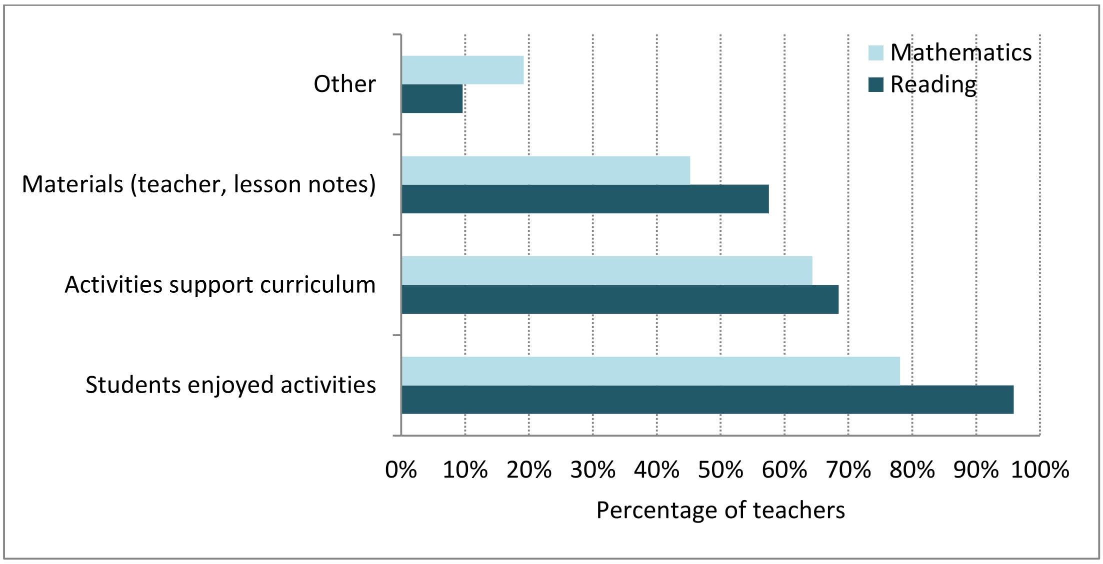
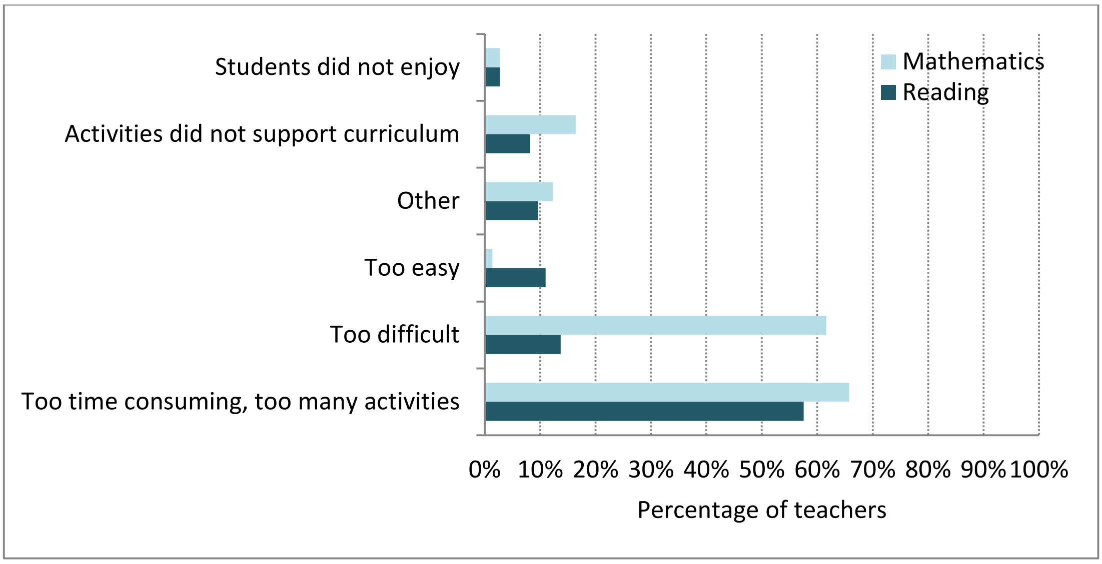


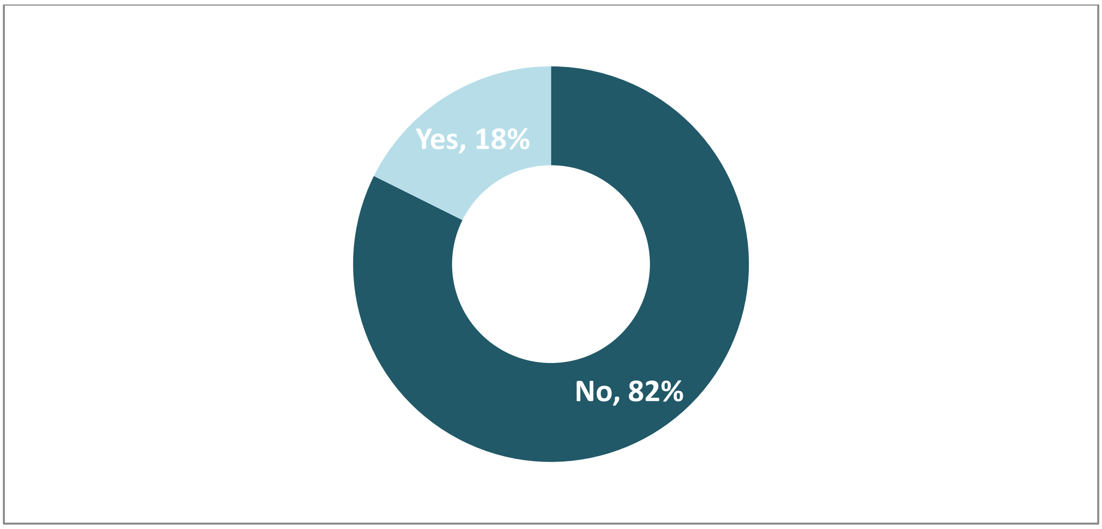
Figure 20 presents the aspects of the reading and mathematics components of the project that were viewed negatively. Although more aspects than those shown were nominated by respondents as negative, most of those views were held by 10% or less of responding teachers. However, two aspects of the reading and mathematics components stand out as more widely held negative views. More than half of responding teachers felt that the mathematics component was too difficult and that both components (reading and mathematics) were too time-consuming. It is interesting, in particular, to note the perceived degree of difficulty associated with the mathematics component as compared to reading; the mathematics component was much more widely perceived by teachers to be too difficult. The concern about the amount of time it takes to implement, and the number of activities required, reflect earlier findings in this report about the frequency with which the routine takes longer than the allotted time in the classroom and about the insufficient planning time available for the intervention. In summary, it appears that time, both in and outside the classroom, is at a premium for teachers in treatment schools, and this intervention may be viewed as adding to the burden of teachers' workloads rather than alleviating their load.

Figure 20. Negative aspects of the program specific to reading and mathematics components—Teachers' perceptions



Perhaps the most direct question posed in the oral teacher questionnaire was whether teachers felt the intervention should be or should not be continued in their schools. ***Figure 21*** shows the responses of teachers in treatment schools, which is overwhelmingly not in favor of continuing, by a margin of almost five to one. Unfortunately, this item did not ask teachers to specify precisely why they responded as they did. It is interesting, however, to juxtapose this response with those of other perception-oriented questions, the results of which were mostly positive. That is, despite mostly widespread agreement by teachers that the intervention led to enhanced student skills in reading and mathematics (and the survey results provide evidence of this effect), to improved teaching, and to being enjoyable for students, the majority of teachers would choose to not continue with the project. It is not completely clear why a majority of teachers made this judgment; however, the teachers' experience of the project as an add-on and hence an additional burden, as reported in discussions and during training, may provide some explanation. If this is the case, these findings should encourage intervention designers to take into account teachers' perceptions of the project and its role, for these perceptions are intrinsically related to stakeholder ownership, implementation fidelity, and ultimately, project sustainability.

Figure 21. Teachers' response to: Do you think your school should continue the approach of the reading and mathematics project?



Teachers' response to the reading and mathematics project in Jordan (written questionnaire)

In addition to the structured oral questionnaire that was completed by all teachers of those treatment students who had been assessed in the endline survey, all grade 1, grade 2, and grade 3 teachers in treatment schools also completed a written questionnaire. This written questionnaire contained mostly open-ended questions that invited freely phrased responses. A total of 233 teachers completed the questionnaire, and the teacher responses were coded and summarized to identify common themes and trends among the responses. Unlike for the oral questionnaires, linking the teachers' responses for the written questionnaires to the proportion of readers and mathematicians in their classes is not possible. The written responses, however, provide more qualitative rather than quantitative data; they nonetheless provide some very clear patterns that should be taken into account when reflecting on this intervention.

In the written questionnaires, teachers were asked to discuss both the reading and the mathematics components of the intervention in terms of the materials, the implementation, their perception of student response to the project, and the role of the supervisors in training and supporting the teachers as they implemented the project. For each of these (materials, implementation, student response, and supervisor role) teachers were asked to list what they had experienced as positive and as negative and to make recommendations for possible changes.

Summary graphs of the main responses to the written questionnaires and their frequencies are provided in **Annex A.** These responses confirm the responses that teachers expressed to the project staff during visits to the training venues in February 2014 and that were communicated by supervisors to the project staff during reflection sessions arranged between the project staff, the supervisors, and MoE personnel on several occasions over the life of the project. The key responses can be summarized as follows:

- Positive features of the intervention:
  - The intervention helped students to improve their reading and mathematics skills.
  - Students enjoyed the activities, experiencing them as meaningful and engaging.
  - The project addressed the identified weaknesses and supported students in developing basic/foundational skills.
  - The intervention exposed teachers to new and effective teaching approaches.
  - The supervisors played a critical role in training and supporting teachers.
- Negative features of the intervention:
  - The intervention added to teachers' workload. Teachers were not happy about this, and more than a few of them asked why they had been "burdened" with this additional load.
  - The mathematics materials were experienced as difficult for students—especially the so-called weaker students. This level of difficulty was not the case with the reading materials—especially not in the first semester.
  - The intervention did not provide sufficiently for the full developmental range of all the students in classes, resulting in some students being bored or frustrated (as was the case for stronger students in reading and for weaker students in mathematics).
  - Teachers could not always see the link between the intervention and their curriculum. In particular, they felt frustrated because the activities in the daily routine were not linked directly to the activities for that day in their curriculum (that is, their textbook).
  - Teachers felt that they did not have the time to complete both the curriculum and the daily routines of the intervention. In particular, many felt that they could not complete the daily intervention activities in the 15 minutes that were allocated to them.
  - Supervisors were unable to model the classroom activities when conducting training and/or were unable to demonstrate them in classrooms during support visits.
- Recommendations for revisions to be made to the intervention:
  - Reduce the load of the activities—especially the time required and the number of activities to be completed.
  - Make the link between the intervention activities and the curriculum (textbook) more explicit.

Despite the teachers raising concerns and making recommendations (as was asked of them), the written questionnaire responses gave the overriding impression that the teachers recognized the value of the intervention and thought more positively about it than negatively.

Those features identified as being negative in teachers' responses to the written questionnaire mainly reflect misunderstandings and misconceptions about the project and its role. For example, the intervention was not intended to take place in parallel to the curriculum, but instead, it was designed expressly to provide daily practice in foundational skills that the 2012 National Survey had identified as being poorly developed. The fact that teachers developed a misunderstanding reflects the way that the training and general advocacy around the intervention did not adequately anticipate and hence address how teachers would experience the intervention. Although the generally positive feedback from the teachers is encouraging and exciting, the negative aspects that were identified, the concerns raised and the suggestions made for improvement, clearly showed aspects that should receive attention in future training and advocacy around such a project. In particular, the concerns and suggestions also reflect something about the ownership of the project. Unless the activity is seen and experienced by teachers as an activity of the MoE, the chances of adoption and success will be compromised from the outset.

In considering the teachers' response on the oral questionnaire about whether or not they felt that their schools should continue with the project, it is clear that as long as an intervention of this kind is experienced by teachers as an add-on to the formal curriculum, teachers will not enthusiastically adopt the program, irrespective of the benefits that they may see in it.

4 Lessons Learned

The intervention has demonstrated unequivocally that it is possible to increase the number of readers and mathematicians in early grade classrooms by providing deliberate, structured, and developmentally appropriate practice in foundational skills for reading and mathematics. While there were virtually no gains in control schools from 2012 to 2014, there were significant gains across treatment schools in reducing the proportion of the lowest performers and increasing the proportion of the highest performers. These results are extremely promising, particularly because the intervention was implemented for only one school year.

This report section provides a summary of the variables that impacted the intervention, including those that increased the effectiveness of the intervention and those that detracted from effective implementation. Furthermore, these variables are analyzed to gain the lessons that they provide for future or similar implementations of this intervention in Jordan.

At the overall or macro level, the results for students in treatment schools show an increase in the proportion of high performers and a decrease in the proportion of low performers for reading and mathematics. At a more detailed level, the results show that the intervention had a significant impact on students' performance for almost all subtasks and that it reduced the percentage of students unable to correctly respond to items in each of the subtasks (zero scores).

This intervention set out to research whether daily practice of foundational skills for reading and mathematics could increase the number of students reading with comprehension and doing mathematics with understanding. From the wide range of evidence collected, it would appear as if, in general, the intervention was implemented with greater fidelity than not and that it had the desired impact. The implication may well be that there is much to be gained by an intervention that systematically addresses only those key elements of a teaching and learning program that has been shown to be deficient, instead of replacing the entire program.

An expectation that was central to the vision of this intervention was that if teachers were introduced to more effective pedagogies through immersion, that is, by asking teachers to implement a limited number of carefully structured routines on a regular (daily) basis, teachers would recognize the benefits of the approach and more generally assimilate some of that approach into their teaching. At this stage of the intervention, it is not possible to know to what extent teachers have actually incorporated the intervention practices more generally into their teaching (although some claim that they have). Nevertheless, it is clear from teachers' responses that they claim to have seen benefits from the intervention. Teachers claim that students enjoyed the intervention activities and that students benefited from the intervention because they appeared to perform better in reading and mathematics as a result of the intervention activities. Teachers also claim that the intervention exposed them to new and more effective teaching approaches.

Encouraged by the positive results, it is nonetheless critical to examine the different components of the intervention to see what lessons can be learned—lessons that will inform future interventions and improve their chances of success. The remainder of this section explores some of the key issues that the study has raised.

4.1 Gender

The issue of gender is complex, and it is beyond the scope of this study to explain some of the findings. However, the findings are so striking that they need to be highlighted and reflected on. The key findings about gender are as follows:

- Girls outperform boys on all EGRA and EGMA subtasks. The difference is statistically significant for eight of the nine EGRA subtasks and three of the six EGMA subtasks.
- Students in all-girls schools perform better than students in mixed schools who, in turn, perform better than students in all-boys schools. These differences are statistically significant on almost all EGRA and EGMA subtasks.
- Although the sub-sample sizes in the data do not allow for rigorous analysis of whether or not there is a statistical relationship between the gender of the teacher and the gender of the student in terms of impact on performance, there is enough evidence to suggest that, in general, the students of female teachers perform better than the students of male teachers.
- The treatment effects for the two genders suggest that while the intervention has had a strong impact on the performance of girls, boys in general appear to have enjoyed little or no benefit from the intervention.

- Among all-girls schools, 65% of the classes were among the top performing classes[9] for reading and 89% of the classes were among the top performing classes for mathematics. By contrast, only 31% and 46% of the classes from all-boys schools were among the top performing classrooms for reading and mathematics. Strikingly, there is not a single class from an all-boys school among the classes in the top performing districts[10] for reading.

This intervention study did not set out to investigate the role of gender, neither the role of student gender, nor the role of teacher gender; however, the results clearly indicate that gender is an issue in Jordanian schools and that further analysis is needed. In terms of lessons learned, it is very clear that future intervention projects will need to better understand the gender dynamics of Jordanian schools and to make conscious design decisions to ensure that boys benefit as much from the intervention as girls do.

4.2 Classroom Support

Key to the intervention implementation design was the role of the supervisors in providing school-based support to teachers implementing the intervention. The same supervisors who trained the teachers also visited the teachers in their classrooms to observe the implementation of the intervention by the teacher and provide advice and support. It is clear from the data that visits by the supervisors contributed to the impact of the intervention on the proportion of readers and mathematicians in a teacher's class. In particular, each additional supervisor visit is associated with an increase of 1.5% in the proportion of readers and 0.8% in the proportion of mathematicians in the class.

In the design of the intervention, it was hoped that supervisors would visit the teachers in their classrooms at least once every second week. For the data that is available over the period of the implementation, such a frequency of visits would have amounted to approximately 16 visits per teacher. Nearly 10% of the teachers were visited 16 or more times over the period of the intervention. Being visited as frequently as 16 or more times over the period of the intervention was associated with a 27% increase in the percentage of readers and a 15% increase in the percentage of mathematicians in the classrooms of those teachers.

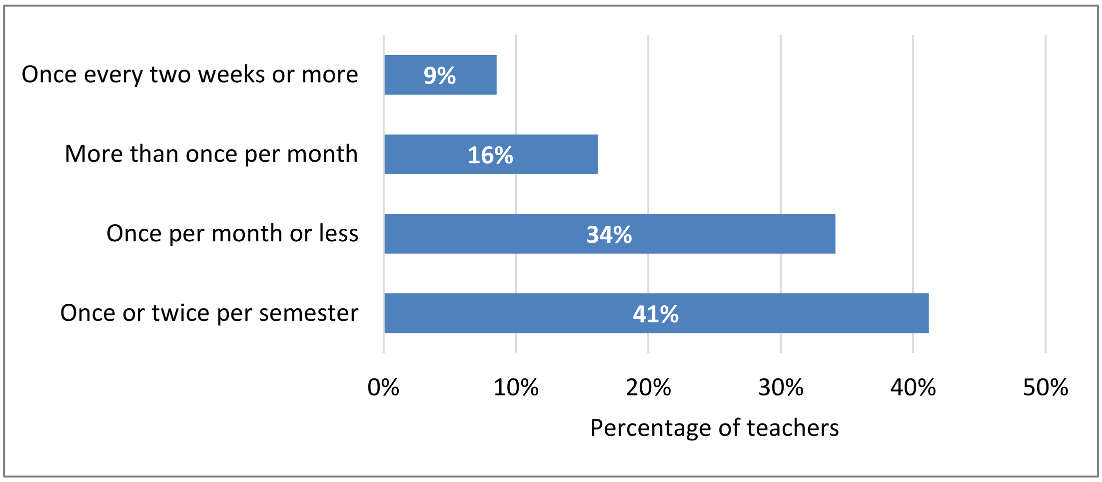
The lesson to be learned is that school- and classroom-based support to teachers, who are implementing an intervention of this type, enhances the successful implementation of the intervention. The more frequent the support is, the more effective the implementation will be. In the case of this intervention, the question remains why some supervisors visited classrooms more often than others, and what factors motivated them to do so. ***Figure 22*** summarizes the proportion of teachers by

[9] Top performing classrooms are defined as those with at least a 10% increase in readers (or mathematicians) in grade 2 or at least a 20% increase in grade 3 readers (or mathematicians).
[10] Top performing districts are defined as the four out of the 12 intervention districts with the largest increases in readers or mathematicians from 2012 to 2014 (that is, the top third of districts in terms of improved performance).

frequency of supervisor visits. It is clear that the majority of teachers were visited no more than once per month.

It is, of course, not enough for supervisors to simply visit teachers. What happens during and after the visit is as important as the visit. The aspects of the implementation, which the supervisors monitored and responded to, are discussed in section 4.4.

Figure 22. Frequency of classroom-based supervisor support visits



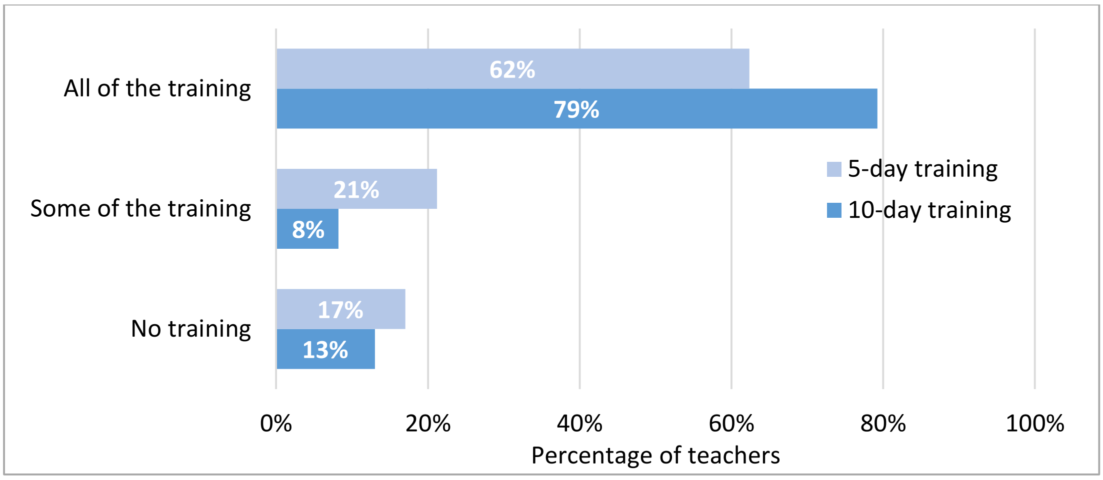
4.3 Teacher Training

Another variable that had a significant impact on the success of the implementation was the proportion of the training that the teachers had attended. Teachers who attended more of the training had a greater proportion of readers and mathematicians in their classes than teachers who attended less training. It is interesting that this is more strongly the case for mathematics than it is for reading. In particular, none of the classrooms associated with teachers who did not attend both trainings were in the top districts for mathematics.

Although teachers attending the training had a significant impact, both on the proportion of readers and proportion of mathematicians in the teacher's class, the impact was greater for mathematics. There are a number of possible explanations for this. First, the mathematics materials were, from the outset, more demanding than the reading materials. In contrast, the reading materials increased in cognitive demand only during the second semester (see discussion in section 1.2.2). Second, it can be assumed that teachers in the early grades are themselves more confident readers than mathematicians: teachers stated this in the focus group discussions conducted during the training visits. Finally, the pedagogy of the mathematics intervention represented a dramatic shift: from the predominant classroom pedagogy where mathematics is seen as *the memorization of facts, rules, formulas, and procedures needed to determine the answers to questions* to a pedagogy where mathematics is experienced as a *meaningful, sense-making, problem-solving activity*.

Although training attendance, in general, was good, considering the logistical challenges faced in arranging the trainings (see discussion in section 1.4), it should nonetheless be noted that more than 20% of the teachers did not attend all of the initial 10-day training and 38% of the teachers did not attend all of the follow-up 5-day training at the start of the second semester.

Figure 23. Teacher attendance at training



The lesson to be learned is that if training is shown to have an impact on the extent of the intervention, then all possible efforts must be made to ensure teacher attendance at the training.

4.4 Fidelity of Implementation

The feedback provided by supervisors about their classroom visits gives a range of different ways of evaluating the fidelity with which teachers implemented the intervention. In particular, supervisors reported about (1) the particular lesson (in the lesson notes) that the teacher was implementing; (2) the extent to which teachers were following the lesson notes as they should have been; (3) whether or not the teacher was actively monitoring student understanding during the lesson; (4) the type of student participation in the lesson; and (5) the extent to which students had worked in their workbooks and teachers had marked the workbooks. In the analysis of the data, all of these variables were positively associated with the intervention's impact.

Being on the expected page of the lesson notes was associated with a 15% increase in the percentage of readers and mathematicians in a classroom. In fact, 70% of the mathematics classrooms in which teachers followed the teacher guide and lesson notes were among the top performing classrooms for mathematics. Classrooms, where teachers encouraged student participation, were more likely to be among the top performing classrooms, both for reading and for mathematics; for classrooms where students were not actively encouraged to participate in the lessons, not a single classroom was in the top performing districts for either reading or mathematics. In addition, 80% of the mathematics classrooms, where teachers monitored students' understanding of learning by asking the students to explain the materials, were among the top performing classrooms, and 44% of these mathematics classrooms were in the top performing districts. Finally, while 85% of the mathematics classrooms, where teachers had marked all the work in the student workbooks, were among the top performing classrooms, and 53% of these mathematics classrooms were in the top performing districts; not a single classroom where teachers had marked less than half or none of the work in the workbooks was in the top performing districts.

Figure 24. Characteristics of classrooms observed

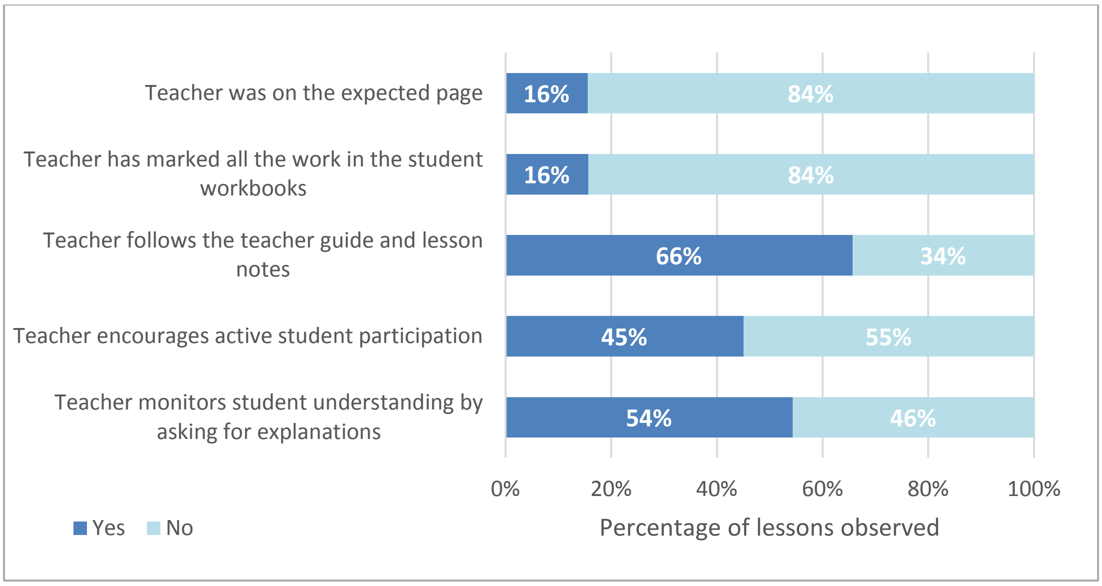


Figure 24 summarizes the extent to which teachers in the observed classrooms (n = 2171) were meeting expectations in implementation, according to criteria discussed above. From the figure, it appears that while more teachers were implementing the participative and student-centered pedagogy of the intervention and were following the routines as described in the teacher guide and lesson notes, the extent to which teachers were on the expected page and marking students' workbooks requires improvement.

The lesson to be learned is that there are a range of teacher actions associated with greater intervention impact. These actions can be monitored by supervisors when they visit teachers; by providing appropriate support to the teachers, supervisors can have a significant impact on the success of the intervention.

4.5 Changes in the Classroom Teacher

Having only one teacher assigned to the class for the year of the intervention was positively associated with intervention impact. Specifically, classrooms with multiple teachers over the course of a year had fewer readers (5.5 percentage points fewer) and mathematicians (3.4 percentage points fewer) than classrooms for which a single teacher was responsible throughout the whole school year. From the dataset for the treatment schools, it could be determined that 78% of classes had a single teacher for the year, and 22% of classrooms had more than one teacher over the course of the year.

While changes in teachers over the school year are unavoidable, the lesson to be learned is that in Jordan there appears to be a high turnover of teachers (more than one-fifth of the classes in the intervention had more than one teacher in the 2013/2014 school year). The implication for an intervention is that the intervention design needs to build in mechanisms for responding to teacher turnover.

4.6 Translating the Vision Into Practice

Much discussion has already been devoted to the challenge of translating the vision of the intervention into classroom practice (see section 1.4). Analysis of teachers' experiences of and responses to the intervention in section 3.5.3 reveals that, in general, although teachers identified more positive than negative aspects to the intervention, when asked whether or not their schools should continue with the implementation of the intervention, the majority said no.

Weighing up all teachers' responses collected through the various questionnaires, focus group discussions, and direct observation of their classes, it is clear that despite all the positive aspects associated with the intervention, teachers regarded the intervention as something to be implemented *in addition to* the curriculum and not necessarily *in support of* the curriculum. As such, they experienced the intervention as a burden. Many teachers asked: "Why are we being burdened with this extra load?"

The lesson learned is that much work needs to be done when implementing an intervention of this type, to ensure that teachers experience the intervention both as supportive of their work and as part of the MoE's program in general.

5 Conclusion and Recommendations

This intervention study set out to explore whether daily practice of foundational skills through deliberate, structured, and developmentally appropriate activities can support children to be able to read with comprehension and do mathematics with understanding. And if so, what the conditions are that help teachers to implement the daily routine and the associated activities with fidelity and confidence.

The results show quite clearly that the intervention did exactly what it was intended to do. While there were virtually no gains in control schools from 2012 to 2014, there were significant gains across treatment schools in reducing the proportion of the lowest performers and increasing the proportion of the highest performers. These results are extremely promising, particularly because the intervention was implemented for only one school year.

5.1 Dissemination Workshop and Recommendations

On September 16–18, 2014, the MoE and USAID hosted a dissemination workshop in Amman. This workshop was conducted to review the findings of the National Early Grade Literacy and Numeracy Survey (Jordan): Intervention Impact Analysis Report, examine the implications arising from those findings, make recommendations for this report, and set benchmarks and targets for students' reading and mathematics achievements in Jordan.

For the first day of the workshop, approximately 40 to 50 participants representing the various ministry departments, representatives from the donor community and from NGOs working in the field of early grade education were in attendance. The workshop's first day served to focus on generating recommendations for this report.

In creating the recommendations that follow, participants were organized into groups that covered the six key lessons learned (see ***Section 4***) emerging from the report. After a period of deliberation, each group reported back to the workshop plenary. Each group's recommendations were then debated by all workshop participants. The following recommendations represent the suggestions of all workshop participants.

Gender

Although neither the survey nor the intervention set out to explore how the role of teacher and student gender affects student performance, the results nonetheless revealed that boys are, in general, not benefiting from early grade instruction in the same way that girls are. Furthermore, the results also indicate that there is a statistically significant relationship between the school type (all-girls, all-boys, and mixed), the gender of the teacher and the performance, by gender, of the students.

Workshop participants' recommendations are that:

1. A qualitative study be conducted to assist the MoE and other stakeholders to better understand why boys do not benefit to the same extent as girls from schooling in the early years. In general, the study needs to explore how and why the school and learning experiences of boys are different from those of girls. Variables that the study needs to take into account include school type, teacher gender, and student gender. The findings of this study should inform the nature of future intervention activities such that those activities should make provision for the issues raised by the study.
2. In the short term, the MoE explore the feasibility of using only female teachers in the early grades, because evidence suggests that female teachers' students generally outperform male teachers' students.
3. In the short term, the MoE explore the feasibility of having only mixed and all-girls schools in the early grades, because evidence suggests that the students in these schools generally outperform the students from all-boys schools.
4. Teacher training modules for inclusion in pre-service, induction, and in-service programs be developed to sensitize teachers to the need to create gender inclusive classrooms that pay equal attention to the needs of boys and girls. In addition, that specific modules be developed to support male teachers in creating supportive classroom atmospheres.
5. The Field Directorates be mandated to increase their support and supervisory visits to all-boys schools.

Classroom support

It is clear from the study data that supervisors' classroom visits contributed to the impact of the intervention on the proportion of readers and mathematicians in a teacher's class. More frequent support resulted in more effective intervention implementation. In the case of this intervention study, however, it was also clear that not all supervisors were able to visit classes as often as the intervention hoped that they would (once every two weeks).

Workshop participants' recommendations are that:

1. The number of support visits for teachers be determined on a "needs" basis, such that teachers identified as being in need of greater support be visited more frequently than teachers who do not need the same level of support.
2. The supervisors and Field Directorates need to actively and frequently monitor how teachers implement new curriculum and programs in their classrooms, to increase the likelihood of effective implementation.
3. The number of teachers for which each supervisor is responsible should be reduced, to allow supervisors to be more effective in providing teacher mentoring and support.
4. Specific training be developed for supervisors that addresses effective mentoring and support strategies.

Teacher participation in in-service training

Teachers participating in the intervention were expected to attend two sets of training sessions: (1) a 10-day training session before the start of the school year and (2) a 5-day training session before the start of the second semester. Although attendance, in general, was good, more than 20% of the teachers did not attend all of the initial 10-day training, and 38% of the teachers did not attend all of the 5-day training at the start of the second semester. The proportion of the training that the teachers had attended was, however, a variable that had a significant impact on the success of the implementation. Teachers who attended more of the training had a greater proportion of readers and mathematicians in their classes than teachers who attended less training.

Workshop participants' recommendations are that:

1. In general, the policy environment for in-service training needs to be more responsive to the needs of teachers, as well as make it easier for teachers to attend in-service training programs. Specific suggestions included:
    - Making the in-service training hours responsive to the needs and realities of teachers, in particular, female teachers who have families to care for.
    - Providing nurseries at training venues for the children of teachers attending the training.
    - Supporting teachers in schools to be able to attend in-service training during school hours, by providing teaching coverage for their classes.

2. Teachers need to be rewarded for participating in in-service training activities. This reward could be either direct, in the form of financial reward, or indirect, in credits earned as part of a continuing professional development program that impacts teacher employment, promotion, and tenure.
3. The need for in-service training on the specific intervention methodologies of this intervention be reduced, by revising the content of pre-service training programs to more deliberately include modules on research-based approaches to teaching early grade reading (phonics approach) and mathematics (problem-based learning).

Fidelity of implementation

Analysis of a wide range of the survey data clearly indicates that a range of teacher actions are associated with greater intervention impact (e.g., being on the correct page of the program, following the activity notes, monitoring student understanding, ensuring that students are working in their workbooks, and marking the work of students in the workbooks). These actions can be monitored by supervisors using simple checklists when they visit teachers. By monitoring teachers on the fidelity with which they perform these actions, supervisors (and school principals) can have a significant impact on the success of the intervention.

Workshop participants' recommendations are that:

1. Supervisors and school principals take greater responsibility for monitoring the fidelity with which teachers implement an intervention. This monitoring is less involved with the pedagogical fidelity of implementation and more with a wide range of simple binary (yes/no) indicators: For example, if teachers are on the correct page (Y/N), etc.

Changes in classroom teacher

Having only one teacher assigned to the class for the year of the intervention was positively associated with intervention impact. In the case of this intervention, 78% of classes had a single teacher for the year, and 22% of classrooms had more than one teacher over the course of the year. Seen realistically, changes in teachers over the school year are, however, unavoidable.

Workshop participants' recommendations are that:

1. Intervention programs need to include mechanisms for responding to teacher turnover. These mechanisms could, at a minimum, address how substitute and replacement teachers will be trained after permanent teachers in a school have already been trained, and whether or not temporary teachers will be expected to participate in the in-service training activities of an intervention.

Translating the vision of the intervention into practice

The challenge of translating the intervention methodology (vision) into classroom practice was not insubstantial. Detailed discussions about these challenges are found both in ***Section 1.4*** and in ***Section 4.6*** of this report. In short, teachers experienced difficulties in assimilating new pedagogies into their practice. And, some of teachers' criticisms about the intervention reflected not so much fundamental problems with the intervention, but rather mismatches between the intervention vision(s) and teachers' predominant teaching styles and approaches.

Workshop participants' recommendations are that:

1. Intervention implementers need to take care to ensure that teachers experience:
   - Intervention activities not as add-ons to the work that they do, but instead as supportive of and integral to what they do.
   - Intervention activities as activities of the MoE and directly linked to the curriculum. Information briefs for teachers linking the methodologies to research evidence would be helpful. Similar briefs for parents would assist teachers and schools in managing parent questions and expectations.
   - Successful program implementation and student performance as linked to their promotion and tenure.
2. Teachers be encouraged to commit to exploring new methodologies; success stories about program impact be shared with teachers through a range of media, including social media web-based communities of practice.
3. Video vignettes be developed. These videos should demonstrate the desired methodologies being successfully implemented by teachers in typical classrooms. Such videos can be used by trainers in in-service training activities and made available to teachers through the MoE intranet and social media web-based communities of practice.

5.2 Benchmarks and Targets

As part of the dissemination workshop hosted by the MoE and USAID on September 16–18, 2014, participants representing the various ministry departments set benchmarks and targets for EGRA and EGMA on the second day of the workshop.

The benchmarks reported in ***Table 21*** were based on the results of grade 2 and grade 3 Jordanian students in the 2014 National Survey and informed by a range of international benchmarks, the participants' experience with and knowledge of the Jordanian context, and technical support provided by the researchers who led the RTI research team.

The 5-year targets reported in ***Table 21*** were set by the participants, with support from the technical experts based on the evidence of the intervention's potential impact, as described in this report. In addition, the targets also assume that many of the recommendations made in this report are implemented and that a concerted effort is exerted by all stakeholders in applying these recommendations.

Table 21. Grade 2 and 3 (combined) performance benchmarks and 5-year targets for reading and mathematics.

		Reading			Mathematics	
		Non-word decoding	Oral reading fluency (ORF)	Comprehension	Addition and subtraction level 2	Missing number
Benchmark		**23 correct words per minute (cwpm)**	**46 cwpm**	**80% correct**	**80% correct**	**70% correct**
Percentage of students at benchmark	2014 actual[+]	5.3%	7.5%	17.9%[+++]	16%	39%
	5 year target	31%	35%	57%	40%	58%
Percentage of zero scores	2014 actual	30.7%	9.1%	34.1%	13.4%	3.1%
	5 year target	13%	5%	13%	5%	__[++]

[+] The 2014 actual is for Grade 2 and Grade 3 combined and is based on the May 2014 national EGRA and EGMA survey.

[++] No five-year target has been set since the baseline performance is already acceptable.

[+++] Note that the 2014 actual percentage of students at benchmark is very similar to the value established in the 2012 national survey (17.1%).

Annex 1: Summary of teacher responses to the written questionnaire (n = 233)

Figure 1.1. Reading materials—what was positive, negative, and could be improved?

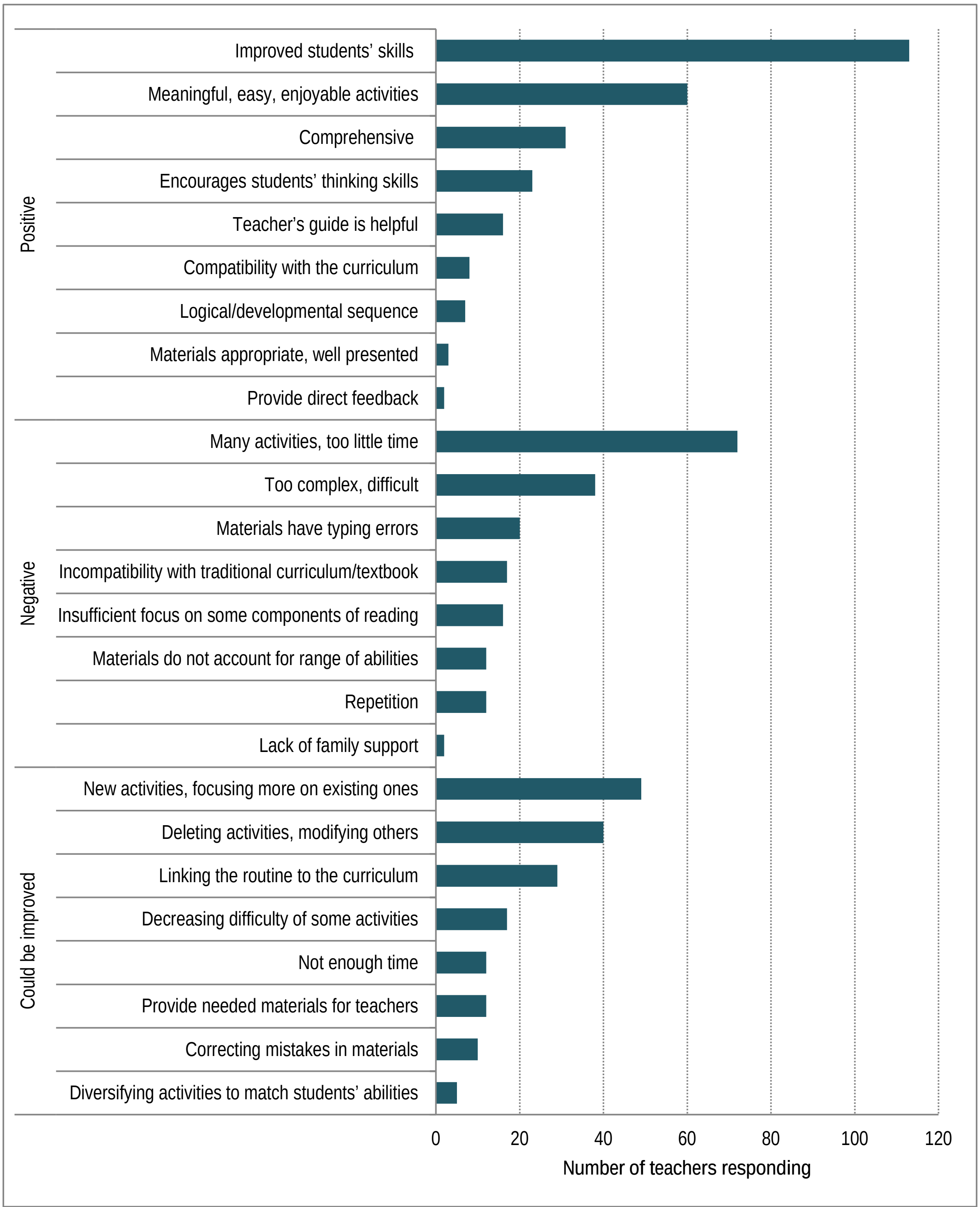


Figure 1.2. Reading implementation—what was positive, negative, and could be improved?

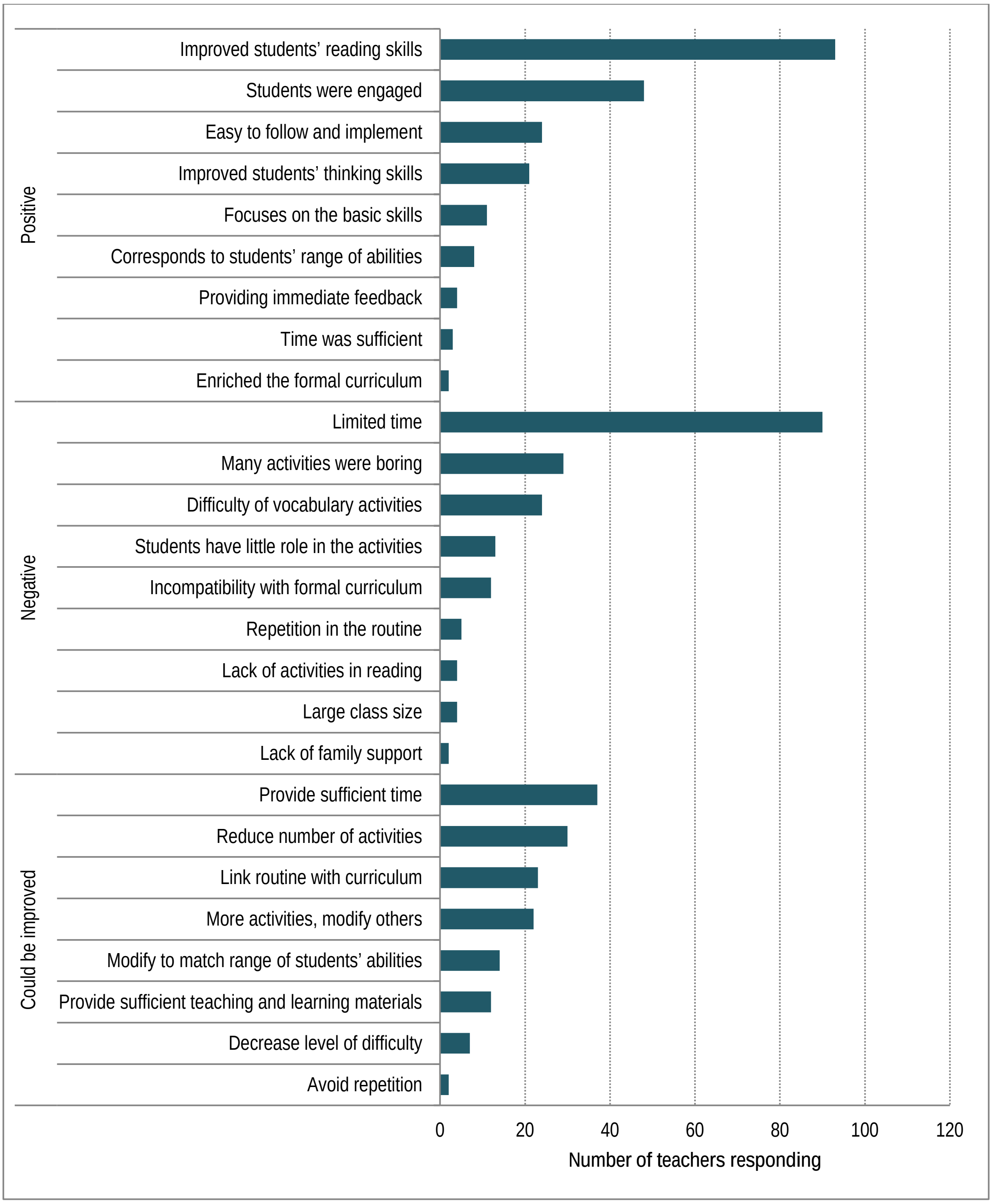


Figure 1.3. Student response to reading—what was positive, negative, and could be changed to improve the students' response?

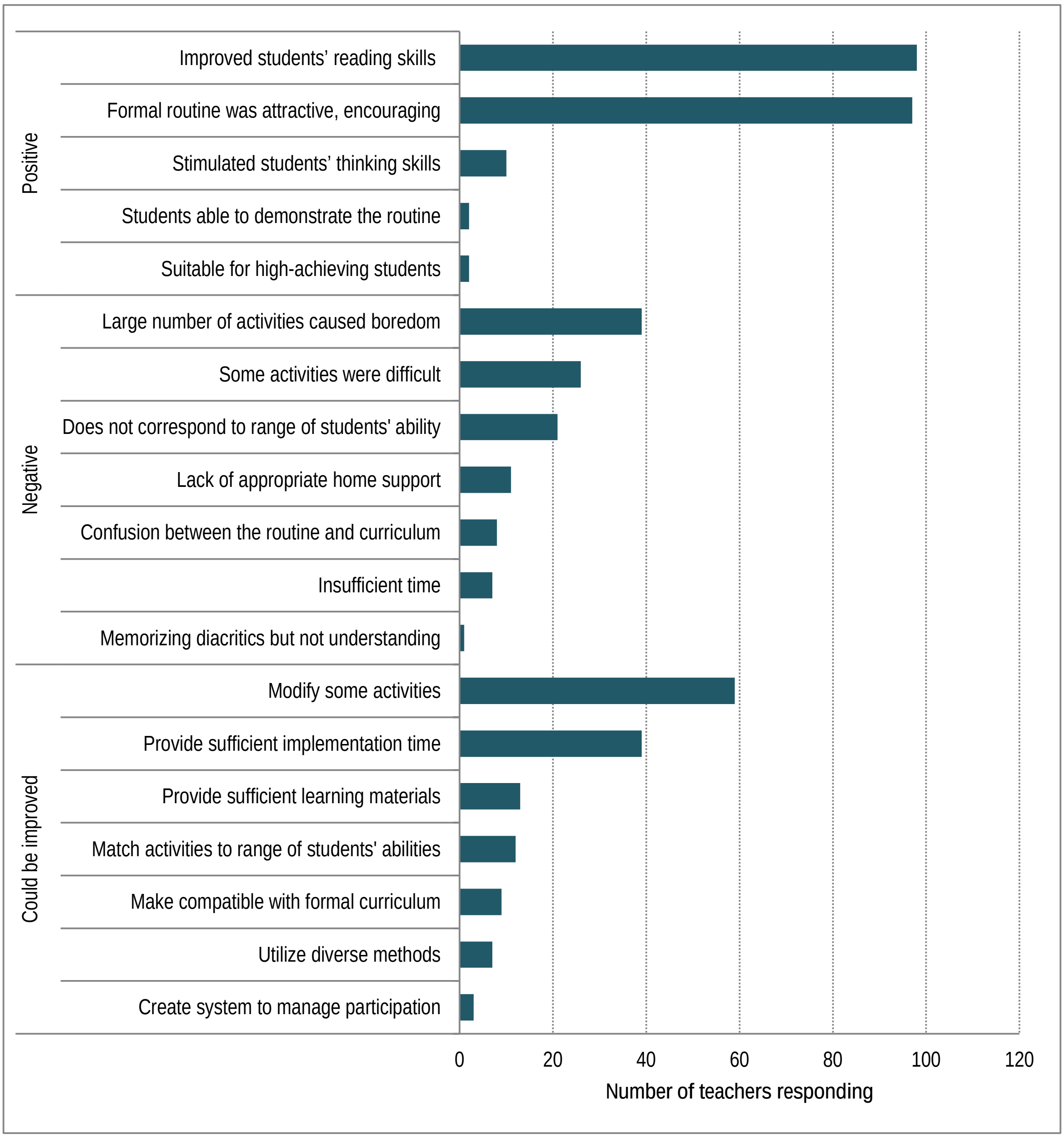


Figure 1.4. Supervisor support to reading—what was positive, negative, and could be changed to improve the supervisors' support of reading?

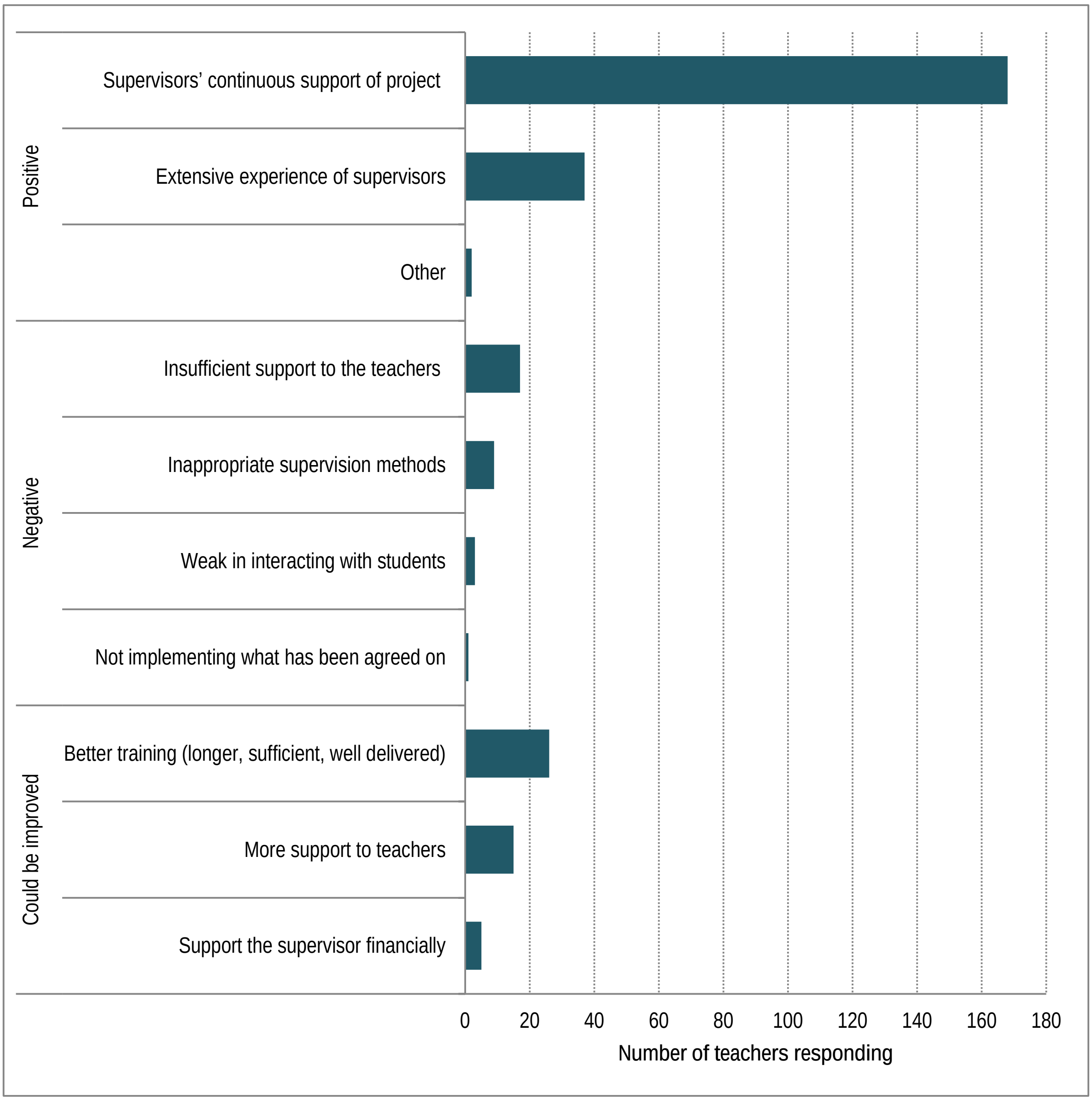


Figure 1.5. Mathematics materials—what was positive, negative, and could be improved?

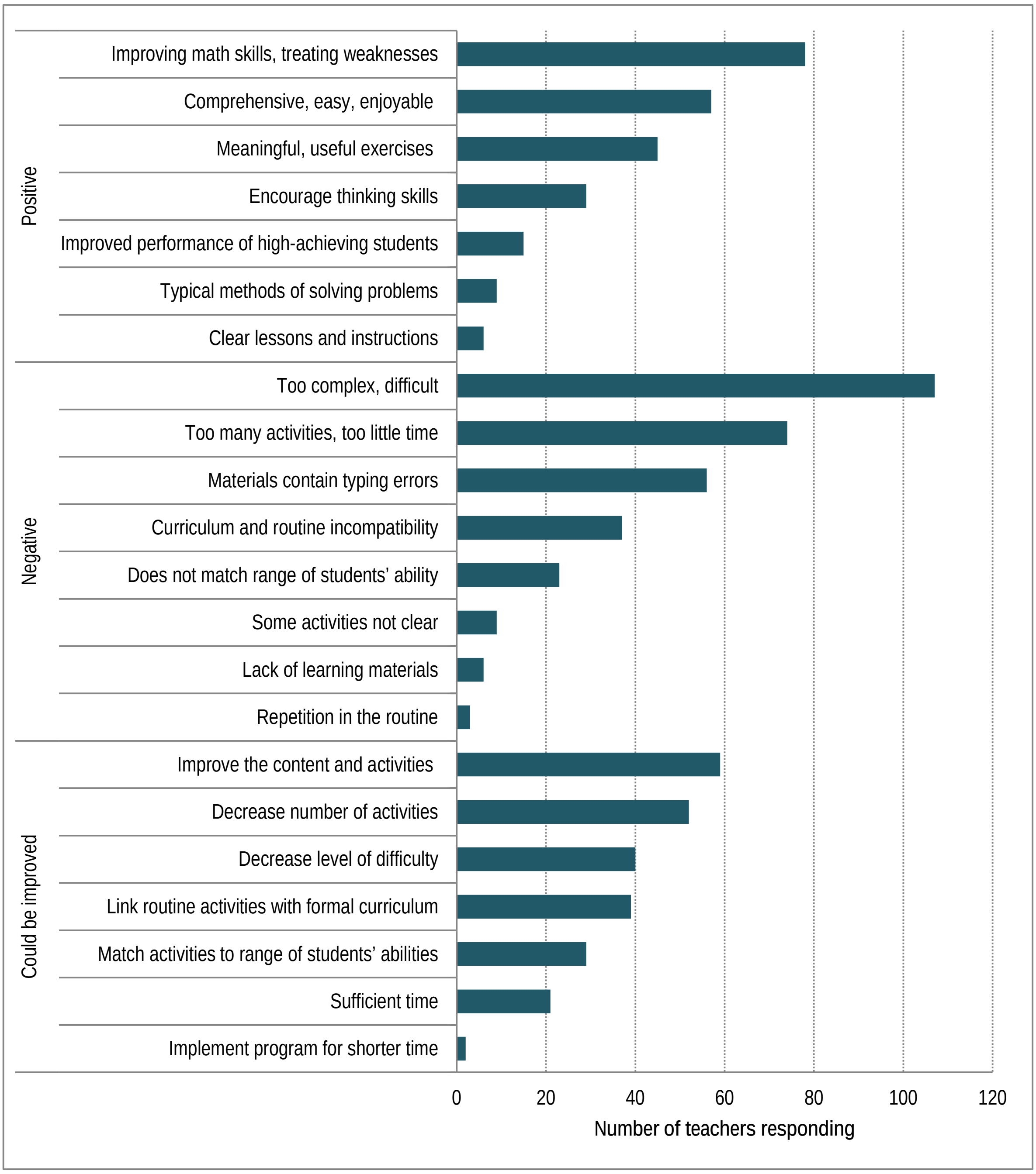


Figure 1.6. Mathematics implementation—what was positive, negative, and could be improved?

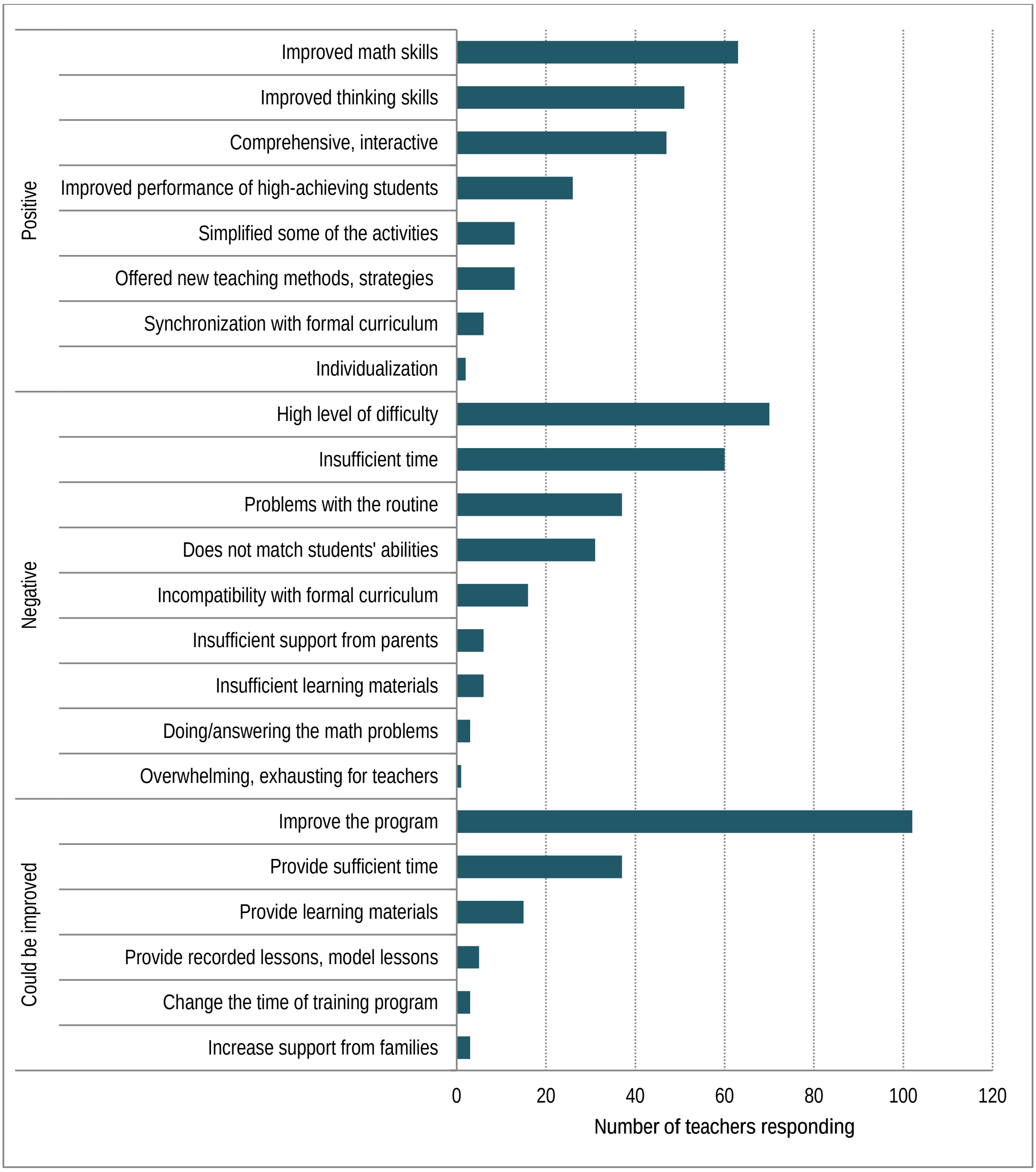


Figure 1.7. Student response to mathematics—what was positive, negative, and could be changed to improve the students' response?

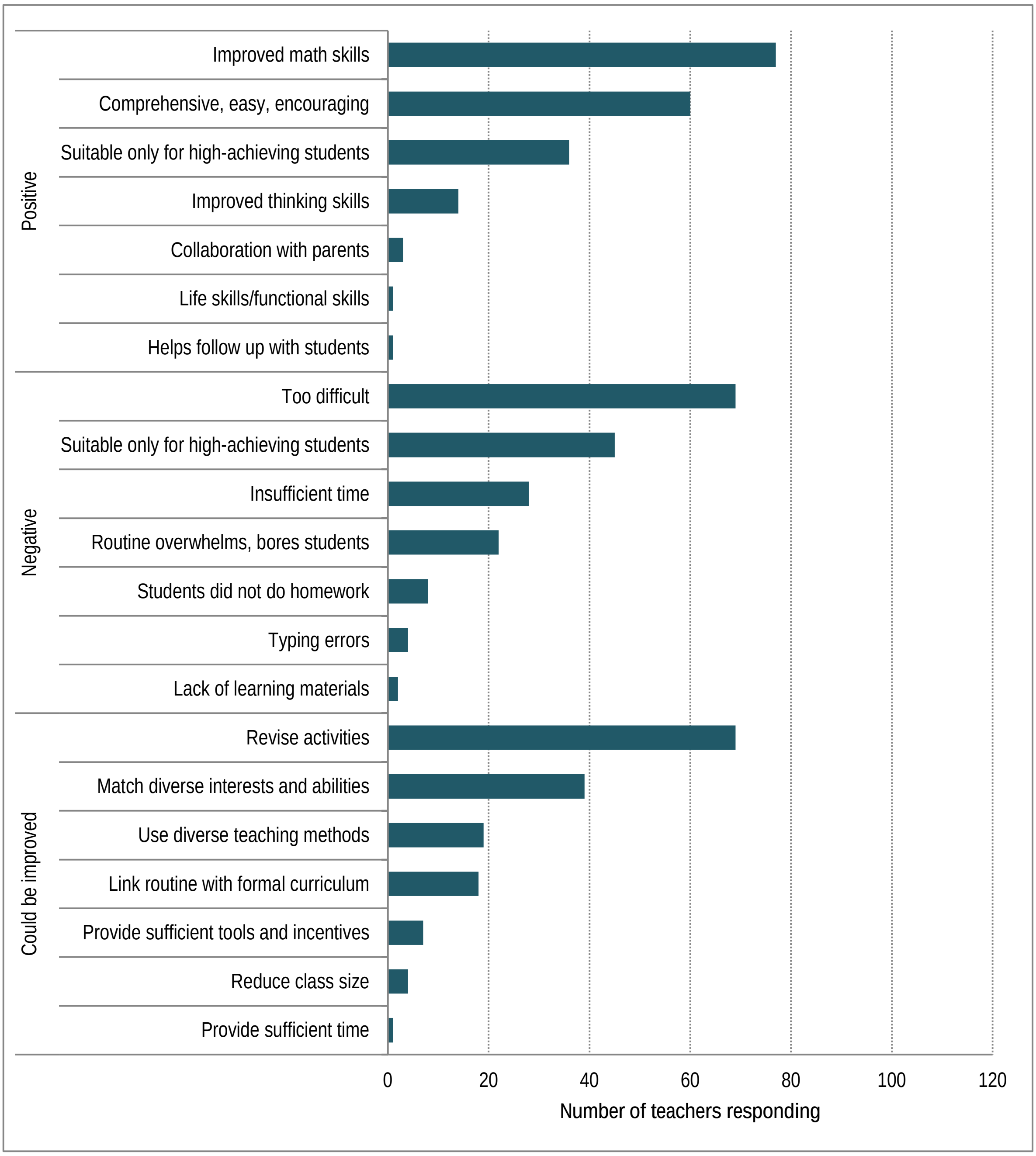
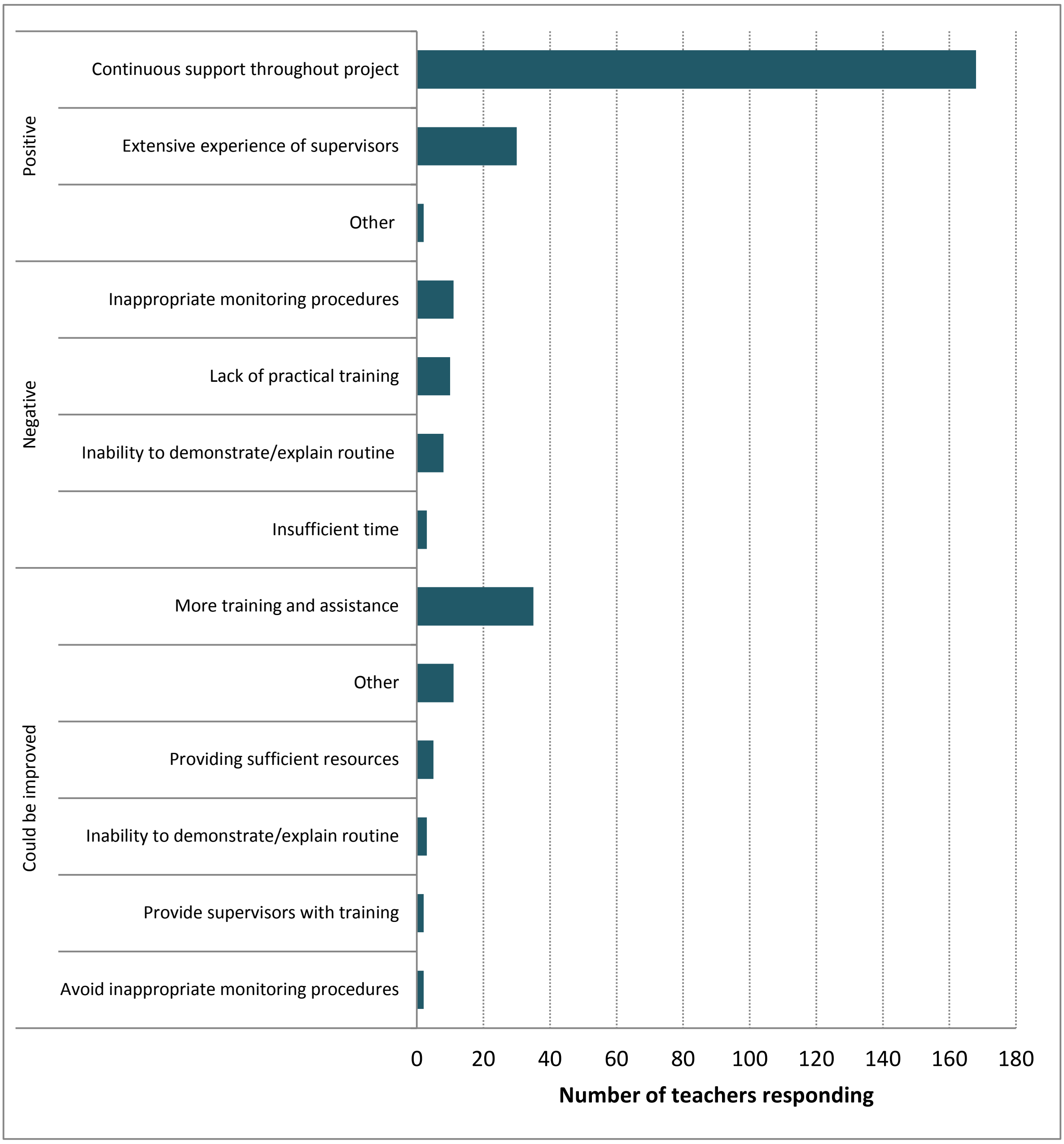


Figure 1.8. Supervisor support to mathematics—what was positive, negative, and could be changed to improve the supervisors' support of mathematics?



Annex 2: All instruments

Annex 3: List of MoE Contributors

Name	Department	Role(s)
Afaf Arar	School	Materials development, survey instrument adaptation
Afaf Eitom	Field Directorate	Teacher training and coaching, survey instrument adaptation, survey administration (assessor)
Ahmad Kalbounah	Field Directorate	Survey administration (assessor), survey instrument adaptation
Ala'a Abu Jaber	Curriculum	Materials development
Ali Khleifat	Field Directorate	Survey administration (assessor)
Bassam B'deir	Field Directorate	Survey administration (assessor)
Dr. Ahmad Al Ajarmeh	Examinations	Materials development, survey instrument adaptation
Dr. Ahmad Al Salamat	Curriculum	Materials development, teacher training and coaching
Dr. Ali Abd El Baqi	Field Directorate	Teacher training and coaching
Dr. Amal Al Bajawi	ETC	Materials development, teacher training and coaching, survey instrument adaptation
Dr. Ayed Al Athamat	Field Directorate	Survey administration (assessor)
Dr. Basma Muammar	Field Directorate	Materials development, survey administration (assessor)
Dr. Fatima Al Boursan	Field Directorate	Teacher training and coaching, survey administration (assessor)
Dr. Hassan Al Rababa	Field Directorate	Materials development, survey instrument adaptation, survey administration (assessor)
Dr. Khaled Al Najjar	Field Directorate	Materials development, teacher training and coaching, survey instrument adaptation, survey administration (assessor)
Dr. Khawla Abu Al Haija	ETC	Project leadership
Dr. Mahmoud Al Jarrah	Field Directorate	Teacher training and coaching
Dr. Maleeha Addamkh	ETC	Project leadership
Dr. Osama Jaradat	Curriculum	Materials development, survey instrument adaptation
Dr. Qaseem Hamadneh	Field Directorate	Teacher training and coaching
Dr. Radi Al Shunnaq	Field Directorate	Teacher training and coaching
Dr. Shaker Al Qaoud	Field Directorate	Teacher training and coaching
Falah Al Mashaqbeh	Field Directorate	Teacher training and coaching
Fathiyyeh Bazbaz	Field Directorate	Survey administration (assessor)
Firyal Aqel	DCU	Project leadership
Hani Al Jabali	Field Directorate	Teacher training and coaching
Imad Al Ardah	ETC	Teacher training and coaching
Imad Naamneh	Curriculum	Materials development
Issam Shatnawi	Curriculum	Materials development, survey instrument adaptation

Name	Department	Role(s)
Jihad Abu Al Rizeq	Field Directorate	Teacher training and coaching
Jihad Abu Al Roukab	Field Directorate	Materials development, Survey instrument adaptation
Khaled Al Jaddou	Curriculum	Materials development
Khitam Al Sawarees	ETC	Teacher training and coaching, survey instrument adaptation
Lafi Al Baqum	Field Directorate	Survey administration (assessor)
Lana Arafah	Field Directorate	Materials development, survey instrument adaptation
Mahmoud Al Qatamin	Field Directorate	Teacher training and coaching, survey administration (assessor)
Mansour Al Oneh	Field Directorate	Teacher training and coaching, survey administration (assessor)
Muhammad Kinana	Examinations	Materials development, survey instrument adaptation
Muna Al Haja	Field Directorate	Materials development, survey instrument adaptation
Muntaha Al Tartir	School	Materials development
Nadera Al Sleibi	School	Materials development, survey instrument adaptation
Naif Al Rifae	ETC	Teacher training
Nawal El Hambooth	Field Directorate	Materials development, survey instrument adaptation
Nawal Madi	Field Directorate	Teacher training and coaching, survey instrument adaptation, survey administration (assessor)
Nehaya Al Rimawi	Field Directorate	Survey administration (assessor)
Nezar Al Doqs	Field Directorate	Survey administration (assessor)
Nowwar Ifteihah	Examinations	Survey instrument adaptation, survey administration (assessor)
Omar Abu Saif	Field Directorate	Survey administration (assessor)
Qasem Shqerat	Field Directorate	Teacher training and coaching, survey administration (assessor)
Rabi'a Al Moumani	Field Directorate	Teacher training and coaching, survey administration (assessor)
Raed Aqel	Field Directorate	Teacher training and coaching
Rima Zreiqat	Field Directorate	Survey administration (assessor)
Rihan Al Mustafa	Field Directorate	Survey administration (assessor)
Salim Al Harahsheh	Field Directorate	Teacher training and coaching
Sameeh Al Momani	Field Directorate	Teacher training and coaching
Shadia Gharaybeh	Curriculum	Materials development
Shatha Al Bau	School	Materials development
Sumayya Jaradat	Field Directorate	Teacher training and coaching, survey administration (assessor)
Thaera Abu Dayyeh	ETC	Survey instrument adaptation

Name	Department	Role(s)
Wafa Al Abdallat	Curriculum	Project leadership
Wafa Sawman	School	Materials development, survey instrument adaptation