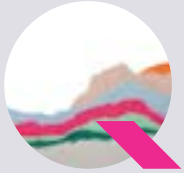
MORPHOLOGICAL AWARENESS
FEASIBILITY STUDY QUANTITATIVE REPORT



QUEEN RANIA FOUNDATION

مؤسسة الملكة رانيا

REPORT AUTHORS

Jumana Shahzadeh
Ghalia Ghawi

ACKNOWLEDGEMENTS

The authors of this report would like to thank the team at the Ministry of Education, the Queen Rania Teacher Academy, and Dr. Abdallah Al Shdaifat for developing the program content and training, and the principals, teachers, and students for opening up the schools and classrooms for this research.

The authors would like to extend their gratitude to the team at QRF for their valuable contributions, including Reema Abed, Nafez Dakkak, Lucie Moore, Raed Omar, and Ruba Samain.

CAPI	Computer-assisted personal interviewing
EGRA	Early Grade Reading Assessment
HRD	Human Resources and Development
PISA	Programme for International Student Assessment
QRF	Queen Rania Foundation
QRTA	Queen Rania Teacher Academy
RAMP	Reading and Math Project
RCT	Randomized Control Trial
RTI	Research Triangle Institute (RTI International)
TIMSS	Trends in International Mathematics and Science Study

LIST OF TABLES AND FIGURES

EXECUTIVE SUMMARY

This report presents the endline findings of a feasibility study of the Morphological Awareness pilot, implemented by the Queen Rania Foundation in partnership with the Queen Rania Teachers Academy and the Ministry of Education.

The Morphological Awareness pilot was designed to feed into the 'Modernization of the Basic and Secondary Curriculum' project, which is part of the National Human Resources and Development strategy (HRD) 2016 - 2025. The HRD strategy outlined the need to develop new curricula in line with international best practice. This pilot was intended to address the issue of low achievement in Arabic language literacy in the country, by creating supplementary textbooks that support the existing national curriculum.

The main goal of this pilot was to improve Arabic language outcomes of 2nd to 5th grade students; mainly comprehension, through instruction on Arabic morphemes. The intervention trained teachers on how to explicitly teach their students morphological rules in the Arabic language.

To evaluate the pilot, a randomized control trial was conducted in 20 schools, and was implemented in 10 of those schools for one academic year (September 2017-May 2018). A mixed method approach consisting of assessments, perception surveys and focus groups was used to evaluate the pilot.

After analysing the results, it was found that the pilot did not improve the literacy outcomes of the students. This could be due to several reasons, including the small sample size, selection bias and assessment tools. Further exploration and a more comprehensive process evaluation on the influence of morphemes and how to integrate it in the Arabic curriculum need to take place before such a pilot is considered for scale.

INTRODUCTION

The Queen Rania Foundation for Education and Development (QRF) is running a series of pilots to explore what works to improve the literacy, numeracy and socioemotional learning outcomes of students in Jordan. One such pilot is the Morphological Awareness pilot, which was tested across a sample of 2nd to 5th grade classrooms, with an aim to increase literacy outcomes.

Two randomized control trials were conducted, each of which targeted ten schools. The first RCT targeted 2nd and 3rd grade students in low-fee private schools and the second RCT targeted 4th and 5th grade student in public schools. A mixed method approach was used to evaluate the impact of the pilot.

RATIONALE FOR THE **MORPHOLOGICAL AWARENESS PILOT**

International assessments such as the Programme for International Student Assessment (PISA), Early Grade Reading Assessment (EGRA), and the Trends in International Mathematics and Science Study (TIMSS) have highlighted the low and declining achievement in reading and literacy for students in Jordan.

In response to the EGRA 2012 results that showed low levels of literacy achievement, the Ministry of Education moved towards a phonetic approach to teaching language in 2014 (i.e., the teaching of letter sounds). As such, the Jordanian education system became focused on teaching the Arabic language phonetically for the early grades.

However, research has shown that such an approach may not be the most effective for developing literacy; such that teaching phonemes (units of sound in a language) alone is not sufficient (Whitehurst & Lonigan, 1999). Teaching language phonetically can be effective in children's reading but not comprehension skills. A more effective approach to teaching language could be including explicit references to morphemes in instruction. Morphemes are the shortest components of language that have a meaning. Extensive international research investigating the influence of morphological awareness on literacy has reflected its importance in reading, vocabulary and comprehension skills in languages. Bowers et al. (2010) reviewed 22 such studies in their meta-analysis. The selected studies examined reading, spelling and vocabulary outcomes as a function of morphological awareness. The findings showed that morphological awareness and instruction positively influenced learners, young and old, and is highly beneficial for less able readers. Additionally, Kirby et al. (2012) found morphological awareness to be a significant predictor of several aspects of literacy, such as text reading speed and reading comprehension. Increases in vocabulary, spelling abilities, reading abilities, and comprehension have been found as a result of morphological instruction (Bowers et al., 2010; Carlisle, 2010; Kirk & Gillon, 2009).

Few studies have examined the influence of morphological awareness on language development in Arabic. Of those studies, Shalhoub-Awwad and Leikin (2016) have found that students' knowledge of root words was a predictor of children's vocabulary. Taha and Saiegh-Haddad (2016) additionally investigated the influence of two interventions, a phonological awareness intervention, and a morphological awareness intervention, on word spelling, in comparison to a control group. Results showed the interventions were both successful in promoting correct spelling and linguistic awareness. Additionally, the skills children use in reading Arabic were examined in a literature review (Al-Ghanem & Kearns, 2014). Morphological skills were examined in only three studies; but it was found that there was some evidence of association with word reading, especially for skilled readers.

Considering that Arabic is a morphologically rich language, and a reader can come across several unfamiliar words in a text due to the extensive inventory of affixes and clitics, introducing explicit morphological instruction in teaching could be greatly beneficial for comprehension. Currently, morphemes in the Arabic language are taught through pattern-recognition. Through this teaching approach, children are expected to deduce rules and grammar through sentences and sentence completion; such as, "The child runs > the child ran," "The child drinks > the child?" In that example, the student is expected to deduce the past tense of "drinks" and complete the sentence. No explicit reference is made to the meaning of the word "drinks". However, introducing explicit instruction of morphemes into teaching (i.e. making explicit reference to the meaning of root words, suffixes, and prefixes) may have extensive benefits towards improving children's literacy and comprehension skills in the Arabic language.

The Morphological Awareness pilot was designed to feed into the 'Modernization of the Basic and Secondary Curriculum' project, which is part of the National Human Resources and Development strategy (HRD) 2016 - 2025. The HRD strategy outlined the need to develop new curricula in line with international best practice. This pilot was intended to address the issue of low achievement in Arabic language literacy in the country, by creating supplementary textbooks that support the existing national curriculum.

THE MORPHOLOGICAL AWARENESS PILOT

BACKGROUND AND OBJECTIVES

SUMMARY	
Type of study	Feasibility study
Grade Range	Grades 2 - 5
Number of Schools	20
Number of Students	928 [1]
Type of intervention	Teacher training and provision of materials targeting the teaching of morphemes
Date of intervention	September 2017 - May 2018
Dosage of delivery	1 hour per week as part of the Arabic classes

The pilot entails a pedagogical intervention targeting the teaching of morphemes. The intervention trained teachers on how to explicitly teach their students morphemes, and morphological transformations and rules. Research has shown that an explicit approach may improve spelling, reading, vocabulary, and comprehension (Bowers et al., 2010; Carlisle, 2010; Kirk & Gillon, 2009). It is worth noting that all the existing research focusing specifically on explicit morphemes instruction in the English and Spanish language. The Arabic Morphemes language studies were correlational and did not assess explicit Morpheme instruction (Shalhoub-Awwad and Leikin, 2016). The pilot hypothesizes that training teachers to explicitly teach morphemes in 2nd to 5th grade classrooms will be beneficial to students' Arabic literacy development.

The overall objective of the pilot was to evaluate whether this learning method helps students increase their overall text comprehension, in addition to reading abilities for students in grades 2-3.

[1] Total number of students surveyed at endline.

PILOT CONTENT AND MATERIALS

The content of the intervention was aligned with the national Arabic curriculum for grades 2-5; i.e., it followed the same levels, progression and lesson plans, but supplemented what was taught by introducing explicit instruction on morphemes. Dr. Abdullah Al Shdaifat[2] created the initial framework for the content, and gave general examples on morphological transformations and exercises. This framework was used to develop the content and materials.

The content was developed by a task force consisting of members of the Queen Rania Teacher's Academy (QRTA) and the Ministry of Education (MoE). The task force developed booklets for the teachers which guided them on what material to develier to the students, giving examples on how to explain the content and included the morphological rules, student booklets which acted as supporting materials to the existing textbooks, and flashcards to be used in class. QRTA were also responsible for teacher training.

The explicit morphological instruction introduced was based on and aligned with the patterns the students are meant to solve in textbooks. Since all public and private schools in Jordan are required to follow the same curriculum, we envisioned this approach would be sustainable as teachers would be able to follow the supplementary material without diverting from their lesson plans.

[2] Assistant Professor at the University on Jordan

THEORY OF CHANGE (ToC)

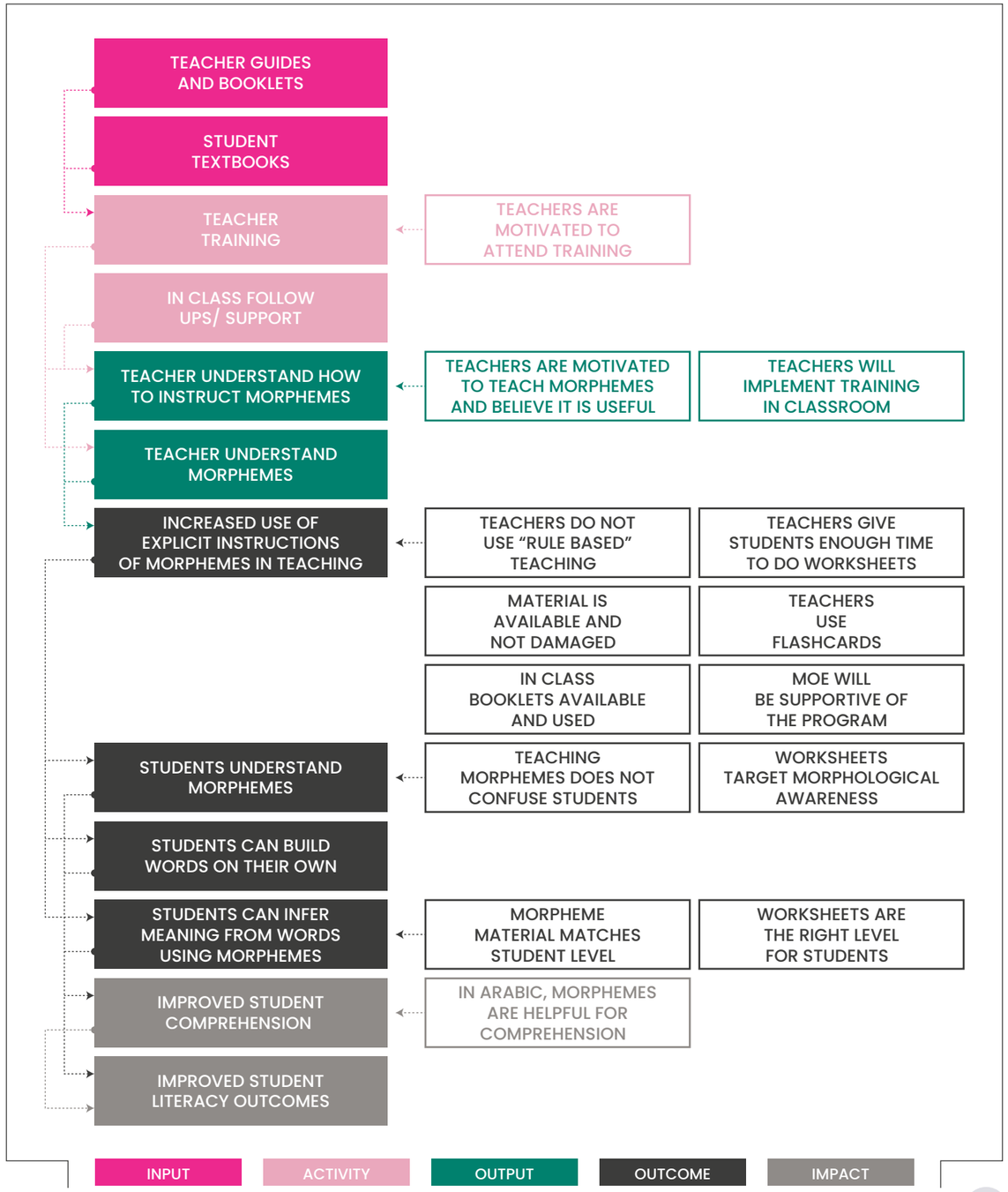
The design of the pilot was informed by a theory of change that illustrated how morphological awareness training for 2nd to 5th grade teachers could result in the desired impact on students' literacy outcomes. Figure 1 depicts the pilot's theory of change. The ToC was developed to identify key hypotheses associated with the ToC and the core areas to probe in the qualitative section of the study. The ToC allowed the setting of indicators to evaluate the outcomes of the pilot.

KEY EVALUATION HYPOTHESES

- ***HYPOTHESIS ONE*:** The training will provide teachers with a clear understanding of morpheme rules, how to develop their teaching methods, and how to use the supplementary material (booklets and flashcards). This was expected to result in teachers being able to instruct on morphemes explicitly.
- ***HYPOTHESIS TWO*:** If teachers can use explicit instruction of morphemes while leveraging the pilot content and material, they will be more confident and motivated to teach morphemes explicitly in the class.
- ***HYPOTHESIS THREE*:** If teachers use explicit instruction of morphemes using the new teaching method and supplementary material, students will be more engaged and interested in Arabic lessons.
- ***HYPOTHESIS FOUR*:** If students are engaged in the class and able to solve the activities in the supplementary textbook, their morphological awareness will improve.
- ***HYPOTHESIS FIVE*:** If the students' morphological awareness improves, their text comprehension skills will improve. This was believed to improve the students' overall literacy level.

FIGURE 1: MORPHOLOGICAL AWARENESS TOC

Context: Teachers are incentivized to finish curriculum and not focus on learning



EVALUATION METHOD

OVERVIEW

The overall aim of this feasibility study was to answer the questions:

EVIDENCE OF PROMISE
To what extent does it appear that the theory of change of the intervention holds in this context?

FEASIBILITY
Can the morphemes awareness pilot be implemented as intended in the Jordanian context?

SCALABILITY
Is the intervention scalable?

To do so, a mixed method approach was used. Quantitative and qualitative tools were used to collect data from school teachers and students to test all the components in the theory of change and answer all the core areas probed in the evaluation matrix.

Two randomized control trials (RCT) were conducted; each targeted ten schools in Amman. The first RCT targeted 2nd and 3rd grade students in low-fee private schools [3], and the second RCT targeted 4th and 5th grade student in public schools. Each RCT had five treatment and five control schools of similar characteristics. The characteristics included: gender, location within Amman, number of students per class.

Both treatment and control schools followed the national curriculum; however, treatment school teachers were given a teacher booklet that included new ways to teach Arabic morphemes, and were trained on how to use the suggested approaches. Additionally, students were supplied with exercise booklets to scaffold their learning. By conducting pre and post surveys for both groups, we were able to measure the impact of the pilot on the students' learning outcomes and attitudes on the new method of learning.

Both quantitative and qualitative data were collected for the evaluation. EGRA was used to assess 2nd and 3rd grade students and tests developed by the MoE were used to test 4th and 5th grade students. Additionally, a perception and technical survey on morphemes were administered to all grade levels. Focus groups were conducted with the treatment school teachers at midline and endline to solicit their opinions and feedback on the design, training, content, and implementation of

[3] Due to the implementation of the Reading and Math Project (RAMP) in grades 1-3 across all public schools in Jordan, the QRF pilot was implemented in low-fee private schools for Grade 2 and 3.

the pilot. QRF's implementation unit visited each school twice per semester and filled out a follow-up sheet, which has information on the implementation. QRTA conducted follow-up visits with the schools to provide technical support.

SAMPLE SIZE AND POWER CALCULATIONS

A total of 20 schools were selected to be part of the pilot. To avoid intervention spill-over, control and treatment classrooms for the same grade level were chosen from different schools. The schools were selected randomly based on the following criteria:

- All schools should be located in Amman and have more than 20 students in the target grade.
- Ten mixed gender, low fee private schools for grades 2 and 3.
- Ten public schools for grades 4 and 5, six of which were female schools, and four were male schools.

Schools were randomly assigned to a treatment. Each school principal was responsible for nominating two teachers, one from each grade to be part of the pilot. The principals chose the teachers based on their level of engagement and their availability as some teachers may have not been available due to prior commitments

POWER CALCULATIONS

The main goal of the pilot was to improve the students' Arabic language outcomes; mainly comprehension, through instruction on Arabic morphemes. Therefore, for our power calculations, we used reading comprehension and listening comprehension indicators[4] for the 2nd and 3rd grade students and used the total score on the MoE test for the 4th and 5th grade students.

POWER CALCULATIONS

Following the random selection of the schools, one section per grade was surveyed. In total, ten classrooms (2 grades per school) were surveyed in each treatment arm. The average number of students per classroom was 20 (min=11, max=27).

The results showed that the minimum detectable effect size (MDE) that this study will be able to detect is around 1.05SD. Since this is much higher than average effect sizes found in other studies, it is quite unlikely that this study will find statistically significant evidence regarding the pilot's impact on literacy. The reason behind the high MDE is the small number of clusters in the sample. Cluster RCT's with low numbers of clusters are usually not recommended since the fewer the clusters, the higher the MDE needed to detect if the pilot had an impact. A larger RCT with more schools involved would have decreased the MDE number for both indicators.

[4] Both indicators are subtasks of the EGRA assessment

EVALUATION TIMELINE

This pilot ran throughout the 2017-2018 academic year. A baseline survey was administered in September 2017, at the beginning of the academic year, to benchmark where the students are at prior to the pilot commencement (in terms of student levels of text comprehension and morphemes understanding). Focus groups with teachers were conducted at the end of each semester (December and May) to gain qualitative information about the pilot (midline and endline). The endline survey was administered at the end of the year to measure student abilities.

ANALYSIS METHOD

To determine if this pilot had an impact on student learning outcomes, we used descriptive and regression analyses.

Four regression models were used. Clustering at the school level was used throughout, considering that the school level is the level that the treatment was assigned at.

Our basic regression model compares treatment and control students at endline with clustering at the school level.

Equation 1

$$Y_{ij}^{endline} = \beta_0 + \beta_1 Tr + \varepsilon$$

Where Y is the outcome variable at endline (time=1) for student i in school j, Tr denotes treatment status and is a dummy variable that indicates treatment status, and ε is the error term.

To control for observable differences between the groups, a similar regression was run using controls. The controls included child age and mothers' and fathers' employment status at baseline. The equation for this model is:

Equation 2

$$Y_{ij}^{endline} = \beta_0 + \beta_1 Tr + \mathrm{X_{ij}}\gamma + \varepsilon$$

Where X_{ij} is a vector of control variables for person i in cluster j, measured at baseline time=0

In order to control for differences at baseline between treatment and control groups, we also ran a set of ANCOVA regressions.

Equation 3

$$Y_{ij1} = \beta_0 + \beta_1 Tr + \beta_2 Y_{ij0} + \varepsilon$$

Where Y_{ijt} is the outcome variable of student i in school j measured at time t, and Tr is the treatment status.

In the final model, the same controls which were used in equation 2 were also accounted for.

Equation 4

$$Y_{ij1} = \beta_0 + \beta_1 Tr + \beta_2 Y_{ij0} + X_{ij}\gamma + \varepsilon$$

RESULTS - **GRADES 2 AND 3**

The survey for grades 2 and 3 consisted of three parts:

- The Early Grade Reading Assessment (EGRA) test which measures student comprehension and vocabulary
- A student satisfaction survey which measures the student satisfaction on Arabic class
- A technical survey that tests the student's morphological awareness.

PART 1: STUDENT COMPREHENSION AND VOCABULARY

The EGRA is an interview based oral assessment of the basic foundation skills that are known to predict reading success for students in the early grades of primary school which includes 1st to 3rd grade (Sprenger-Charolles, 2008). This assessment was created by RTI International in 2007.

In order to assess the students' abilities and skills, three different measures were used:

- The number of correct responses per minute
- Percentage of correct responses from items attempted
- Percentage of students with zero scores.

NUMBER CORRECT RESPONSES

We found that the pilot did not have a statistically significant effect on any of six EGRA subtasks for grade 2 or grade 3 students. Both treatment and control students improved throughout the year, but the improvement in treatment schools was generally no greater than the improvement in control schools and as such, the regression results suggested that the pilot did not have a statistically significant effect on students' literacy outcomes.

Figure 2, Figure 3 and Table 6 show the baseline and endline averages for students in grades 2 and 3 in more detail. The results of the regression are displayed in Table 9.

FIGURE 2: GRADE 2 - EGRA AVERAGE SCORE

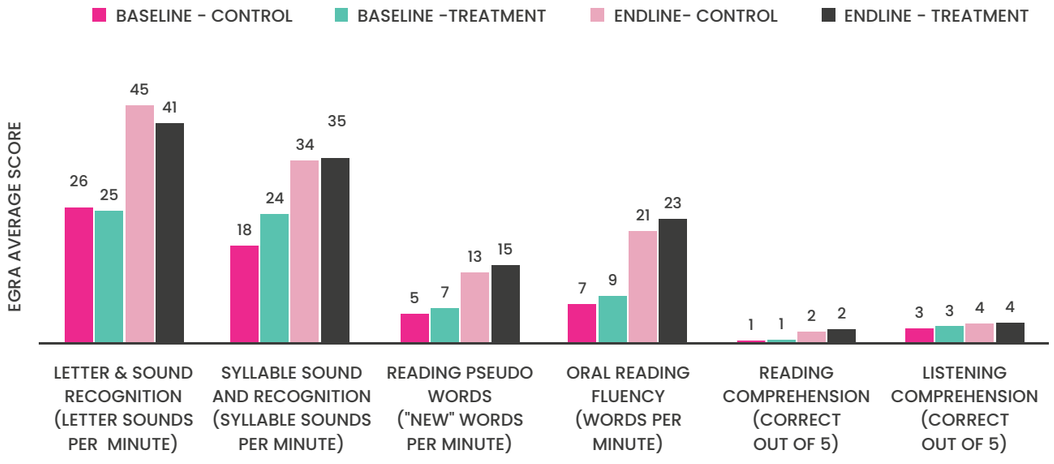
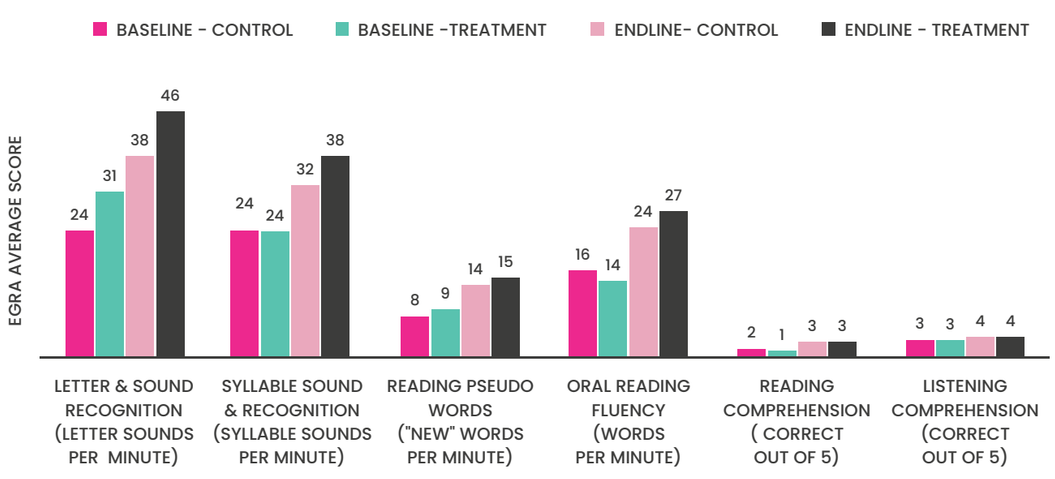


FIGURE 3: GRADE 3 - EGRA AVERAGE SCORE



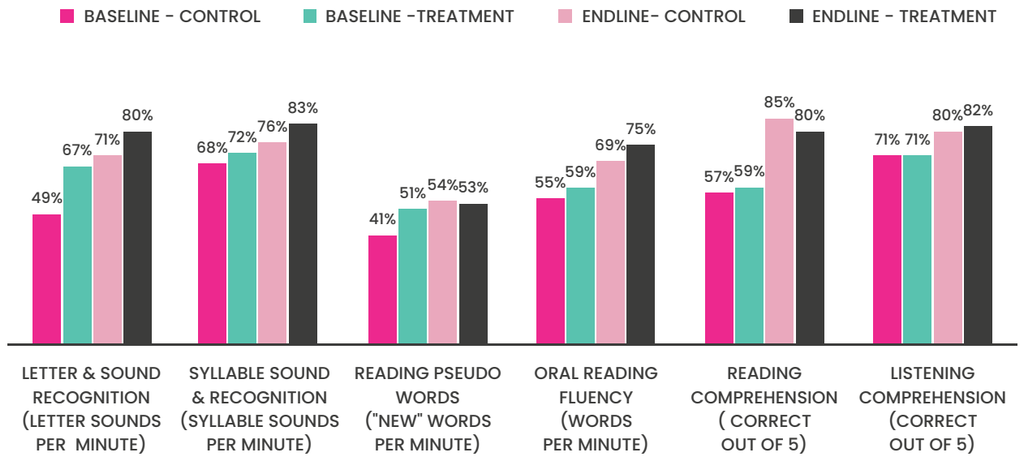
PERCENTAGE OF CORRECT ITEMS FROM ITEMS ATTEMPTED

Another measure used to calculate the change in EGRA score was the percentage of correct items from the number of items attempted. This measure gave different numbers from the average score, since the students might not have had enough time to finish all the items in the section given the allotted time. To obtain this percentage, the number of correct items was divided over the number of items attempted. Figure 4, Figure 5, and Table 7 list the averages of students in grades 2 and 3 in more detail.

After running the regressions, it was found that there was no statistically significant difference between the treatment and control students' percentage correct out of number of items attempted (Table 10).

FIGURE 4: GRADE 2 - PERCENTAGE CORRECT FROM ITEMS ATTEMPTED

FIGURE 5: GRADE 3 - PERCENTAGE CORRECT FROM ITEMS ATTEMPTED



PERCENTAGE OF STUDENTS WITH ZERO SCORES

The percentage of students with a zero score was calculated. A zero score signifies that the student was not able to correctly answer a single question within a subtask. The figures below (Figure 6 and Figure 7) show that a large percentage of students have zero scores in the subtasks, which can depress the overall average scores.

There was no statistically significant difference between the treatment and control students' percentage of zero scores at end-line (Table 11). The results did not change after accounting for control variables.

In general, across treatment and controls schools, the percentage of students with zero scores decreased across most subtasks at endline meaning that students are more capable of correctly answering at least one question in each of the subtasks.

The largest decrease in the percent of students scoring zero (across both treatment and control groups) was in the reading comprehension subtask. At baseline 61% of the grade 2 treatment school students were unable to answer one question correctly, this decreased to only 16% at endline. At baseline 42% of grade 3 treatment students were unable to answer a question correctly, which decreased to 16% at endline.

The percentage of the students scoring zero has decreased for all subtasks except reading invented words for the treatment students of grade 3. The average and percentage scores have improved for this subtask. Therefore, we can conclude that the students who were capable of answering at baseline have improved, and the students who were able to answer this question at baseline still cannot answer it at end-line.

FIGURE 6: GRADE 2 - PERCENTAGE OF STUDENTS WITH ZERO SCORES

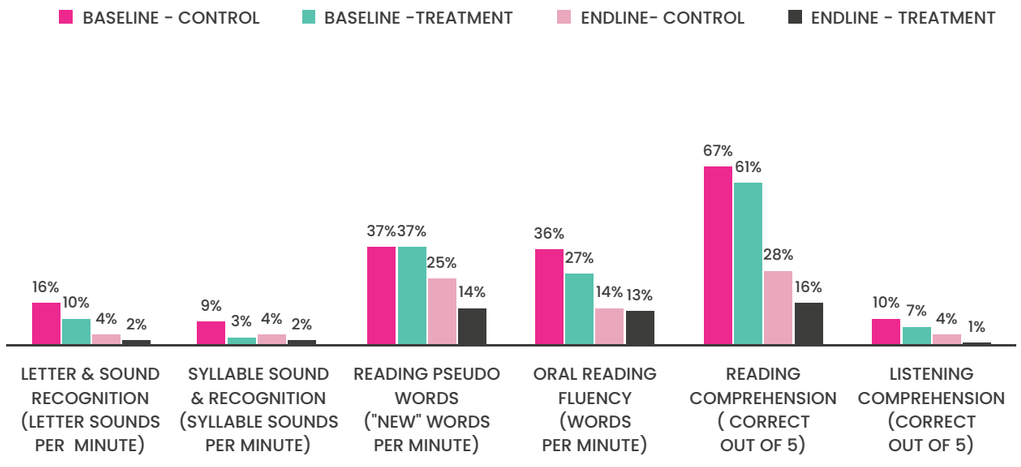
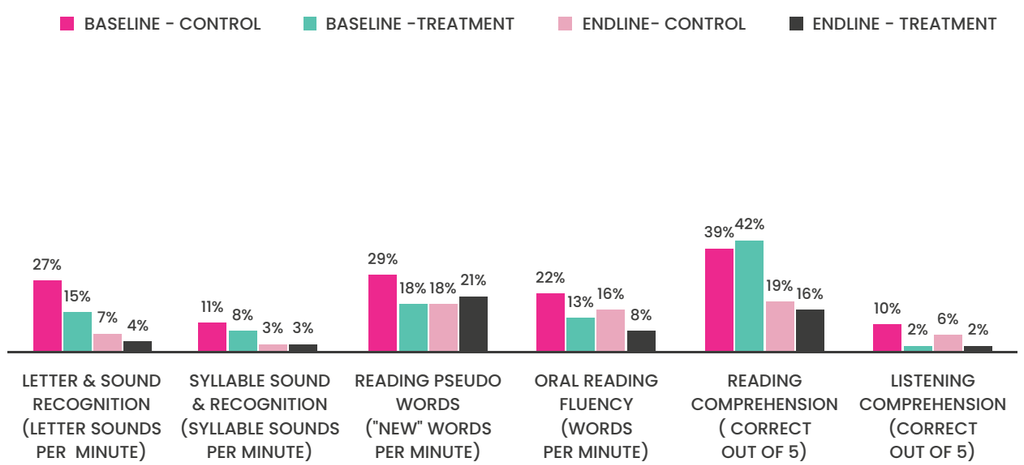


FIGURE 7: GRADE 3 - PERCENTAGE OF STUDENTS WITH ZERO SCORES



COMPARING THE TOP 50 PERCENT VERSUS THE BOTTOM 50

We wanted to see whether the top and/or bottom performing students were able to benefit more from the pilot. To do so, we classified all students who scored above the median score at baseline as top performing students and all students who scored below the median as low performing students. All regressions were run again, once only including top-performing students, and once with the low-performing students. The results suggested that the pilot has not been particularly beneficial to either the top or low performing students.

PART 2: STUDENT SATISFACTION OF ARABIC CLASS

2nd and 3rd grade students were asked a series of questions at baseline and endline regarding their satisfaction of the Arabic language class and their Arabic teachers. At endline, treatment school students were asked additional questions regarding their satisfaction with the morphemes pilot and the workbooks that they had to use. All the questions were close-ended and the student had to choose between a 4 point frequency scale (always, sometimes, rarely, never).

It was found that students were generally satisfied with the pilot, and found some evidence of that the pilot led to grade 3 students being more satisfied with their teacher and Arabic class. However, this result did not hold for 2nd grade students.

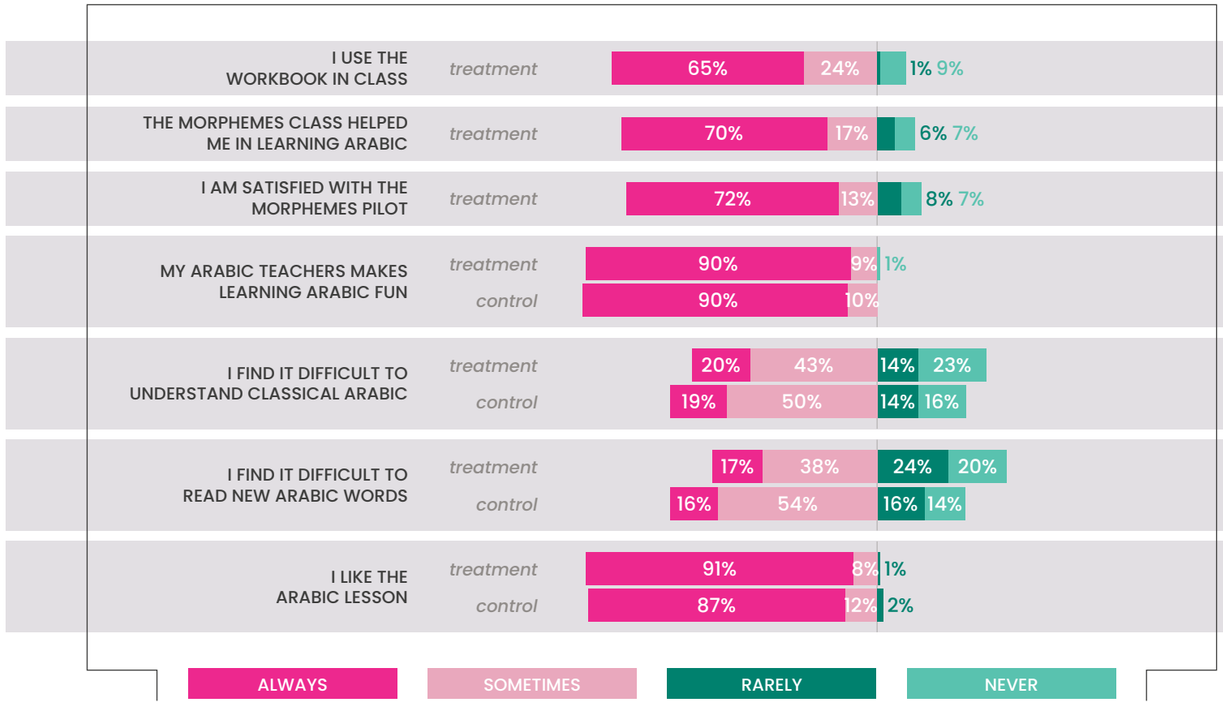
Most 2nd and 3rd grade treatment students were satisfied with the morphemes pilot and believed that it had helped them in reading, writing and learning Arabic (Figure 8, Figure 9, Table 12, and Table 13).

The use and satisfaction with the workbook provided by the pilot was also assessed. More than 60% of students indicated that they always used the workbooks in class. The vast majority of students indicated that the workbooks were clear and easy to follow. However, only about half the students indicated that were satisfied with the workbooks' difficulty level [5]. The majority of students believed that they were always given enough time to finish their workbook.

Multiple regression models were run to see if there was a significant difference in the satisfaction of the students due to the pilot. There were no significant differences in the satisfaction of grade 2 students between treatment and control schools (Table 14).

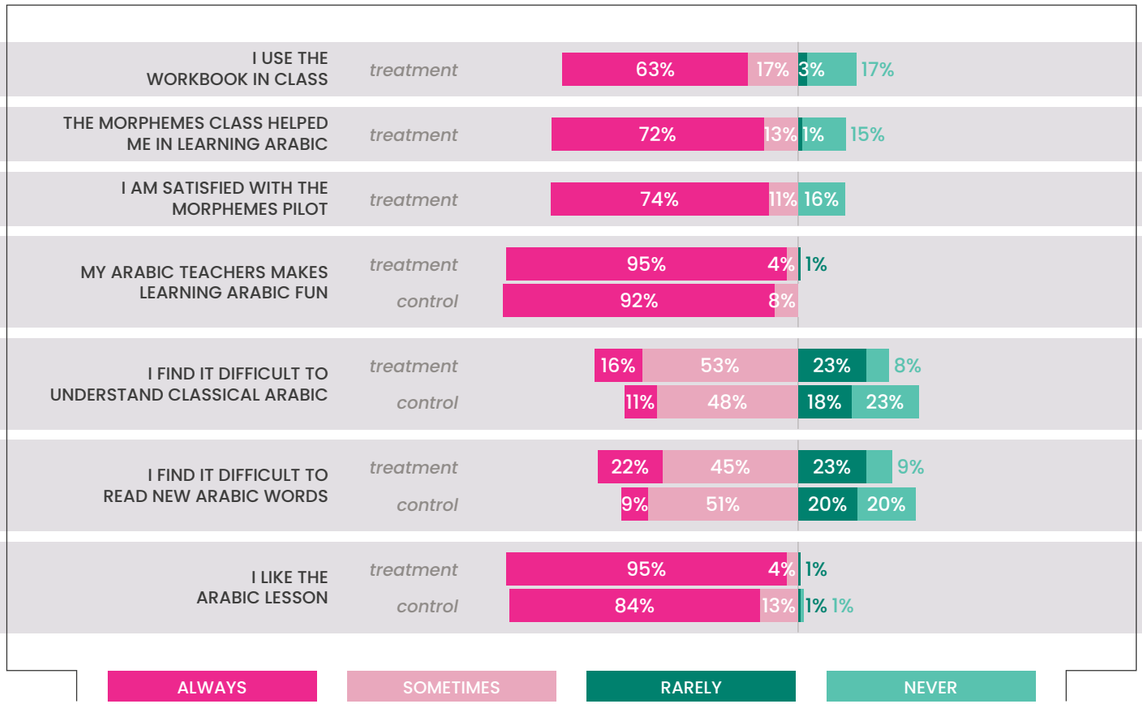
We found some evidence of higher levels of satisfaction in grade 3 students on a few dimensions as a result of the pilot. Grade 3 treatment school students enjoyed reading in Arabic more than the control school students; the difference was significant. Also, grade 3 treatment school students were more satisfied with their Arabic teacher at the end of the year. Based on the perception survey, students reported that more treatment school teachers were explaining the goal of the lesson, and helping students understand various concepts by explaining them differently. The details of the regression can be found in Table 15.

FIGURE 8: GRADE 2 STUDENT SATISFACTION AT ENDLINE



[5] The perception survey did not ask students whether the workbooks' difficulty was too high or low. Therefore, we cannot conclude what difficulty the students perceived the workbooks to be at.

FIGURE 9: GRADE 3 STUDENT SATISFACTION AT ENDLINE



PART 3: STUDENTS' MORPHOLOGICAL AWARENESS

The technical survey consisted of 5 questions that test the student's morphological awareness.

GRADE 2 SURVEY:

QUESTION 1: Understanding the root of the word

QUESTION 2: Converting a sentence from singular to plural form

QUESTION 3: Converting a word from plural to singular form

QUESTION 4: Deriving the root of the word

QUESTION 5: Producing morphological transformation from the root word

GRADE 3 SURVEY:

QUESTION 1: Understanding the root of the word

QUESTION 2: Converting a sentence from plural to singular form

QUESTION 3: Converting a word from singular to 2-person form

QUESTION 4: Deriving the root of the word

QUESTION 5: Producing morphological transformations from the root word

The first three questions were asked at baseline and endline while the last two questions were added at end-line.

The morphemes pilot did not have a significant impact on students' ability to answer these questions (Table 18, Table 19, Table 20, Figures 7 and 8). The only questions where the treatment students outperformed the control students were questions that were covered in detail in the morphemes booklet. Combining these findings with the EGRA results above, we have evidence that the content of the morphemes pilot was covered but that it did not translate into improved literacy outcomes.

COMPARING THE TOP 50 PERCENT VERSUS THE BOTTOM 50

The influence of the pilot on the top or low performing students was investigated, but no differences were found.

FIGURE 10: GRADE 2 PERCENTAGE CORRECT FROM MORPHEMES QUESTIONS

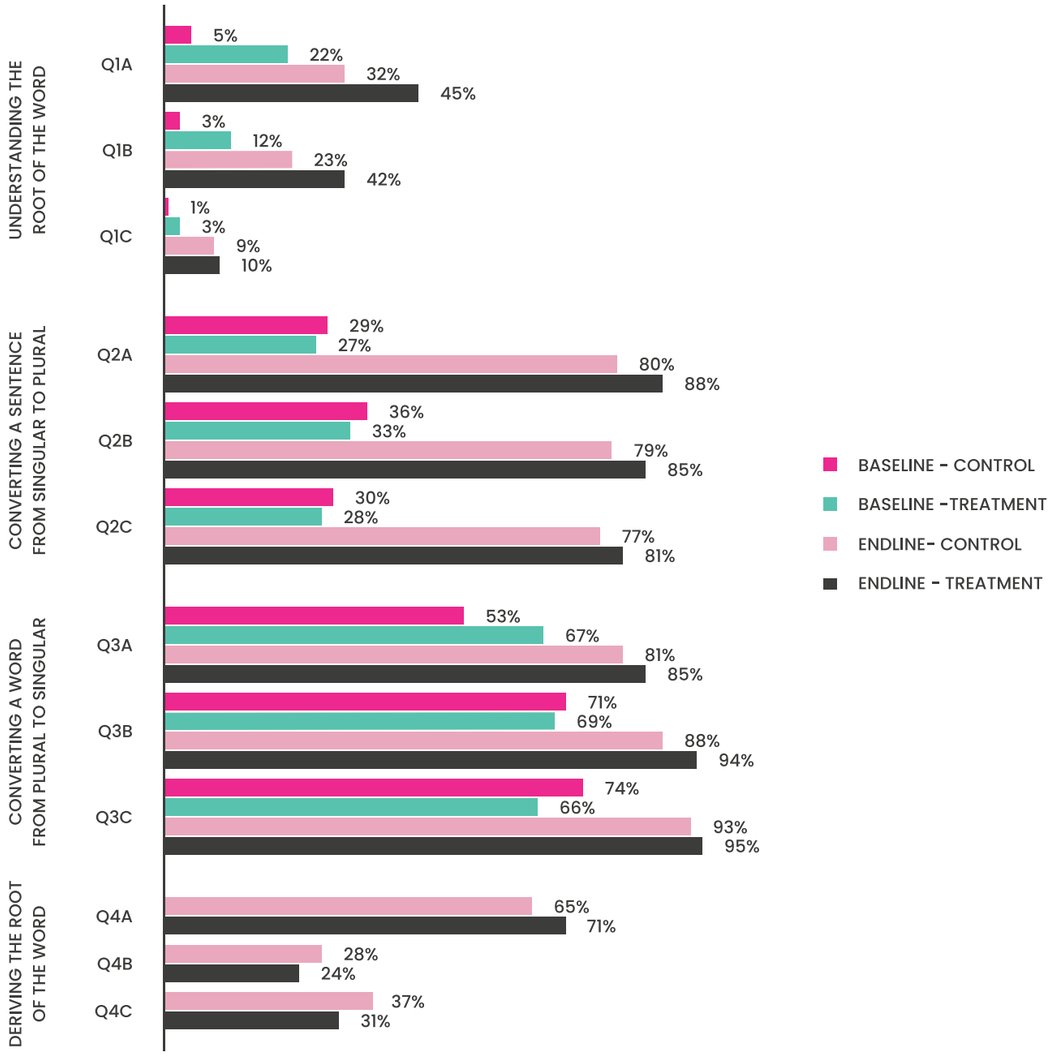
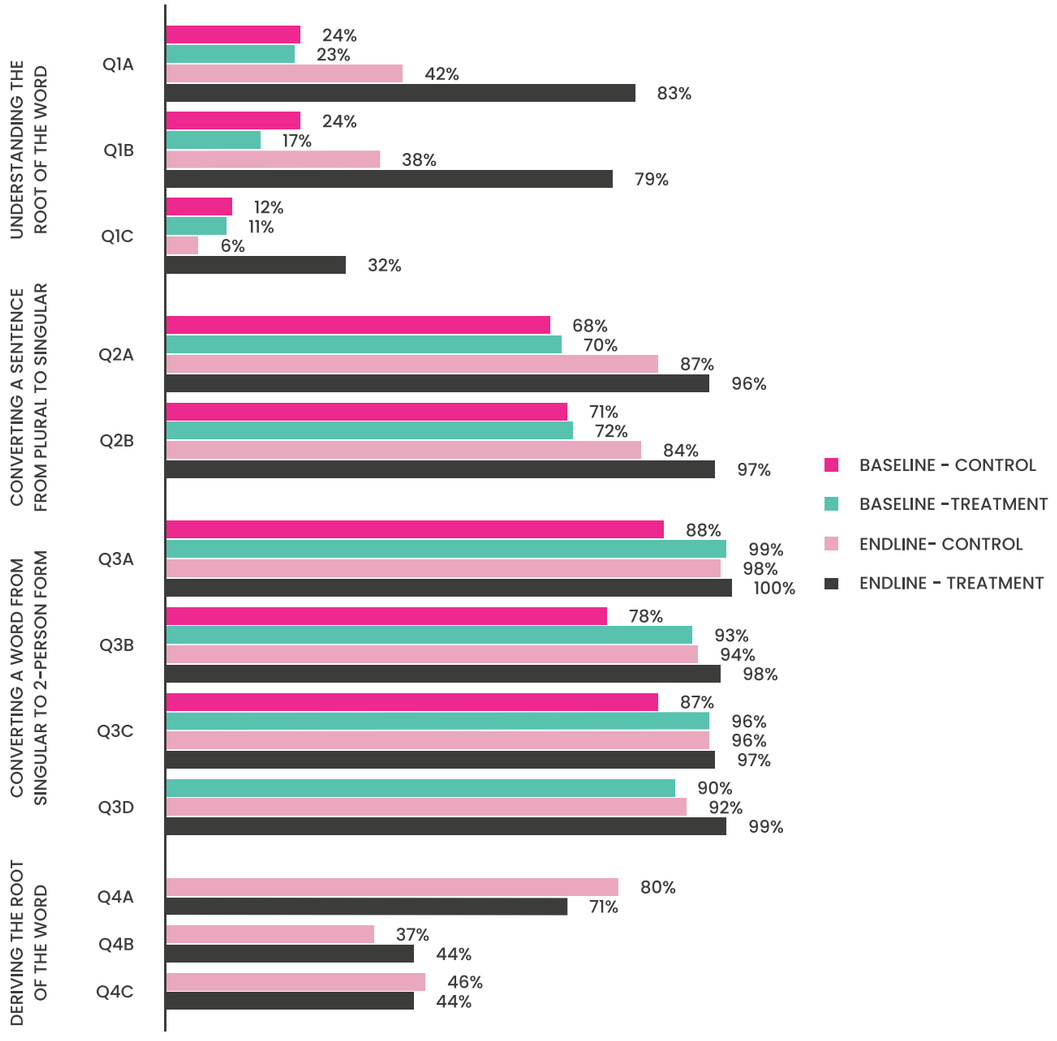


FIGURE 11: GRADE 3 PERCENTAGE CORRECT FROM MORPHEMES QUESTIONS



RESULTS - GRADES 4 AND 5

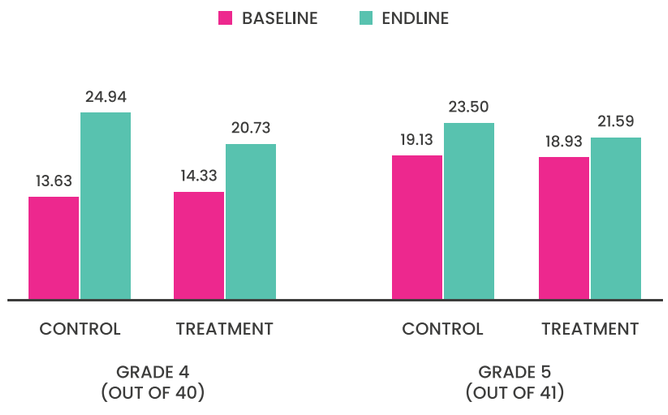
The tests for grade 4 and 5 consisted of 40 multiple choice Arabic tests (41 questions for 5th grade students) which were written by the Ministry of Education, which tested students' reading comprehension, spelling, and grammar abilities.

A satisfaction survey (similar to the one administered to the grade 2 and 3 students) was also administered. Students were given the same test and survey at baseline and endline. At endline, the satisfaction survey had additional questions about the Morphemes pilot. The test and survey were paper administered.

PART 1: MOE TEST

The total score was calculated by summing the number of correct questions. The maximum number is 40 for 4th-grade students and 41 for 5th-grade students. The percentage correct is the total score expressed in a percentage. The percentage zero is the number of students who got a total score of zero; did not answer a single question correctly. The grammar score is the sum of all the grammar questions in the test; the maximum possible score is 34.

FIGURE 12: MEAN OF THE TOTAL SCORE IN THE MOE TEST



The pilot did not have a statistically significant effect on the scores of the grade 4 and 5 students. Both treatment and control students improved throughout the year, (Figure 12) but the improvement in treatment schools was generally no greater than the improvement in control schools. The regression results showed that the pilot did not have a statistically significant effect on the assessment results (Table 23).

COMPARING THE TOP 50 PERCENT VERSUS THE BOTTOM 50

Similar to what was done with the 2nd and 3rd grades, the top and/or bottom performing students were explored in more detail to identify whether they were more able to benefit more from the pilot. To do so, all students who scored above the median score at baseline as top performing students and all students who scored below as low performing students were classified. All the regressions were run again, once including only top-performing students, and once with the low-performing students.

The results of the regression showed that there was no significant difference in achievement, meaning that the pilot had not been particularly beneficial to either the top or bottom performing students.

PART 2: SATISFACTION SURVEY

4th and 5th grade students were asked the same satisfaction questions as the 2nd and 3rd grade students. There were no significant changes in the students' satisfaction with their teacher and the Arabic class (Table 26 and Table 27).

Generally, most of the grade 4 and grade 5 treatment students were satisfied with the morphemes pilot and believed it helped them in reading, writing and learning Arabic (Table 24 and Table 25).

More than half of the students indicated that they always used the workbooks in class. The vast majority of students indicated that the workbooks were clear and easy to follow. However, less than half the students indicated that were satisfied with the workbooks' difficulty level. The majority of students believed they were always given enough time to finish the exercises in the workbook.

FIGURE 13: GRADE 4 STUDENT SATISFACTION AT ENDLINE

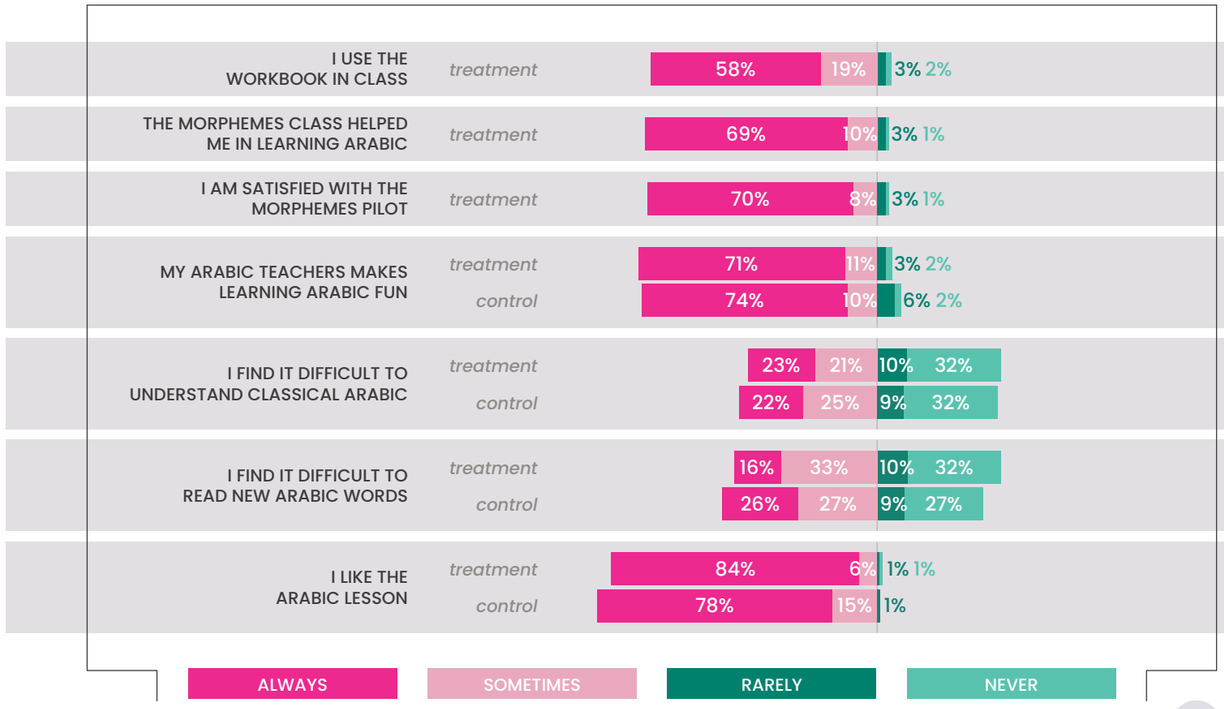
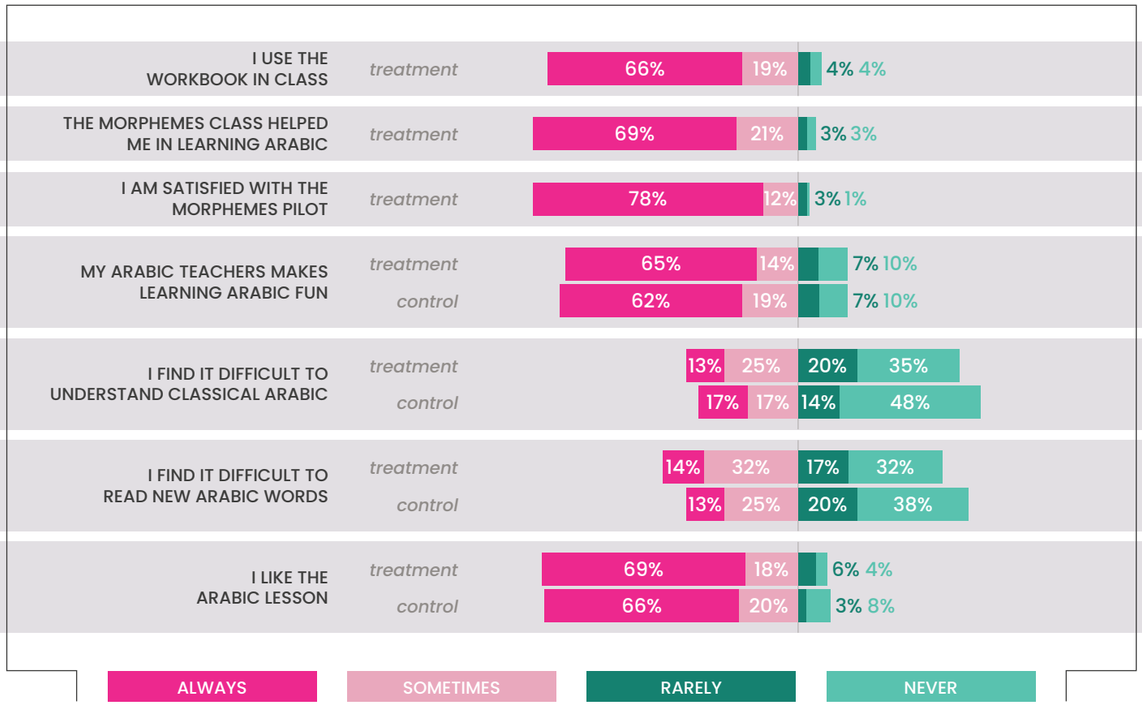


FIGURE 14: GRADE 5 STUDENT SATISFACTION AT ENDLINE



DISCUSSION AND CONCLUSION

This feasibility study aimed to answer the following research questions related to the evidence of promise, feasibility, and scalability:

EVIDENCE OF PROMISE:
To what extent does it appear that the intervention's theory of change holds in this context?

The theory underlying the intervention was that employing explicit instructions of morphemes in teaching would lead to students understanding morphemes, students being able to build words on their own, and students inferring meaning from words using morphemes. The long-term outcomes/impacts expected were improved student comprehension and improved student literacy outcomes.

Based on the results from the **satisfaction survey**, we can state most of our assumptions from the theory of change hold. The majority of students of all four grades were satisfied with the pilot and believed that the Morphemes class had helped them in learning Arabic. The majority of students indicated that they had used the workbooks in class. However, not all students were satisfied with the difficulty level of the workbook. During the design of the pilot, it was assumed that the workbooks were at the right level for students.

On the other hand, based on the results from **the EGRA, morphemes technical survey and MoE tests**, we found that the pilot did not improve the literacy outcomes of the students. Although the ANCOVA regressions showed that there were a small number of significant positive changes, the impact was no different from what would be experienced by chance.

FEASIBILITY:
Can the morphemes awareness pilot be implemented as intended in the Jordanian context?

This question is answered in the qualitative report.

SCALABILITY:
Is the intervention scalable?

Given our findings, we suggest that the pilot should not be scaled at this time. Although the students and teachers indicated that they enjoyed and benefited from the pilot, the students' outcomes did not reflect any learning gains.

LIMITATIONS AND RECOMMENDATIONS

- Although the schools were randomly chosen, the principals had the final word in whether they would like to participate in the pilot, which teachers they would like to nominate for the training, and which class sections were to be included in the pilot. The principals may have chosen the teachers and students that were performing better, which could lead to treatment schools performing better than the control schools. This also reduces the generalizability of the results to the whole of Jordan.

- Due to the small sample and cluster size, our power calculation led to having large MDEs, ones which were much higher than other studies. Therefore, it was highly unlikely that this study will find statistically significant evidence regarding the pilot's impact on literacy.

- At endline, some students who were surveyed at baseline were absent or moved to another section. This reduced our overall sample.

- At baseline, we surveyed the 4th-grade section of School 6 [6], but throughout the pilot, we found that the pilot is not being implemented in that section. At endline, both sections were surveyed the section which received the pilot and the same section we surveyed at baseline. For our analysis, the section in which the pilot was not implemented in was dropped.

- The 5th grade teacher of School 10 was absent for the majority of the second semester. The school was dropped from our analysis since students did not receive the full treatment.

- Some of the students indicated that they were not satisfied with the workbook's difficulty level. If the pilot were rerun in the future, we would recommend exploring how students' perceived the workbooks (too difficult or too easy) and updating them to ensure they are at the correct student level.

- The assessments used may not have been the most accurate measures of the pilot's goals. If the pilot is to be rerun in the future, we would recommend hiring an assessment expert to map out what was being taught and what was measured in the assessments to see if the assessments were capable of capturing the change in literacy.

[6] School names have been excluded for privacy and confidentiality purposes.

BIBLIOGRAPHY

Al Ghanem, R., & Kearns, D. M. (2014).
Orthographic, Phonological, and Morphological Skills and Children's Word Reading in Arabic: A Literature Review. Reading Research Quarterly, 50(1), 83–109.
doi:10.1002/rrq.84

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010).
The Effects of Morphological Instruction on Literacy Skills: A Systematic Review of the Literature. *Review of Educational Research,* 80(2), 144–179.
https://doi.org/10.3102/0034654309359353

Carlisle, J. F. (2010).
Effects of instruction in morphological awareness on literacy achievement: An integrative review. *Reading Research Quarterly*, 45(4), 464–487.
https://doi.org/10.1598/RRQ.45.4.5

Goodwin, A. P., & Ahn, S. (2013).
A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of reading*, 17(4), 257-285.

Hemming, K., & Marsh, J. (2013).
A menu-driven facility for sample-size calculations in cluster randomized controlled trials. *Stata J,* 13(1), 114-35.

Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012).
Children's morphological awareness and reading ability. *Reading and Writing,* 25(2), 389–410.
https://doi.org/10.1007/s11145-010-9276-5

Kirk, C., & Gillon, G. T. (2009).
Integrated morphological awareness intervention as a tool for improving literacy. *Language, Speech, and Hearing Services in Schools,* 40(3), 341–351.
https://doi.org/10.1044/0161-1461(2008/08-0009)

RTI International (2012),
Student Performance in Reading and Mathematics, Pedagogic Practice, and School Management in Jordan, *EdData II Technical and Managerial Assistance, Task Number 16, Contract Number: AID-278-BC-00019*

RTI International (2012),
All Children Learning: Early Grade Learning Brief: Jordan.

Shalhoub-Awwad, Y., & Leikin, M. (2016).
The lexical status of the root in processing morphologically complex words in Arabic. *Scientific Studies of Reading,* 20(4), 296-310.

Sprenger-Charolles, L. (2008).
Early Grade Reading Assessment (EGRA) results from Senegalese primary school students learning to read in French and in Wolof - Report for the World Bank, (February).

Taha, H., & Saiegh-haddad, E. (n.d.).
The Role of Phonological versus Morphological Skills in the Development of Arabic Spelling: An Intervention Study.
https://doi.org/10.1007/s10936-015-9362-6

USAID. (2012).
Early Grade Learning Brief: JORDAN.

Whitehurst, G. J., & Lonigan, C. J. (2001).
Emergent literacy: Development from prereaders to readers. *Handbook of early literacy research*, 1, 11-29.

TABLE 1: EGRA AVERAGE SCORE

	Grade 2 - BL		Grade 2 - EL		Grade 3 - BL		Grade 3 - EL	
Subtask	**Control Mean (SD)**	**Treatment Mean (SD)**	**Control Mean (SD)**	**Treatment Mean (SD)**	**Control Mean (SD)**	**Treatment Mean (SD)**	**Control Mean (SD)**	**Treatment Mean (SD)**
Number of Students	**100**	**67**	**113**	**86**	**79**	**89**	**90**	**95**
Letter and Sound Recognition (letter sounds per minute)	25.51 (17.23)	24.91 (21.82)	44.77 (22.46)	41.28 (20.60)	23.95 (22.38)	31.13 (21.22)	37.80 (22.33)	46.08 (22.67)
Test of equality between treatment arms	0.85		0.22		0.10*		0.03 **	
Syllable Sound and Recognition (syllable sounds per minute)	18.24 (12.75)	24.39 (17.64)	34.26 (20.89)	34.64 (14.41)	23.87 (13.65)	23.64 (14.89)	32.44 (15.78)	37.85 (18.24)
Test of equality between treatment arms	0.09		0.91		0.95		0.18	
Reading Pseudo Words ("new" words per minute)	5.46 (6.06)	6.63 (7.71)	13.33 (12.33)	14.71 (9.53)	7.8 (7.18)	9.04 (7.29)	13.68 (9.97)	15.02 (11.77)
Test of equality between treatment arms	0.16		0.55		0.45		0.7	
Oral Reading Fluency (words per minute)	7.35 (9.08)	8.87 (8.21)	21.08 (13.96)	23.3 (14.10)	16.37 (13.99)	14.27 (12.41)	24.47 (14.51)	27.45 (14.10)
Test of equality between treatment arms	0.15		0.47		0.42		0.41	
Reading Comprehension (Correct out of 5)	0.57 (1.00)	0.69 (1.03)	2.13 (1.80)	2.64 (1.78)	1.57 (1.61)	1.37 (1.53)	3 (1.88)	2.96 (1.87)
Test of equality between treatment arms	0.28		0.16		0.52		0.92	
Listening Comprehension (Correct out of 5)	2.91 1.53	3.13 1.39	3.73 1.48	3.88 1.22	3.38 1.6	3.27 1.38	3.74 1.43	3.92 1.24
Test of equality between treatment arms	0.47		0.61		0.74		0.66	

Notes:

All tests of equality are based on OLS regressions including state fixed effects. Standard errors are clustered at the school level.

*Stars indicate levels of significance of the differences between treatment arms: ***=1%, **=5%, *=10%.*

TABLE 2: EGRA PERCENTAGE CORRECT OUT OF ATTEMPTED

	Grade 2 - BL		Grade 2 - EL		Grade 3 - BL		Grade 3 - EL	
Subtask	Control Mean (SD)	Treatment Mean (SD)	Control Mean (SD)	Treatment Mean (SD)	Control Mean (SD)	Treatment Mean (SD)	Control Mean (SD)	Treatment Mean (SD)
Number of Students	100	67	113	86	79	89	90	95
Letter and Sound Recognition (letter sounds per minute)	64% (0.35)	58% (0.37)	81% (0.26)	79% (0.29)	49% (0.39)	67% (0.37)	71% (0.34)	80% (0.27)
Test of equality between treatment arms	0.35		0.73		0.01*		0.21	
Syllable Sound and Recognition (syllable sounds per minute)	63% (0.30)	72% (0.25)	76% (0.29)	83% (0.23)	68% (0.31)	72% (0.31)	76% (0.27)	83% (0.24)
Test of equality between treatment arms	0.05**		0.2		0.4		0.24	
Reading Pseudo Words ("new" words per minute)	33% (0.33)	37% (0.33)	50% (0.36)	61% (0.31)	41% (0.33)	51% (0.33)	54% (0.35)	53% (0.35)
Test of equality between treatment arms	0.46		0.10*		0.09		0.94	
Oral Reading Fluency (words per minute)	36% (0.34)	44% (0.32)	65% (0.35)	68% (0.33)	55% (0.36)	59% (0.35)	69% (0.33)	75% (0.30)
Test of equality between treatment arms	0.15		0.79		0.46		0.48	
Reading Comprehension (Correct out of 5)	37% (0.44)	42% (0.46)	71% (0.40)	81% (0.29)	57% (0.40)	59% (0.44)	85% (0.26)	80% (0.31)
Test of equality between treatment arms	0.48		0.22		0.71		0.46	
Listening Comprehension (Correct out of 5)	65% (0.29)	70% (0.27)	80% (0.26)	84% (0.22)	71% (0.31)	71% (0.25)	80% (0.26)	82% (0.22)
Test of equality between treatment arms	0.45		0.38		0.99		0.71	

Notes:

All tests of equality are based on OLS regressions including state fixed effects. Standard errors are clustered at the school level.

*Stars indicate levels of significance of the differences between treatment arms: ***=1%, **=5%, *=10%.*

TABLE 3: EGRA PERCENTAGE OF STUDENTS WITH ZERO SCORE

	Grade 2 - BL		Grade 2 - EL		Grade 3 - BL		Grade 3 - EL	
Subtask	Control	Treatment	Control	Treatment	Control	Treatment	Control	Treatment
Number of Students	100	67	113	86	79	89	90	95
Letter and Sound Recognition (% of students)	16%	10%	4%	2%	27%	15%	7%	4%
Test of equality between treatment arms	0.43		0.43		0.003***		0.49	
Syllable Sound and Recognition (% of students)	9%	3%	4%	2%	11%	8%	3%	3%
Test of equality between treatment arms	0.12		0.71		0.39		0.96	
Reading Pseudo Words (% of students)	37%	37%	25%	14%	29%	18%	18%	21%
Test of equality between treatment arms	0.96		0.12		0.22		0.74	
Oral Reading Fluency (% of students)	36%	27%	14%	13%	22%	13%	16%	8%
Test of equality between treatment arms	0.24		0.85		0.13		0.25	
ReadingComprehension (% of students)	67%	61%	28%	16%	39%	42%	19%	16%
Test of equality between treatment arms	0.22		0.24		0.7		0.58	
Listening Comprehension (% of students)	10%	7%	4%	1%	10%	2%	6%	2%
Test of equality between treatment arms	0.52		0.39		0.03**		0.31	

Notes:

All tests of equality are based on OLS regressions including state fixed effects. Standard errors are clustered at the school level.

*Stars indicate levels of significance of the differences between treatment arms: ***=1%, **=5%, *=10%.*

TABLE 4: AVERAGE EGRA SCORE – ANCOVA REGRESSION

		Treatment Status (No Controls)		Treatment Status (Controls)		N
Grade 2	Letter and Sound Recognition	-6.365*	(2.013)	-4.265	(3.011)	167
	Syllable Sound and Recognition	-2.902	(3.833)	-2.81	(3.703)	167
	Reading Pseudo Words	0.367	(2.351)	0.122	(2.352)	167
	Oral Reading Fluency	0.98	(3.175)	0.534	(2.978)	167
	Reading Comprehension	0.294	(0.341)	0.263	(0.286)	167
	Listening Comprehension	0.0743	(0.269)	0.11	(0.322)	167
Grade 3	Letter and Sound Recognition	5.306	(3.046)	3.251	(3.071)	166
	Syllable Sound and Recognition	5.246	(3.346)	4.403	(3.438)	166
	Reading Pseudo Words	0.284	(3.425)	-0.607	(3.402)	166
	Oral Reading Fluency	4.835	(3.245)	4.649	(3.226)	166
	Reading Comprehension	0.117	(0.397)	0.073	(0.392)	166
	Listening Comprehension	0.116	(0.343)	0.0165	(0.332)	166

*Standard errors in parentheses Stars indicate levels of significance: ***=1%, **=5%, *=10%.*

TABLE 5: PERCENTAGE CORRECT OUT OF ATTEMPTED – ANCOVA REGRESSION

		Treatment Status (No Controls)		Treatment Status (Controls)		N
Grade 2	Letter and Sound Recognition	-0.0446	(0.0465)	-0.0328	(0.0549)	165
	Syllable Sound and Recognition	0.0327	(0.0396)	0.0359	(0.039)	167
	Reading Pseudo Words	0.088	(0.0675)	0.0738	(0.0637)	166
	Oral Reading Fluency	-0.0176	(0.0881)	-0.039	(0.085)	167
	Reading Comprehension	0.000858	(0.0790)	-0.0579	(0.0669)	113
	Listening Comprehension	0.0359	(0.0341)	0.0388	(0.0403)	164
Grade 3	Letter and Sound Recognition	0.0412	(0.0584)	0.0109	(0.0582)	166
	Syllable Sound and Recognition	0.0493	(0.0543)	0.0425	(0.0555)	166
	Reading Pseudo Words	-0.0504	(0.1120)	-0.0742	(0.11)	166
	Oral Reading Fluency	0.0497	(0.0820)	0.0351	(0.0795)	163
	Reading Comprehension	-0.0849	(0.0771)	-0.0771	(0.0735)	126
	Listening Comprehension	0.0113	(0.0522)	-0.00328	(0.0485)	162

*Standard errors in parentheses Stars indicate levels of significance: ***=1%, **=5%, *=10%.*

TABLE 6: PERCENTAGE OF STUDENTS WITH ZERO SCORE – ANCOVA REGRESSION

		Treatment Status (No Controls)		Treatment Status (Controls)		N
Grade 2	Letter and Sound Recognition	-0.00209	(0.0197)	-0.00802	(0.0199)	167
	Syllable Sound and Recognition	-0.0109	(0.0237)	-0.0174	(0.0237)	167
	Reading Pseudo Words	-0.101	(0.0709)	-0.106	(0.0599)	167
	Oral Reading Fluency	0.0259	(0.0796)	0.0253	(0.0784)	167
	Reading Comprehension	-0.0919	(0.1110)	-0.0576	(0.107)	167
	Listening Comprehension	-0.0324	(0.0362)	-0.0333	(0.0404)	167
Grade 3	Letter and Sound Recognition	0.00945	(0.0266)	0.0312	(0.0235)	166
	Syllable Sound and Recognition	-0.000789	(0.0313)	0.00572	(0.0265)	166
	Reading Pseudo Words	0.0439	(0.0997)	0.0719	(0.0973)	166
	Oral Reading Fluency	-0.0841	(0.0476)	-0.0902	(0.0464)	166
	Reading Comprehension	-0.0383	(0.0552)	-0.0404	(0.0616)	166
	Listening Comprehension	0.00613	(0.0155)	0.00711	(0.0160)	166

*Standard errors in parentheses Stars indicate levels of significance: ***=1%, **=5%, *=10%.*

TABLE 7: GRADE 2 SATISFACTION SURVEY RESPONSES

	Baseline - Control				Baseline - Treatment				Endline - Control				Endline - Treatment			
Scale*	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
I like reading in Arabic	77	20	2	1	82	18	0	0	84	14	2	0	90	8	2	0
I like the Arabic Lesson	83	14	2	1	88	10	1	0	87	12	2	0	91	8	1	0
I find it difficult to read new Arabic words	14	54	21	11	12	48	12	28	16	54	16	14	17	38	24	20
I find it difficult to understand classical Arabic	15	58	13	14	15	48	18	19	19	50	14	16	20	43	14	23
My Arabic teachers make learning Arabic fun	84	15	1	0	91	4	3	1	90	10	0	0	90	9	0	1
My Arabic teacher helps me in understanding the meaning of new words	70	25	3	2	73	27	0	0	81	15	2	3	80	16	1	2
My Arabic teacher explains my mistakes	60	37	2	1	67	24	3	6	81	17	2	0	72	23	3	1
My Arabic teacher explains to me how I can do better	60	35	3	2	75	18	4	3	81	14	3	2	83	14	2	1
My Arabic teacher uses examples that I can easily understand	66	25	7	2	70	30	0	0	81	17	3	0	83	13	3	1
My Arabic teacher explains the goals of different lessons	56	36	3	5	79	19	0	1	80	19	2	0	85	13	2	0
When I do not understand a particular concept, my Arabic teacher explains it in a different way	62	33	0	5	73	22	1	3	76	21	2	1	80	17	1	1
My Arabic teacher gives us the opportunity to participate in the class	77	22	1	0	82	18	0	0	88	11	1	0	88	10	1	0
Ilike participating in the Arabic class	88	10	1	1	88	9	3	0	95	5	0	0	94	5	1	0
In general, my grades in school are higher than most of my colleagues	53	42	4	1	52	40	7	0	40	50	10	1	43	47	9	1
I want to get the top marks in my class	89	11	0	0	87	12	1	0	90	10	0	0	93	7	0	0
I am satisfied with the morphemes pilot													72	13	8	7
I applied what I've learned when reading or writing													71	16	6	7
The morphemes class helped me in learning Arabic													70	17	6	7
I used the workbook in class													65	24	1	9
The workbooks were clear													83	10	1	6
I was given enough time to complete the workbook													70	21	2	7
I am satisfied with the workbook's difficulty level													50	35	7	8

Numbers are percentages, Scale 1=Always, 2=Sometimes, 3=Rarely, 4 Never

TABLE 8: GRADE 3 SATISFACTION SURVEY RESPONSES

	Baseline - Control				Baseline - Treatment				Endline - Control				Endline - Treatment			
Scale*	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
I like reading in Arabic	85	10	1	3	87	12	0	1	81	18	1	0	96	4	0	0
I like the Arabic Lesson	81	14	3	1	89	9	0	2	84	13	1	1	95	4	1	0
I find it difficult to read new Arabic words	16	53	13	16	16	66	9	9	9	51	20	20	22	45	23	9
I find it difficult to understand classical Arabic	15	52	11	20	15	63	11	11	11	48	18	23	16	53	23	8
My Arabic teachers make learning Arabic fun	82	14	1	1	90	10	0	0	92	8	0	0	95	4	1	0
My Arabic teacher helps me in understanding the meaning of new words	80	19	0	0	79	21	0	0	78	20	2	0	85	13	2	0
My Arabic teacher explains my mistakes	77	18	0	4	76	22	0	1	76	19	3	2	81	16	2	1
My Arabic teacher explains to me how I can do better	71	24	3	1	64	34	1	1	73	26	1	0	83	13	2	2
My Arabic teacher uses examples that I can easily understand	75	22	1	1	80	18	2	0	73	24	2	0	85	14	1	0
My Arabic teacher explains the goals of different lessons	77	19	3	0	71	27	1	1	67	31	2	0	85	12	3	0
When I do not understand a particular concept, my Arabic teacher explains it in a different way	71	23	5	0	80	20	0	0	62	30	6	2	84	14	1	1
My Arabic teacher gives us the opportunity to participate in the class	71	28	0	0	85	15	0	0	79	21	0	0	93	7	0	0
Ilike participating in the Arabic class	89	8	1	1	93	6	0	1	92	8	0	0	99	1	0	0
In general, my grades in school are higher than most of my colleagues	54	39	5	0	53	39	8	0	47	42	10	1	35	52	13	1
I want to get the top marks in my class	85	13	1	0	92	6	1	1	96	4	0	0	94	6	0	0
I am satisfied with the morphemes pilot													74	11	0	16
I applied what I've learned when reading or writing													67	16	1	16
The morphemes class helped me in learning Arabic													72	13	1	15
I used the workbook in class													63	17	3	17
The workbooks were clear													73	12	3	13
I was given enough time to complete the workbook													58	28	1	13
I am satisfied with the workbook's difficulty level													56	31	1	13

Numbers are percentages, Scale 1=Always, 2=Sometimes, 3=Rarely, 4 Never

TABLE 9: GRADE 2 ANCOVA SATISFACTION QUESTIONS

	Treatment Status (No Controls)		Treatment Status (Controls)		N
I like reading in Arabic	0.00755	(0.102)	0.00763	(0.0895)	167
I like the Arabic Lesson	0.00438	(0.0918)	-0.0119	(0.0721)	167
I find it difficult to read new Arabic words	0.222	(0.293)	0.24	(0.305)	167
I find it difficult to understand classical Arabic	0.142	(0.264)	0.0814	(0.291)	167
My Arabic teacher makes learning Arabic fun	0.0444	(0.0451)	-0.00871	(0.0447)	167
MyArabic teacher helps me in understanding the meaning ofnew words	-0.0205	(0.156)	-0.0488	(0.151)	167
My Arabic teacher explains my mistakes	0.0972	(0.104)	0.0572	(0.0904)	167
My Arabic teacher explains to me how I can do better	-0.0362	(0.121)	-0.0293	(0.109)	167
My Arabic teacher uses examples that I can easily understand	0.0376	(0.087)	0.0185	(0.0769)	167
My Arabic teacher explains the goals of different lessons	-0.0431	(0.103)	-0.0378	(0.0945)	167
When I do not understand a particular concept, my Arabic teacher explainsit in a different way	-0.0803	(0.0814)	-0.0921	(0.0839)	167
My Arabic teacher gives us the opportunity to participate in the class	0.00892	(0.0626)	-0.0226	(0.0719)	167
Ilike participating in the Arabic class	-0.00512	(0.0296)	-0.029	(0.0171)	167
In general, my grades in school are higher than most of my colleagues.	-0.0564	(0.062)	-0.119	(0.0964)	167
I want to get the top marks in my class	-0.0178	(0.0545)	-0.0314	(0.0619)	167

Scale 1=Always, 2=Sometimes, 3=Rarely, 4 Never

*Standard errors in parentheses Stars indicate levels of significance: ***=1%, **=5%, *=10%.*

TABLE 10: GRADE 3 ANCOVA SATISFACTION QUESTIONS

	Treatment Status (No Controls)		Treatment Status (Controls)		N
I like reading in Arabic	-0.153*	(0.0577)	-0.172*	(0.0609)	166
I like the Arabic Lesson	-0.0519	(0.0525)	-0.0403	(0.0417)	166
I find it difficult to read new Arabic words	-0.251	(0.194)	-0.164	(0.231)	166
I find it difficult to understand classical Arabic	-0.231	(0.205)	-0.194	(0.214)	166
My Arabic teacher makes learning Arabic fun	0.011	(0.0461)	0.0123	(0.0477)	166
My Arabic teacher helps me in understanding the meaning of new words	-0.0739	(0.0995)	-0.0614	(0.0799)	166
My Arabic teacher explains my mistakes	-0.0708	(0.101)	-0.0573	(0.109)	166
My Arabic teacher explains to me how I can do better	-0.00171	(0.0942)	-0.0136	(0.0907)	166
My Arabic teacher uses examples that I can easily understand	-0.113	(0.0588)	-0.112	(0.0623)	166
My Arabic teacher explains the goals of different lessons	-0.189*	(0.0803)	-0.181*	(0.0736)	166
When I do not understand a particular concept, my Arabic teacher explainsit in a different way	-0.265*	(0.0926)	-0.224*	(0.0978)	166
My Arabic teacher gives us the opportunity to participate in the class	-0.119*	(0.049)	-0.103	(0.051)	166
Ilike participating in the Arabic class	-0.0623	(0.0389)	-0.0613	(0.0462)	166
In general, my grades in school are higher than most of my colleagues.	0.139	(0.104)	0.142	(0.122)	166
I want to get the top marks in my class	0.0143	(0.0434)	0.0197	(0.04)	166

Scale 1=Always, 2=Sometimes, 3=Rarely, 4 Never

*Standard errors in parentheses Stars indicate levels of significance: ***=1%, **=5%, *=10%.*

TABLE 11: GRADE 2 TECHNICAL QUESTIONS

	Baseline - Control			Baseline - Treatment			Endline - Control			Endline - Treatment		
N	100			67			113			86		
	A	B	C	A	B	C	A	B	C	A	B	C
q1a	5%	52%	43%	22%	36%	42%	32%	48%	20%	45%	34%	21%
q1b	3%	45%	52%	12%	48%	40%	23%	55%	22%	42%	36%	22%
q1c	1%	44%	55%	3%	58%	39%	9%	64%	27%	10%	65%	24%
q2a	29%	40%	31%	27%	49%	24%	80%	16%	4%	88%	8%	3%
q2b	36%	42%	22%	33%	46%	21%	79%	19%	3%	85%	9%	6%
q2c	30%	43%	27%	28%	48%	24%	77%	19%	4%	81%	15%	3%
q3a	53%	33%	14%	67%	30%	3%	81%	17%	2%	85%	14%	1%
q3b	71%	19%	10%	69%	28%	3%	88%	10%	2%	94%	6%	0%
q3c	74%	17%	9%	66%	30%	4%	93%	6%	1%	95%	5%	0%
q4a							65%	26%	9%	71%	21%	8%
q4b							28%	58%	14%	24%	64%	12%
q4c							37%	51%	12%	31%	57%	12%

A=Correct, B=Incorrect, C=No Answer/Did not know

TABLE 12: GRADE 3 TECHNICAL QUESTIONS

	Baseline - Control			Baseline - Treatment			Endline - Control			Endline - Treatment		
N	68			71			90			95		
	A	B	C	A	B	C	A	B	C	A	B	C
q1a	24%	21%	54%	23%	31%	46%	42%	34%	23%	83%	6%	11%
q1b	24%	16%	59%	17%	37%	46%	38%	36%	27%	79%	12%	9%
q1c	12%	22%	65%	11%	31%	58%	6%	68%	27%	32%	58%	11%
q2a	68%	21%	10%	70%	17%	13%	87%	9%	4%	96%	3%	1%
q2b	71%	18%	10%	72%	15%	13%	84%	11%	4%	97%	2%	1%
q2c	88%	9%	1%	99%	1%	0%	98%	1%	1%	100%	0%	0%
q3a	78%	16%	4%	93%	7%	0%	94%	4%	1%	98%	2%	0%
q3b	87%	9%	3%	96%	4%	0%	96%	3%	1%	97%	3%	0%
q3c	77%	19%	4%	90%	8%	1%	92%	6%	2%	99%	1%	0%
q4a							80%	17%	3%	71%	22%	7%
q4b							37%	56%	8%	44%	48%	7%
q4c							46%	47%	8%	44%	48%	7%

A=Correct, B=Incorrect, C=No Answer/Did not know

TABLE 13: GRADE 2 AND 3 MORPHEMES QUESTION 5

		Grade 2			Grade 3		
		q5a	q5b	q5c	q5a	q5b	q5c
Endline Control	One word	26%	29%	26%	22%	23%	24%
	Two words	22%	16%	17%	27%	34%	34%
	Three words	5%	4%	4%	20%	14%	13%
	Four words	1%	0%	0%	4%	2%	1%
	Five words	1%	0%	0%	1%	0%	0%
	Incorrect Answer	20%	27%	27%	12%	14%	14%
	Did not know	25%	23%	27%	13%	11%	12%
Endline Treatment	One word	15%	20%	28%	17%	26%	23%
	Two words	31%	35%	33%	25%	24%	25%
	Three words	17%	14%	8%	21%	22%	24%
	Four words	6%	0%	0%	15%	7%	5%
	Five words	1%	0%	0%	3%	0%	0%
	Incorrect Answer	13%	17%	17%	13%	12%	14%
	Did not know	16%	14%	14%	6%	8%	8%

TABLE 14: GRADE 2 ANCOVA MORPHEMES QUESTIONS

	Treatment Status (No Controls)		Treatment Status (Controls)		N
q1a	0.175*	(0.061)	0.176*	(0.0624)	167
q1b	0.192*	(0.0827)	0.196*	(0.0761)	167
q1c	-0.00382	(0.0394)	0.0109	(0.0367)	167
q2a	0.0588	(0.099)	0.0521	(0.0991)	167
q2b	0.0846	(0.0983)	0.104	(0.112)	167
q2c	0.017	(0.0927)	0.0254	(0.106)	167
q3a	-0.0181	(0.0891)	-0.0245	(0.0887)	167
q3b	0.0592	(0.0731)	0.0678	(0.0765)	167
q3c	0.0337	(0.0488)	0.0327	(0.0441)	167

*Standard errors in parentheses Stars indicate levels of significance: ***=1%, **=5%, *=10%.*

TABLE 15: GRADE 3 ANCOVA MORPHEMES QUESTIONS

	Treatment Status (No Controls)		Treatment Status (Controls)		N
q1a	0.348*	(0.111)	0.311*	(0.11)	167
q1b	0.361**	(0.106)	0.334**	(0.0961)	167
q1c	0.214**	(0.0453)	0.218**	(0.0524)	167
q2a	0.0673	(0.0539)	0.0698	(0.0581)	167
q2b	0.093	(0.0478)	0.0883	(0.0445)	167
q3a	0.0323	(0.0173)	0.0375	(0.0248)	167
q3b	0.0275	(0.0226)	0.0285	(0.0266)	167
q3c	0.0211	(0.0291)	0.0318	(0.0282)	167
q3d	0.0867*	(0.0276)	0.0920*	(0.029)	137

*Standard errors in parentheses Stars indicate levels of significance: ***=1%, **=5%, *=10%.*

TABLE 16: RCT REGRESSION FOR THE 4TH MORPHEMES QUESTION

		Treatment Status (No Controls)		Treatment Status (Controls)		N
Grade 2	Q4-A	0.0544	(0.093)	-0.0131	(0.107)	199
	Q4-B	-0.039	(0.0625)	-0.0853	(0.0698)	199
	Q4-C	-0.0577	(0.106)	-0.0484	(0.124)	199
Grade 3	Q4-A	-0.0947	(0.0513)	-0.0683	(0.0595)	185
	Q4-B	0.0754	(0.0969)	0.106	(0.0933)	185
	Q4-C	-0.0135	(0.0712)	-0.013	(0.0729)	185

*Standard errors in parentheses Stars indicate levels of significance: ***=1%, **=5%, *=10%.*

TABLE 17: TOTAL SCORE, PERCENTAGE CORRECT AND ZERO SCORES

	Grade 4 - BL		Grade 4 - EL		Grade 5 - BL		Grade 5 - EL	
	Control Mean (SD)	Treatment Mean (SD)	Control Mean (SD)	Treatment Mean (SD)	Control Mean (SD)	Treatment Mean (SD)	Control Mean (SD)	Treatment Mean (SD)
Number of Students	127	105	132	145	119	123	133	134
Total score	13.63 (7.04)	14.33 (5.97)	24.94 (9.87)	20.73 (8.66)	19.13 (8.26)	18.93 (8.01)	23.5 (8.59)	21.59 (8.38)
Test of equality between treatment arms	0.75		0.33		0.95		0.57	
Percentage correct	34% (0.18)	36% (0.15)	62% (0.25)	52% (0.22)	47% (0.2)	46% (0.2)	57% (0.21)	53% (0.20)
Test of equality between treatment arms	0.75		0.33		0.95		0.57	
Percentage zero	9% (0.29)	0% -	0% -	0% -	4% (0.2)	2% (0.13)	1% (0.09)	0% -
Test of equality between treatment arms	0.31				0.3		0.32	
Grammar total	10.34 (5.62)	10.04 (5.05)	19.67 (8.54)	16.23 (7.18)	15.34 (6.77)	15.07 (6.7)	18.98 (7.38)	17.66 (7.08)
Test of equality between treatment arms	0.87		0.36		0.91		0.62	

Notes:

All tests of equality are based on OLS regressions including state fixed effects. Standard errors are clustered at the school level.

*Stars indicate levels of significance of the differences between treatment arms: ***=1%, **=5%, *=10%.*

TABLE 18: ANCOVA REGRESSION FOR MOE TEST

		Treatment Status (No Controls)		N	Treatment Status (Controls)		N
Grade 4	Total score	-2.156	(2.647)	232	-1.817	(1.52)	242
	Percentage correct	-0.0539	(0.0662)	232	-0.0443	(0.0371)	242
	Percentage of students with zero score	0	(.)	232	0	(.)	242
Grade 5	Total score	-2.421	(1.952)	93	-1.605	(1.598)	158
	Percentage correct	-0.0605	(0.0488)	93	-0.0391	(0.039)	158
	Percentage of students with zero score	0	(.)	93	0	(.)	158

*Stars indicate levels of significance: ***=1%, **=5%, *=10%.*

TABLE 19: GRADE 4 SATISFACTION SURVEY RESPONSES

	Baseline - Control				Baseline - Treatment				Endline - Control				Endline - Treatment			
Scale*	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
I like reading in Arabic	43	14	2	0	39	20	3	0	83	11	2	1	83	11	1	0
I like the Arabic Lesson	43	17	1	1	37	15	4	2	78	15	1	0	84	6	1	1
Understanding Arabic lessons aremore important than memorizing them	24	22	6	5	26	28	5	4	55	27	6	0	53	30	4	2
I can create correct words in classical Arabic	35	16	7	2	30	22	6	4	48	34	3	5	50	28	7	4
I find it difficult to read new Arabic words	17	9	14	19	14	17	21	10	26	27	9	27	16	33	10	32
I find it difficult to understand classical Arabic	13	15	9	19	10	22	15	14	22	25	9	32	23	21	10	32
My Arabic teacher makes learning Arabic fun	41	13	2	2	36	22	1	3	74	10	6	2	71	11	3	2
MyArabic teacher helps me in understanding the meaning ofnew words	43	13	1	0	34	26	5	3	82	4	2	2	74	12	1	1
My Arabic teacher explains my mistakes	24	19	9	5	22	23	11	10	71	11	4	6	61	12	8	7
My Arabic teacher explains to me how I can do better	31	18	2	4	36	22	6	2	67	10	2	1	73	10	0	1
My Arabic teacher uses examples that I can easily understand	31	19	2	1	35	22	5	4	63	16	1	0	63	21	1	0
My Arabic teacher asked us questions to make sure we were following along	38	13	1	2	27	26	9	3	62	11	5	3	66	14	2	2
My Arabic teacher asks us questions to make sure we are thinking about the things that are we read or write	33	17	4	1	30	26	7	2	60	14	5	1	58	21	3	2
My Arabic teacher explains the goals of different lessons	26	17	5	6	30	24	9	5	59	10	7	5	58	17	3	4
When I do not understand a particular concept, my Arabic teacher explainsit in a different way	31	13	7	2	20	25	11	9	51	11	4	9	50	21	6	7
My Arabic teacher gives us the opportunity to participate in the class	35	14	1	2	33	21	7	1	68	8	2	2	69	10	3	1
Ilike participating in the Arabic class	42	10	2	1	38	20	3	2	71	5	0	2	74	7	1	1
In general, my grades in school are higher than most of my colleagues.	36	15	3	0	34	25	4	1	38	35	5	2	46	32	5	1
I want to get the top marks in my class	39	10	2	2	43	11	6	4	69	7	3	0	68	14	1	1

	Baseline - Control				Baseline - Treatment				Endline - Control				Endline - Treatment			
Scale*	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
I am satisfied with the morphemes pilot													70	8	3	1
I applied what I've learned when reading or writing													72	6	3	1
The morphemes class helped me in learning Arabic													69	10	3	1
I used the workbook in class													58	19	3	2
The workbooks were clear													70	7	6	0
I was given enough time to complete the workbook													61	14	2	2
I am satisfied with the workbook's difficulty level													42	27	8	2
The flash cards were clear													60	16	3	3
I enjoyed using the flashcards													61	15	4	1

Numbers are percentages, Scale 1=Always, 2=Sometimes, 3=Rarely, 4 Never

TABLE 20: GRADE 5 SATISFACTION SURVEY QUESTION RESPONSES

	Baseline - Control				Baseline - Treatment				Endline - Control				Endline - Treatment			
Scale*	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
I like reading in Arabic	70	20	0	0	50	25	2	1	71	23	1	3	78	19	1	0
I like the Arabic Lesson	57	29	2	2	41	34	2	1	66	20	3	8	69	18	6	4
Understanding Arabic lessons aremore important than memorizing them	53	29	6	2	47	18	7	2	54	31	9	2	70	17	7	3
I can create correct words in classical Arabic	47	37	4	3	46	24	3	2	56	29	7	5	62	25	4	4
I find it difficult to read new Arabic words	18	24	17	31	10	18	31	18	13	25	20	38	14	32	17	32
I find it difficult to understand classical Arabic	13	15	21	39	11	14	26	24	17	17	14	48	13	25	20	35
My Arabic teacher makes learning Arabic fun	66	18	2	0	41	17	7	3	62	19	7	10	65	14	7	10
MyArabic teacher helps me in understanding the meaning ofnew words	70	16	3	2	49	17	2	2	72	11	8	5	79	14	3	2
My Arabic teacher explains my mistakes	55	18	11	9	44	15	8	3	62	17	4	13	72	13	6	7
My Arabic teacher explains to me how I can do better	67	15	4	3	43	20	3	1	63	23	4	5	66	19	4	5
My Arabic teacher uses examples that I can easily understand	61	21	4	6	41	16	5	6	62	23	4	4	72	19	2	4
My Arabic teacher asked us questions to make sure we were following along	61	21	4	4	41	19	5	2	65	19	8	4	65	16	7	9
My Arabic teacher asks us questions to make sure we are thinking about the things that are we read or write	58	29	4	1	41	18	6	4	59	24	5	4	60	25	6	5
My Arabic teacher explains the goals of different lessons	53	28	5	6	39	20	7	2	62	20	8	4	62	19	10	6
When I do not understand a particular concept, my Arabic teacher explainsit in a different way	51	19	9	11	36	15	11	5	52	21	8	14	55	17	9	14
My Arabic teacher gives us the opportunity to participate in the class	62	22	5	6	48	12	4	5	65	17	8	3	74	16	4	3
Ilike participating in the Arabic class	73	15	2	3	52	15	2	1	68	16	2	6	75	15	3	2
In general, my grades in school are higher than most of my colleagues.	62	26	5	1	46	18	5	2	41	45	2	5	50	39	5	1
I want to get the top marks in my class	77	11	1	4	56	9	3	2	76	9	5	4	90	4	1	0

	Baseline - Control				Baseline - Treatment				Endline - Control				Endline - Treatment			
Scale*	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
I am satisfied with the morphemes pilot													78	12	3	1
I applied what I've learned when reading or writing													65	22	6	2
The morphemes class helped me in learning Arabic													69	21	3	3
I used the workbook in class													66	19	4	4
The workbooks were clear													67	19	4	2
I was given enough time to complete the workbook													66	25	2	2
I am satisfied with the workbook's difficulty level													51	25	7	8
The flash cards were clear													78	11	4	1
I enjoyed using the flashcards													69	25	1	1

Numbers are percentages, Scale 1=Always, 2=Sometimes, 3=Rarely, 4 Never

TABLE 21: GRADE 4 ANCOVA SATISFACTION QUESTIONS

	Treatment Status (No Controls)		N	Treatment Status (Controls)		N
I like reading in Arabic	0.00975	(0.194)	232	-0.0152	(0.044)	93
I like the Arabic Lesson	-0.0533	(0.360)	232	0.12	(0.053)	93
Understanding Arabic lessons aremore important than memorizing them	-0.268	(0.413)	232	-0.179	(0.330)	93
I can create correct words in classical Arabic	-0.157	(0.302)	232	0.0977	(0.263)	93
I find it difficult to read new Arabic words	0.162	(0.392)	232	0.313	(0.245)	93
I find it difficult to understand classical Arabic	0.0778	(0.366)	232	-0.115	(0.296)	93
My Arabic teacher makes learning Arabic fun	0.139	(0.591)	232	0.367	(0.250)	93
MyArabic teacher helps me in understanding the meaning ofnew words	-0.0182	(0.419)	232	0.166	(0.162)	93
My Arabic teacher explains my mistakes	0.126	(0.471)	232	0.0506	(0.129)	93
My Arabic teacher explains to me how I can do better	-0.448	(0.685)	232	0.0777	(0.362)	93
My Arabic teacher uses examples that I can easily understand	-0.416	(0.758)	232	0.0515	(0.385)	93
My Arabic teacher asked us questions to make sure we were following along	-0.393	(0.752)	232	0.0835	(0.340)	93
My Arabic teacher asks us questions to make sure we are thinking about the things that are we read or write	-0.429	(0.807)	232	-0.015	(0.358)	93
My Arabic teacher explains the goals of different lessons	-0.251	(0.710)	232	0.164	(0.327)	93
When I do not understand a particular concept, my Arabic teacher explainsit in a different way	-0.611	(0.846)	232	-0.714*	(0.295)	93
My Arabic teacher gives us the opportunity to participate in the class	-0.268	(0.866)	232	0.223	(0.451)	93
Ilike participating in the Arabic class	-0.511	(0.724)	232	-0.125	(0.379)	93
In general, my grades in school are higher than most of my colleagues.	-0.395	(0.700)	232	-0.011	(0.584)	93
I want to get the top marks in my class	-0.38	(0.745)	232	-0.0151	(0.594)	93

*Standard errors in parentheses Stars indicate levels of significance: ***=1%, **=5%, *=10%.*

TABLE 22: GRADE 5 ANCOVA SATISFACTION QUESTIONS

	Treatment Status (No Controls)		N	Treatment Status (Controls)		N
I like reading in Arabic	-0.114	(0.168)	242	-0.168	(0.127)	158
I like the Arabic Lesson	-0.0617	(0.355)	242	0.133	(0.388)	158
Understanding Arabic lessons aremore important than memorizing them	-0.238	(0.219)	242	-0.289	(0.225)	158
I can create correct words in classical Arabic	-0.075	(0.268)	242	-0.19	(0.143)	158
I find it difficult to read new Arabic words	-0.117	(0.255)	242	-0.359	(0.169)	158
I find it difficult to understand classical Arabic	0.0497	(0.238)	242	-0.23	(0.165)	158
My Arabic teacher makes learning Arabic fun	0.0275	(0.484)	242	0.246	(0.477)	158
MyArabic teacher helps me in understanding the meaning ofnew words	-0.253	(0.294)	242	-0.132	(0.279)	158
My Arabic teacher explains my mistakes	-0.319	(0.309)	242	-0.181	(0.348)	158
My Arabic teacher explains to me how I can do better	-0.0195	(0.315)	242	-0.135	(0.382)	158
My Arabic teacher uses examples that I can easily understand	-0.511	(0.416)	242	-0.353	(0.435)	158
My Arabic teacher asked us questions to make sure we were following along	0.0297	(0.362)	242	-0.102	(0.491)	158
My Arabic teacher asks us questions to make sure we are thinking about the things that are we read or write	-0.0395	(0.348)	242	-0.213	(0.384)	158
My Arabic teacher explains the goals of different lessons	-0.155	(0.302)	242	-0.151	(0.409)	158
When I do not understand a particular concept, my Arabic teacher explainsit in a different way	-0.0714	(0.422)	242	-0.1	(0.514)	158
My Arabic teacher gives us the opportunity to participate in the class	-0.093	(0.299)	242	-0.246	(0.431)	158
Ilike participating in the Arabic class	-0.32	(0.361)	242	-0.611	(0.355)	158
In general, my grades in school are higher than most of my colleagues.	-0.208	(0.193)	242	-0.32	(0.248)	158
I want to get the top marks in my class	-0.434	(0.293)	242	-0.679*	(0.229)	158

*Standard errors in parentheses Stars indicate levels of significance: ***=1%, **=5%, *=10%.*