SYRIA
EDUCATION
PROGRAMME

Syria Education Programme Learning Assessment Report 2021

JANUARY 2022

SCHOOL-TO-SCHOOL INTERNATIONAL

This report has been prepared by School-to-School International for Chemonics International, Inc. This study has been funded by UK Aid from the United Kingdom Government's Foreign, Commonwealth and Development Office; however, the views expressed do not necessarily reflect the government's official policies.

Disclaimer

This document has been redacted to protect the individuals involved in the Syria Education Programme. All names of people and locations have either been altered or removed, as has any information that may identify people or locations.

Project Description

The Syria Education Programme (SEP), also known as Manahel, provides access to safe, inclusive, and quality learning opportunities. Across its lifecycle the project will reach half a million primary-school-age children in Syria.

SEP enables teachers, school staff, and education sector leadership to deliver quality education. In response to the ever-changing landscape of conflict and crisis in Syria, SEP invests in and applies research to respond to the educational, psychological, and protection needs of Syria's children.

From the specialised requirements of disabled children to the psychological demands of childhood within conflict, students' needs are as diverse as they are urgent. SEP takes a broad and nuanced approach to the myriad needs of individual children and groups. By broadening educational access, promoting a safe and secure environment, and creating quality learning opportunities, SEP strives to meet children's holistic needs at scale.

UKaid from the British people

The Syria Education Programme is funded by UK aid from the UK government.

Acknowledgments

The authors would like to acknowledge the many individuals who supported this study of student learning performance in Syria. This report would not have been possible without the support of the United Kingdom Government's Foreign, Commonwealth and Development Office.

We thank the enumerators who travelled throughout Province A and Province B to collect data for this assessment. We also thank the hundreds of students who took time from their school days to be assessed and speak with enumerators about their experiences of the war, as well as their teachers and head teachers who made this study possible.

This study was implemented with the technical assistance of School-to-School International and Chemonics International, Inc.

Table of Contents

List of Figures

List of Tables

LIST OF ACRONYMS

CADDPM	Correct addition problems per minute
CLSPM	Correct letter sounds per minute
COVID-19	Coronavirus disease 2019
CNRPM	Correct numbers recognised per minute
CNWPM	Correct nonwords per minute
CSUBPM	Correct subtraction problems per minute
CWPM	Correct words per minute
ED	Education Directorates
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
EU	European Union
FCDO	United Kingdom's Foreign, Commonwealth & Development Office
GESI	Gender Equality and Social Inclusion
IRR	Inter-rater reliability test
ORF	Oral reading fluency
SGO	Safeguarding Officer
STS	School-to-School International
SVR	Simple view of reading model
TOT	Training of trainers
USAID	United States Agency for International Development

Executive Summary

Introduction

This report presents the results of a study of learning outcomes for students served by the Manahel Syria Education Programme in Province A and Province B. Manahel is a 5-year project funded by the United Kingdom Government's Foreign, Commonwealth and Development Office (FCDO) and implemented by Chemonics International. Manahel provides access to safe, inclusive, and quality learning opportunities to children in Syria whilst strengthening education actors to manage education effectively.

The study examines the performance of grade 3 and 4 students in reading and mathematics across Manahel-supported schools. The results provide Manahel with insights to ensure the programme meets the needs of the schools and students it serves. The study's results also serve as a point of comparison to the 2020 Manahel midline assessment.[1] The study was conducted by Manahel partner, School-to-School International (STS).

Four research questions guided the study:

1. What proportion of grade 3 students in 2020 and in 2021 are classified as 'progressing' and 'proficient' readers and what proportion are non-readers?

2. How has last year's cohort progressed in reading and mathematics outcomes (grade 3 in 2020, now in grade 4)?

3. How do this year's grade 4 students compare to this year's grade 3 students?

4. To what extent is there a gender gap in reading and mathematics performance among this year's grade 3 and grade 4 students, respectively? Does the gender gap widen or narrow from grade 3 to grade 4?

[1] Early Manahel assessments in 2019 and 2020 included comparisons with the 2017 Idarah assessment. Whilst no comparisons with the Idarah assessment are made in this report, the assessment uses reading proficiency benchmarks established under that programme.

The study draws on data collected in November 2021 from 751 grade 3 students and 739 grade 4 students across 75 Manahel project schools in Province A and Province B. Student and school data was collected using three tools: an Early Grade Reading Assessment (EGRA), an Early Grade Mathematics Assessment (EGMA), and a head teacher survey.

Findings

RQ1: Proportion of students who are progressing and proficient readers and those who are non-readers

1. Grade 3 students in 2021 have comparable reading profiles to grade 3 students in 2020. Overall, there were no statistically significant differences between the reading proficiency classification of students in grade 3 in 2020 and students in grade 3 in 2021, although there was gradual improvement. In 2020, 21.7% of students met the reading proficiency benchmark of scoring 80% or higher on the reading comprehension subtask compared with 25.4% of students in 2021. This may point to teachers' mastery of combining in-class teaching with remote teaching techniques as school closures and multi-shift teaching sessions were prevalent in the 2020/21 school year and will likely continue in the 2021/22 school year. Internationally, school closures related to COVID-19 have had a negative impact on student performance, particularly in early grades. This highlights the significance of any improvement, however small.

   > **Related Recommendation:** The proportion of students achieving satisfactory levels of reading proficiency in 2021 has improved by 3.7% over 2020. However, given the impact of COVID-19 related school closures and prior to that of conflict, direct comparisons are problematic. The improvement in student performance is inadequate and statistically insignificant. As COVID-19 restrictions and the impact of conflict seem to be reducing the teaching of literacy and numeracy in the early grades (grade 1 and above), there needs to be a stronger focus. Teachers need to use continuous assessment more strategically to analyse the status of their students and then implement targeted remedial actions to help increase the proportion of students reaching proficiency in both reading and mathematics.

   Although more students became proficient readers, more students also fell into the non-reader category with 22.7% of grade 3 being non-readers in 2021 compared to 18.9% in 2019. This is also not statistically significant, but it is concerning. It may indicate that students who had moved to being beginning readers before COVID-19 have lapsed back to non-readers as they were unable to progress through remote learning.

   > **Related Recommendation:** Teachers should monitor non-readers more closely in class and particularly when schools revert to remote learning. This will support students most at risk of not gaining the reading skills that are key to greater fluency and comprehension. Manahel will implement specific school-based support to teachers and interventions to reduce the number of non-readers. Community consultations will also target this concern. These targeted services are particularly necessary for grade 3 and 4 students in both Province A and Province B. However, Manahel will also consider providing after school literacy

clubs to students in grade 2 in both provinces. Manahel is in the process of developing a teacher guide for grade 2 Arabic language classes. This is a user-friendly teaching and learning material that appropriately corresponds to the scope and sequence of the national curriculum. It is also part of the process of setting exemplars which can be replicated for other grades, to drive the improvement of literacy in the classroom. Teachers will also be trained and supported in identifying non-readers and providing appropriate remedial reading activities for them. Summer clubs and other after school interventions will also target this group of students.

RQ2: In 2021, grade 4 students outperformed grade 3 students

2. As expected, grade 4 students (both boys and girls) outperformed grade 3 students in every subtask on the EGRA and EGMA in 2021 in both provinces, indicating that students in grade 4 are building on the literacy and numeracy competencies learned in earlier grades. However, there are indications in the grade 4 EGMA results that stronger foundations in more complex skills need to be built in grade 3. It was noted that grade 3 students in Province A achieved higher accuracy scores on the letter name identification, oral reading fluency (ORF), reading comprehension, and listening comprehension subtasks compared to their peers in Province B. However, grade 4 Province A schools had more non-readers.

   Related Recommendation: To sustain these gains through the end of the programme, the Manahel team should continue to help teachers increase the amount of time spent on reading with the help of online tools and targeted interventions for non-readers. In addition, these data indicate that Manahel staff should reflect on whether teachers are using their classroom-based formative assessments to diagnose weaknesses in their students' learning, to inform their remedial or additional activities. Consequently, Manahel could assist teachers and students through the learning circles, remedial numeracy sessions and weekly mathematics challenges and homework. Such interventions could be targeted to where they are most needed based on EGRA/EGMA results – Province A for grade 4 and Province B for grade 3. In addition, the Manahel team should continue emphasising and innovating remote learning content that helps students practice and improve their reading skills, based on skills learned in the classroom.

RQ3: Girls in grades 3 and 4 outperformed boys in reading in 2021, and the difference was greater among grade 4 students

3. As in the 2020 Manahel assessment, the 2021 assessment showed girls outperforming boys in most reading subtasks (all except listening comprehension) in grades 3 and 4. Furthermore, differences between girls and boys in accuracy scores increased from grade 3 to grade 4. This finding indicates that the gap between boys and girls in literacy increases as students transition from grade 3 to grade 4.

> Related Recommendation: The early grade teachers should work with boys through to the end of the Manahel programme to build reading fluency skills and strengthen their foundational skills towards reading proficiency. In grade 3, teachers should focus on building the skills needed to attain comprehension, which will be crucial for students to build later academic skills. In grade 4, support should focus on more advanced fluency and comprehension to ensure that students are prepared for the transition to higher grades, where they are more vulnerable to drop out. Greater focus will also be put on monitoring the attendance of grade 3 and 4 boys as this appears to be a cause for concern. This initiative could be extended to monitor dropouts from school, although this is complicated by the mobility of students between schools and regular student absenteeism.

RQ4: Boys in both grades outperformed girls in mathematics in 2021, but the difference was less significant in grade 4

4. Data from the 2021 assessment showed that the gender gap in mathematics might be narrowing as students progress in grade level. Boys in grade 3 outperformed girls in all mathematics subtasks except addition 1. However, grade 4 boys and girls performed comparably on five of the eight subtasks: number recognition, missing numbers, addition 2, subtraction 2, and word problems.

> **Related Recommendations:** Teachers in grades 1 and 2 should make sure that all learners have understood the basic mathematical functions (addition/subtraction). Teachers in grade 3 should focus on more complex mathematics skills to ensure that students master mathematics operations and real-world thinking. This would also help students be better prepared for the more complex maths taught in grade 4. Manahel is creating girl-focused after-school centres and will measure reading and mathematics outcomes amongst grade 5 and 6 girls. Based on these results, Manahel may also monitor grade 4 girls' mathematics performance, especially in Province A. This would address the mathematics gender gap at an early stage and give the programme and teachers time to work with girls who struggle to match boys in their numeracy outcomes. Manahel will also work with the teachers to identify and help reduce possible bias in relation to assumptions about girls' ability to undertake mathematics. Manahel could observe specific teachers teaching maths to analyse their interaction with both boys and girls to see if the actions and bias of the teachers are related to female under-performance. This should focus on who is being asked questions, who is speaking in small group work, who is coming up to the board, and whose work is being celebrated.

Additional Recommendations Include:

System Related Recommendations:

- All of the above school-based recommendations will be more successful if supported from within the system. In particular, the system actors should assist schools in interpreting and analysing their continuous assessment results in the early grades to inform their teaching and to build remediation measures around the areas of weakness.

Manahel Related Recommendations:

- Manahel should work with schools following assessments (both internal continuous assessments and EGRA/EGMA) to assist the teachers in using the test data to inform remediation efforts. This will require Manahel to train teachers in how to analyse their students' assessment results to inform remediation and how to use the summer school clubs and after school lessons to maximum effect.

- Manahel should remove nonword subtasks in future EGRAs. Research on measuring reading in Arabic indicates that decoding may not contribute to reading comprehension because of the nature of the language.[2, 3] Thus, it is unsurprising that nonword fluency and accuracy scores remained relatively low. This confirms other research that shows that nonword tests are not a good predictor of learner reading performance, unlike letter sound identification, ORF and reading comprehension.

- This points to students in Province B pulling ahead in mathematics results, Manahel needs to analyse what elements of the intervention in Province B had an impact on learner performance and replicate these in Province A where students appear to be falling behind in relative terms – particularly in District 1, where there has been a notably high level of conflict over the past year.

The report also details recommended research activities for the remainder of the project, based on the findings. These are:

Learning Assessment (EGRA/EGMA):

This will be conducted in November 2022 and in May 2023 and will also serve as an endline measurement for the project. These assessments will include both schools which Manahel has supported but in which teachers are not being paid by the project and those where teachers are being paid. This would provide some measure of understanding of how providing teacher pay impacts on learner performance.

---

[2] Arabic is a diglossia language, meaning it has two variants for different situations. In this case, the first variant is Modern Standard Arabic (MSA), used for reading and writing, while another variant is the spoken colloquial dialect which can differ significantly from MSA. The simple view of reading (SVR) model, on which the EGRA is based, explains reading comprehension as the product of decoding (the ability to apply knowledge of letter-sound relationships, including knowledge of letter patterns, to correctly pronounce written words) and listening comprehension. However, the validity of SVR for Arabic has not been tested.

[3] Asadi, Ibrahim A., Asaid Khateb, and Michal Shany. 'How simple is reading in Arabic? A cross-sectional investigation of reading comprehension from first to sixth grade.' Journal of Research in Reading S1, no. 40 (2017): S1–S22. https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-9817.12093

Importance of Teacher Pay:

This research has been completed and is being further analysed to better answer questions about the relationship between payment of teachers and learner performance.

School Case Studies:

Manahel intends to develop case studies of schools where teachers are being paid and those where they are not being paid as part of the intervention. This would allow for a deep dive to better understand the triangular relationship between learners' performance, teachers' payment, and teachers' commitment.

Positive Deviance Studies:

To supplement the research, Manahel will develop detailed case studies of individual project schools which have seen a robust improvement of learner results and/or are maintaining high levels of learner performance in EGRA and EGMA in conditions where other schools are failing to do the same. This would help better understand the conditions that lead to improved and sustained learner performance.

Student Gender and Vulnerability Research:

Three pieces of work are proposed with a focus on gender and disability:

1. A study to explore attendance by girls in the early grades and/or attendance of children with disabilities in the early grades over time in unsupported schools (to test the assumption that the weight of supporting the payment and support of teachers by parents falls disproportionately on parents of girls and children with disability) and compare that to attendance of these two groups in supported schools using a case study approach.

2. Small-scale research to understand if girls' well-being is comparable to boys in the later years and widen the time-on-task/lesson observation work to a small number of upper primary teachers (approximately 30) to see if there is a discernible difference in teaching. Manahel will prioritise introducing learning circles to support teachers to create gender-responsive pedagogy and a growth mindset and measure how these interventions are perceived.

3. Manahel is planning a GESI review during the extension period and reflecting on improved GESI focused activities that can be applied during the extension period as well as making recommendations for future programming.

Introduction and Background

Conflict and Education in Syria

Since March 2011, the Syrian Arab Republic has been embroiled in a conflict between the government of Syria and opposition forces, which has fractured governance in Syria. Currently, the opposition coalition's Syria Interim Government provides civil services – including education management through the Education Directorates and the Ministry of Education – in Provinces A and B.

More than a decade of acute crisis has devastated the education sector in the region. Airstrikes punctuate school days. Children are burdened by the emotional and physical toll of personal loss and continued instability. Teachers, bearing the same burdens as their students, choose to go to schools in the face of danger, sporadic pay, and the challenge of providing a semblance of normalcy for their students.[4]

These challenges only increased in the spring of the 2019/20 academic year when COVID-19 disrupted education worldwide. Schools closed in mid-March 2020 and were required to pivot to an online-learning approach quickly. Although schools reopened for the 2020/21 academic year in November 2020 – two months later than scheduled – schooling continues to be punctuated by temporary closures due to spikes in COVID-19 infection rates.

Manahel Programme Background

The five-year Manahel Syria Education Programme is funded by the United Kingdom's Foreign, Commonwealth and Development Office (FCDO), benefits from targeted support from the Qatar Fund for Development, and is implemented by Chemonics International. The Manahel programme builds upon the previous Idarah project, funded by the United Kingdom and the European Union between 2014 and 2018.

[4] 'Education', United Nations Children's Fund, www.unicef.org/syria/education

From February 2018 to June 2022, the Manahel programme is providing access to safe, inclusive, and quality learning opportunities for children in conflict-affected, opposition-held areas of Syria whilst strengthening educational actors to manage education effectively. Manahel focuses on pedagogy, curriculum and planning, inclusion, and child protection. Initiatives within Manahel's intervention structure include psycho-social care workshops and activities, child safeguarding and protection activities, fixed and mobile library sponsorship, teacher training and coaching, monthly teacher learning circles, accommodations for children with mild or moderate disability, and reading and mathematics instruction. On average, teachers deliver 12 literacy sessions per month using Manahel materials. These sessions are in addition to the standard Arabic and mathematics lessons.

With the school closures due to COVID-19 in 2020, Manahel built on their non-formal education experience to swiftly roll out a suite of online and remote learning tools. By the autumn of 2021, Manahel had combined in-person and online education and protection opportunities for 189,912 children at 430 schools and 40 tent schools. Online learning continues to supplement in-school learning. This is supplemented with support for parents and guardians on supporting their children's reading development.[5] Overall, Manahel has had an impact on some 560,000 children over the course of the project in Provinces A and B in Syria.

[5] NWS has a high literacy rate and parents' ability to read is not a factor delaying children's reading development. However, parents struggle to find time for reasons related to securing livelihood and the fact that average family size is 5-6 children. Manahel has found that often a sibling supports the younger children and when the project structured a clear routine with a clear entry point, parents' engagement increased significantly - for instance reading a bedtime story that Manahel shares via WhatsApp to parents every day at the same time.

Methodology

Study Purpose and Research Questions

The 2021 Manahel Learning Assessment examines the reading and mathematics performance of students attending Manahel-supported schools. This study seeks to provide Manahel schools and EDs, along with FCDO and Manahel staff, insights on factors influencing student performance to ensure the programme meets the shifting needs of the schools and students it serves. The study was conducted by School-to-School International (STS).

The study draws on data collected in November 2021 from grade 3 and grade 4 students. Data collection included three tools: an Early Grade Reading Assessment (EGRA), an Early Grade Mathematics Assessment (EGMA), and a head teacher survey.[6]

Four research questions guided the study, each with a distinct purpose. These are presented in Table 1.

Table 1. 2021 Learning Assessment Research Questions

Research Question	Purpose
1. What proportion of G3 students in 2020 and in 2021 are classified as 'progressing' and 'proficient' readers, and what proportion as non-readers?	To measure progress against the programme's Impact Indicator and to compare the changes in the percentage of students who can read over time.[7]
2. How have last year's cohorts progressed in reading and mathematics outcomes (grade 3 in 2020, now in grade 4)?[8]	To track growth within a cohort across an additional year of schooling and time in the Manahel programme. This information further provides insights to ensure the programme meets the evolving needs of the schools and students it serves.

[6] Tools and their uses are included in Annex D: Assessment and Survey Tools

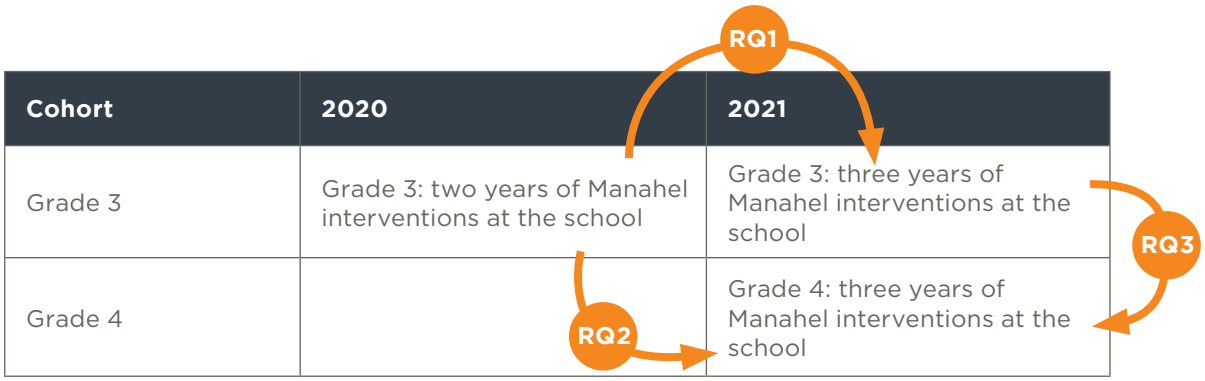
[7] The impact indicator is 'Percentage of students in the top two categories in proficient reader, and advanced progressive reader of early grades students as measured by EGRA results.'

[8] Note that this study will not track individual students from previous studies.

3. How do this year's grade 4 students compare to this year's grade 3 students?	To serve as a proxy comparison group. It also helps identify additional learning that might occur with another year of Manahel intervention.
4. To what extent is there a gender gap in reading and mathematics performance among this year's grade 3 and grade 4 students, respectively? Does the gender gap widen or narrow from grade 3 to grade 4?	To understand differences in performance based on gender and identify any gaps in performance related to gender.

Figure 1 describes the comparisons and assumptions contained within these research questions.

Figure 1. Comparisons and Assumptions in Research Questions



Cohort	2020	2021
Grade 3	Grade 3: two years of Manahel interventions at the school	Grade 3: three years of Manahel interventions at the school
Grade 4		Grade 4: three years of Manahel interventions at the school

The school head teachers were only informed the evening before the visit and the call only provided broad details of what the assessment process would involve. The Manahel access team confirmed that the attendance of students on the day of the assessment in sampled schools was not obviously different from other school days. Informing the schools at short notice helps ensure that weak or disabled students are not asked to stay away on the day of the assessment - this would impact the validity of the results.

Sampling

Sampling was undertaken using a 2-stage approach:

First, STS randomly selected 52 schools in Province A and 23 in Province B, which is proportional to the complete Manahel school list. Replacement schools were randomly selected by STS from the full list of schools in the same way as the sample schools. When there was a need for replacement, Manahel informed STS and STS confirmed the replacement school.

Next, enumerators randomly selected 20 students per school to complete the learning assessments – 10 grade 3 students and 10 grade 4 students, with as equal gender distribution as possible. This made sure that the teachers did not pre-select the

strongest students for the assessment. The head teacher at each school also completed their selected survey. Table 2 provides a summary of the target and final sample.

Table 2. Target and Final Sample

Province	Target Sample						Final sample					
	Schools	Students					Schools	Students				
		Total	G3 Girls	G3 Boys	G4 Girls	G4 Boys		Total	G3 Girls	G3 Boys	G4 Girls	G4 Boys
A	52	1,040	260	260	260	260	56	1,117	282	280	284	271
B	23	460	115	115	115	115	19	373	91	98	93	91
Total:	75	1,500	375	375	375	375	75	1,490	373	378	377	362

The student sample is generalisable at the province level or by gender. Results at lower subgroup levels are associated with lower levels of confidence. Whilst the targeted number of boys and girls to be assessed varied within a province, the numbers in the final sample do not deviate significantly from the planned numbers. Therefore, results by gender and province are valid.

Assessment Tools

The Manahel learning assessment used tools previously developed for early grade reading and mathematics research conducted within Syria and the broader region. Idarah conducted an EGRA and EGMA in 2017 with a version of a 2012 tool developed by the MAHARAT project in Iraq. Manahel used the same EGRA and EGMA tools for the 2019, 2020, and 2021 learning assessments to maintain consistency across studies. However, Manahel did make changes to the nonword tool and minor changes to the administration of the tests to improve their quality and, thus, the accuracy of students' results. A summary of these changes is captured in Table 3.

Table 3. Summary of Assessment Changes Across Studies

	2017 Idarah Study	2019 Manahel Study	2020 Manahel Study	2021 Manahel Study
Student Assessment Timepoint	END of grade 3	START of grade 3[9] to be closer in line with the Manahel logical framework indicators and reflects international best practices	START of grade 3 and grade 4 for the inclusion of grade 4 as a proxy comparison group to grade 3	

[9] International standards recommend assessing students consistently at the end of grade 2 or beginning of grade 3.

	2017 Idarah Study	2019 Manahel Study	2020 Manahel Study	2021 Manahel Study
EGRA Tools	EGRA tool adapted from tools from the 2012 MAHARAT project in Iraq	Updated EGRA with Arabic modifiers for greater precision	2019 EGRA plus the Idarah letter sound identification subtask without modifiers added to provide more nuance around students' reading ability of Arabic modifiers	2020 EGRA plus a revised nonword reading subtask from the USAID-funded Quality Instruction Towards Access and Basic Education Improvement (QITABI) to more closely adhere to the patterns of Arabic words[10] All the other subtasks were the same as in 2019 and 2020
EGMA Tools	EGMA tool adapted from tools from the 2012 MAHARAT project in Iraq	Same EGMA tool as Idarah		
Supporting Surveys	Included: War Stressor Survey Head Teacher Survey Teacher Survey Classroom Observation School Observation	Included: Student Stressor Survey Head Teacher Survey	Included: New Student Survey with focus on access to learning Head Teacher Survey New Teacher Survey with focus on displacement and teaching practices Safeguarding Officer Tool to triangulate data with student responses	Included: Head Teacher Survey
Assessment Administration	Paper administration with stopwatches and timers		Tablet administration using Tangerine®[11, 12]	
Accommodation for Students with Disabilities	None		Timed subtasks were extended from one to two minutes All stimuli were printed with larger font	

[10] A full discussion of the rationale for including a new nonwords subtask is included in Research Question 3

[11] Tangerine® is an open-source software developed by RTI International specifically for the administration of EGRA and EGMA.

[12] Manahel chose to collect the data electronically on tablets to ensure more accurate scoring and better overall data quality. This change required extensive updates to all instructions. STS updated the instructions for the tablet administration in line with the Early Grade Reading Assessment Toolkit, Second Edition and the Early Grade Mathematics Assessment Toolkit.

The final tools for the 2021 assessment include an updated EGRA, consistent EGMA, and consistent head teacher survey. These tools are described in Table 4. A copy of all evaluation tools appears in Annex D: Assessment and Survey Tools.

Table 4. Description of the 2021 Evaluation Tools

Instrument	Description
EGRA and EGMA	The EGRA and EGMA are comprised of subtasks that each measure a foundational skill of reading or mathematics. They are used to determine where a student is in their progression towards proficiency.
Head Teacher Survey	The head teacher survey includes a brief survey on student enrolment. It is used to apply sampling weights to the student data.

Data Collection and Analysis

The enumerator training followed a training-of-trainer (TOT) cascade model. The TOT took place remotely over Zoom© on 20-21 October 2021 for four hours each day. The STS team in the United States trained four Manahel trainers on data collection procedures – two based in Gaziantep, Turkey, and two based in Province A, Syria. The TOT was given in English with interpretation provided by the Manahel programme leads. The four Manahel trainers, in turn, trained 24 enumerators who were selected from the Manahel programme staff, in Arabic, with materials provided by STS on 30–31 October 2021. Standard EGRA/EGMA training approaches were used including practice assessments with feedback and inter-rater reliability (IRR) tests. The outlier trainee enumerators were not included in the final data collection process.

Between 2-24 November 2021, enumerators visited 75 Manahel-supported schools. Enumerators were divided into five teams of four. Each team visited one school per school day and assessed 10 students in grade 3 and 10 students in grade 4. Enumerators uploaded data daily from their tablets via Wi-Fi to a secure, password-protected server maintained by STS staff.

Supervision and Quality Control: Throughout data collection, enumerators were closely supervised to ensure data quality. The Manahel programme leads tracked the progress of the data collection daily. The Manahel trainers performed site visits to ensure enumerators were following protocols. STS staff monitored the data uploaded to the server daily. An additional means of data quality control was using inter-rater reliability (IRR) measures during data collection with 10% of the sampled students, per standard EGRA practice.[13] Results showed that enumerators administered the tools consistently.

- **Child Protection and Research Ethics:** Throughout the programme, Manahel staff ensured children were protected, and the research was conducted in line with research ethics and child protection practices. The Manahel team reviewed the study tools before data collection to ensure that the study adhered to applicable ethical rules and societal norms. All enumerators received training on the programme's

[13] Inter-rater reliability is the degree of agreement between two enumerators who are assessing the same student independently. It allows the data collection monitors to identify and resolve problems within enumerator teams during data collection to improve quality.

code of conduct and child protection policies and procedures. Affirmative informed consent was obtained from all head teachers. Teachers and all children provided affirmative assent to be assessed. They could opt out of the assessment at any time. Students were selected randomly on the day of the data collection and schools were only informed the evening before the assessment to make sure weak students and those with disabilities were not excluded from the random sampling procedure or participating. As a result, accommodations for students with disabilities – such as extended time for the timed subtasks and large print stimuli – were provided to all students throughout data collection.

Data Analysis

After data collection, STS staff cleaned the data to remove invalid observations, resulting in a complete, accurate and internally consistent final data set. STS followed a multistage data cleaning plan to ensure data values were within the allowable range. STS developed a master codebook and merged EGRA and EGMA data sets with the head teacher survey data.

The STS team applied sampling weights to the students' data to produce more representative estimates. To compute sampling weights, STS used the following information about all the schools in the relevant population: education authority or district; the number of students enrolled in grade 3 and grade 4; and the number of students in attendance in grade 3 and grade 4 on the day of testing. This data was collected through the school's head teacher survey at the beginning of each school visit. Weights were computed using SPSS version 25.

After applying the weighting functions, STS analysts produced descriptive statistics disaggregated by variables of interest. Descriptive results were analysed for statistically significant differences by sex, province, and grade using chi-square tests and t-tests.[14] Associations between respondent characteristics and student performance were further analysed using Pearson bi-variate correlations. All analyses were conducted using SPSS version 25.

[14] The chi-square test is a statistical test comparing the proportion of students who did not respond correctly to any items on a subtask –known as zero scores – with what was expected. The independent-sample t-tests compare the difference between the means of two independent groups on the same dependent variable.

Challenges and Limitations

The following limitations should be considered when reviewing the findings of the 2021 Learning Assessment:

1. The study is not a randomised control trial design. Schools were not randomly assigned to the treatment groups at the beginning of the study. Data analysis methods attempt to correct for the non-random approach to sampling by controlling for any confounding variables. However, it is always possible that a major confounding variable is not identified and appears in the analysis.

2. Results cannot confidently be ascribed to continuous student engagement in Manahel programming. While **schools** assessed in the 2019 and 2020 studies were included in the 2021 study sample, the study design did not identify individual **students** who participated in previous studies for reassessment. Previous studies of Manahel students found that large majorities – 83.2% in 2019 –had moved one or more times in the past academic year, indicating a high rate of student turnover within these schools. The Manahel team believes the level of mobility in 2021 is likely to have been considerably lower.

3. The 2021 assessment did not collect data regarding students' exposure to or dosage of Manahel interventions. Given student mobility, the sample will include some students who have not received the full dosage of the intervention. As a result, findings cannot be directly attributed to programme activities, and results should be interpreted with caution. Furthermore, in the absence of a comparison or 'control' group, the research cannot determine how the progression of students participating in Manahel interventions compares to expected progression between grade 3 and grade 4.

4. Learning loss resulting from the COVID-19 pandemic may have affected results, but this study did not attempt to understand that phenomenon. Schools closed entirely in the spring and summer of 2020 due to the outbreak of COVID-19. They remained closed through the first weeks of the 2020/21 and 2021/22 school years. Additionally, to implement social distancing requirements schools operated a double or triple shift system. Overall Manahel estimates that 10 – 15% of normal class time was lost for each grade 3 student over this period. Therefore, it can be assumed that some learning loss resulted from these changes; however, this study does not aim to understand such effects.

5. Ongoing instability in the region due to the conflict presented logistical challenges during data collection. For example, one sampled school closed halfway through the day due to air strikes and needed to be replaced.

Findings

This section reports findings according to the study's four main research questions. Results statistically significant at the $p < 0.05$ level are referred to as 'significantly' lower or higher in the text.

Description of Sample

The sample was equally balanced between grade 3 and grade 4 students, each group representing approximately 50.0% of the overall sample. Girls made up 50.3% of the sample, while boys accounted for the remaining 49.7%. Students ranged in age from 6 to 16 years old. Most (72.7%) were on-age for their grade, but 0.2% were underage and 27.1% were over-age.

Mirroring the relative population distribution in the two provinces, the majority of the sample came from Province A (75.0%) compared to Province B (25.0%). Within Province A, students were relatively equally divided between District 7, District 2, District 3, and District 8 (6.7%, 13.4%, 13.3% and 5.3% of the overall sample, respectively). A slightly larger proportion came from the Province A city limits and District 1 (17.4% and 18.8%, respectively). About 81.3% of students from the Province B sample came from District 9, while 18.8% came from District 10.

Because the sample's distribution is proportional to the relative populations of the provinces and districts, the overall effect is that results are driven mainly by trends seen in Province A as three quarters of the sampled schools are in that province.

Research Question 1:
Progressing and Proficient Readers

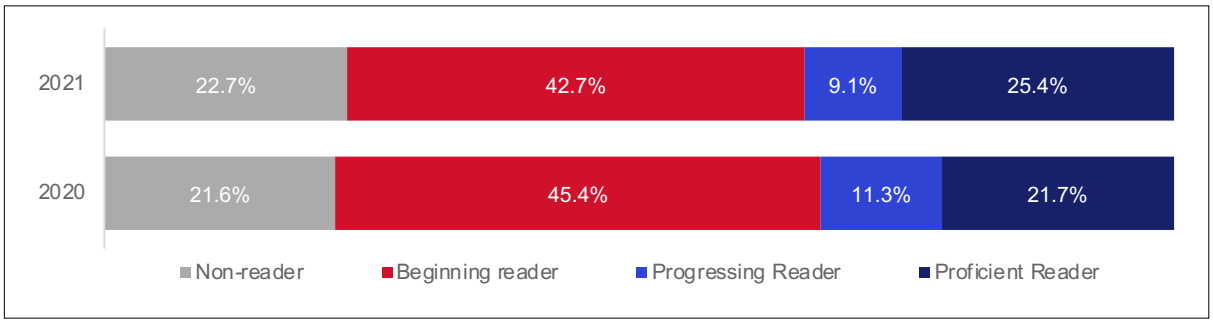
Research Question 1: What proportion of grade 3 students in 2020 and in 2021 are classified as 'progressing' and 'proficient' readers?

As with the 2019 and 2020 Manahel learning assessments, the 2021 assessment classified students' EGRA scores into proficiency bands established by the 2017 Idarah assessment. All data is derived solely from the oral reading fluency (ORF) and reading comprehension subtasks. These bands tie the ability to read at a 'proficient' level – the highest band – directly to comprehension, while lower bands also consider fluency. The reading proficiency bands are defined as follows:

- Non-readers are students who did not read a single word of the ORF passage.
- Beginning readers are students who read between 1 and 22 correct words per minute (CWPM) on ORF and answered fewer than 80% of questions correctly on the reading comprehension subtask.
- Progressing readers are students who read 23 CWPM or more on ORF and answered less than 80% of the reading comprehension subtask correctly.
- Proficient readers are students who answered 80% or more of questions correctly on the reading comprehension subtask.

Overall, there were no significant differences between the reading proficiency classification of students in grade 3 in 2020 and students in grade 3 in 2021. In 2020, 21.7% of students met the reading proficiency benchmark of scoring 80% or higher on reading comprehension compared to 25.4% of students in 2021 (see Figure 2).

Figure 2. Proportion of 2021 Grade 3 and 2020 Grade 3 Students by Reading Proficiency Level



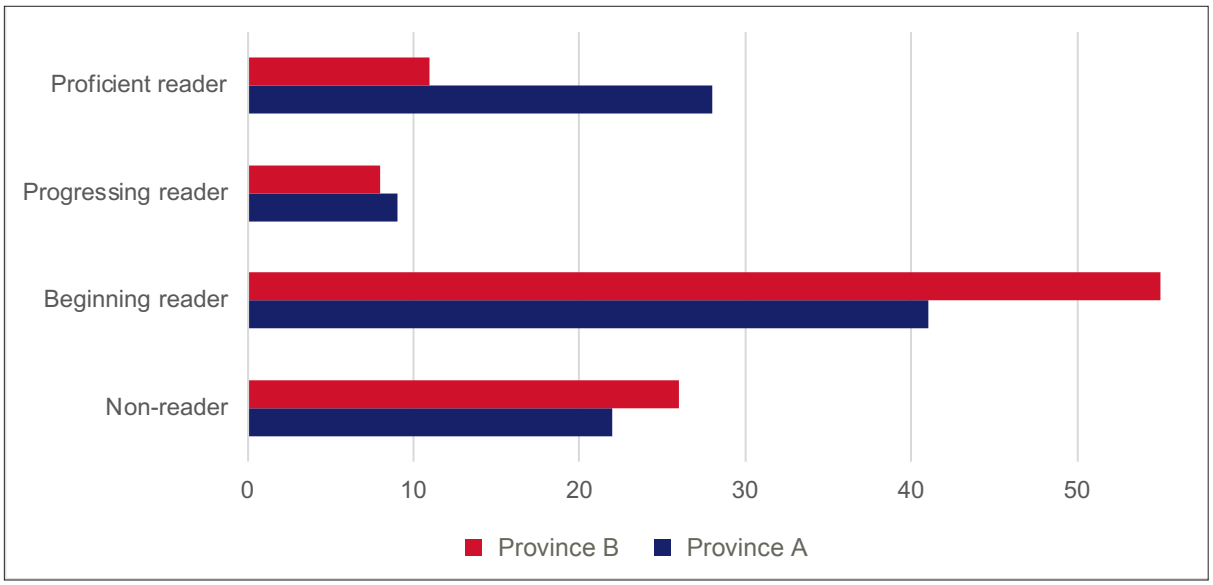
It is worth noting that both 2020 and 2021 grade 3 EGRA results show a significant and consistent improvement on those of 2019. The percentage of proficient readers has risen 11.5% from 13.9% in 2019. However, the proportion of students with zero scores has also increased by a few percentage points from 18.9% in 2019. This may indicate that

for about a fifth of students remote learning has been a real challenge and without the initial reading foundations they have not been able to start decoding.

Results by Province. Province A had a significantly higher proportion of proficient readers than Province B in both years. However, readers in Province B improved in 2021 compared to 2020.

In Province A, more students read at the proficient level in 2020 and 2021 (23.8% of proficient readers in Province A in 2020, compared to 10.8% in Province B; 28.0% of proficient readers in Province A in 2021, compared to 11.3% in Province B). In 2020, Province B had a significantly higher proportion of non-readers than Province A (32.9% compared to 19.4%, respectively.) However, in 2021, the statistical difference was no longer among non-readers but beginning readers. In Province B, 54.9% of students were beginning readers compared to 40.5% in Province A in 2021. Province B had 25.8% of students as non-readers, similar to the 22.2% in Province A (see Figure 3). Readers in Province B were slightly more advanced in the 2021 cohort of 3rd graders compared to 2020.

Figure 3: Comparison of Grade 3 Student Reading Proficiency Levels in Province A and Province B Schools in 2021



Results by sex. A significantly larger proportion of boys met the benchmark in 2021 compared to 2020. In 2020, only 13.1% of boys met the reading benchmark, compared to 23.4% in 2021. There were no significant changes in the proportion of girls attaining the reading benchmark between 2020 and 2021.

Additional results are included in Annex E: Disaggregated Results.

Research Question 2:
Student Progression From Grade 3 to Grade 4

Research Question 2: How has last year's cohort progressed in reading and mathematics outcomes (grade 3 in 2020, now in grade 4)?

The purpose of this research question is to understand growth within a cohort of students across an additional year of Manahel interventions.

Overall, EGRA and EGMA results show that students in grade 4 in 2021 outperformed students in grade 3 in 2020 based on the administration of the same tests to both grades. This indicates that students improve their learning with an additional year of schooling, as is expected. Additional school disruptions due to the ongoing COVID-19 pandemic and conflict in Syria have not resulted in learning regression. However, these results do not indicate if students in grade 4 are performing at the expected level.

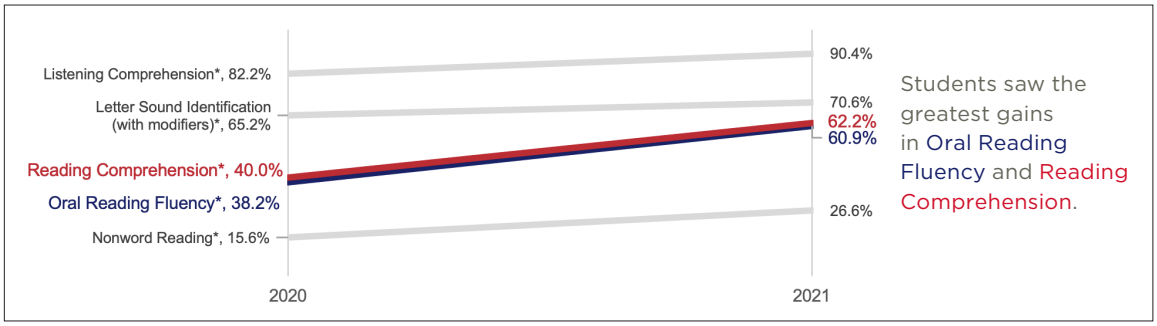
READING OUTCOMES

Overall, a higher proportion of grade 4 students met the reading proficiency benchmark in 2021 than did grade 3 students in 2020 (52.0% compared to 21.7%, respectively). There was also a statistically significantly higher proportion of non-readers and beginning readers amongst grade 3 students in 2020 than grade 4 students in 2021. In 2020, 21.6% of grade 3 students were non-readers, compared to 12.6% of grade 4 students in 2021. Similarly, 45.4% of grade 3 students were beginning readers in 2020, compared to 23.3% of grade 4 students in 2021.

Amongst EGRA fluency scores, students in 2021 achieved a letter sound fluency score of 43.2 correct letter sounds per minute (CLSPM); a nonword reading fluency score of 7.0 correct nonwords per minute (CNWPM); and an ORF score of 31.3 CWPM. In 2020, students read 34.4 CLSPM; 4.0 CNWPM; and 17.8 CWPM. These differences were all statistically significant.

Amongst accuracy scores, students in 2021 had statistically significantly higher scores than in 2020 on every EGRA subtask (see Figure). In 2021, students had an average of 70.6% of letter sounds with modifiers correct; 26.6% of nonwords correct; 60.9% of ORF words correct; 62.2% of reading comprehension questions correct; and 90.4% of listening comprehension questions correct. In contrast, students in 2020 averaged 62.5% of letter sounds with modifiers correct; 15.6% of nonwords correct; 38.2% of ORF words correct; 40.0% of reading comprehension questions correct; and 82.2% of listening comprehension questions correct. The lowest scores in 2021 remained on the nonword reading subtask, as was the case in 2020 and 2019.

Figure 4. Reading Accuracy Scores by Subtask and Year

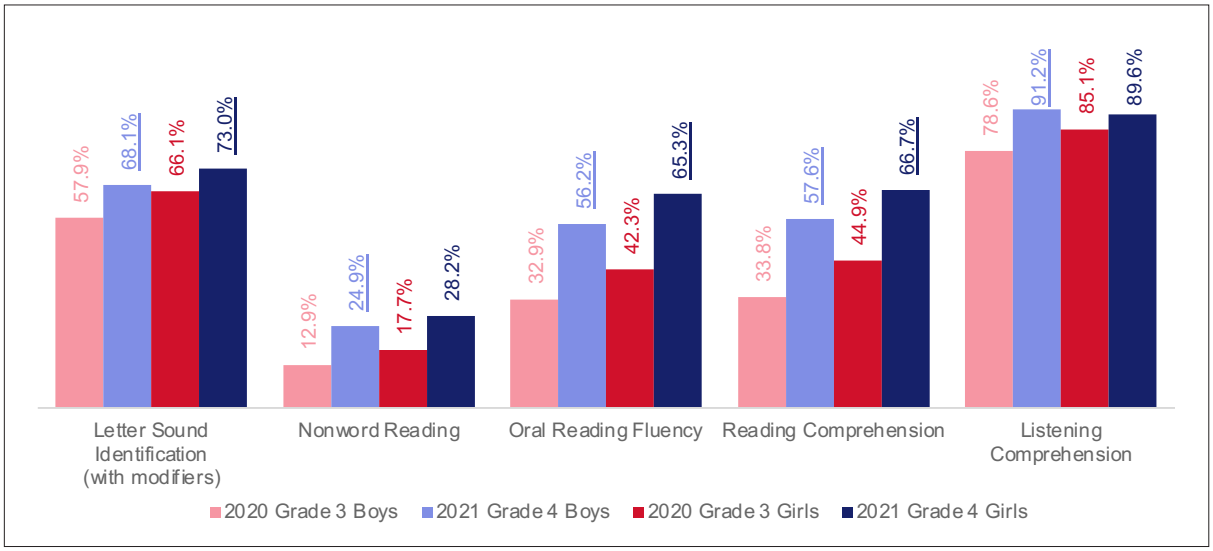


Note: Differences for subtasks with an asterisk (*) are statistically significant at p<0.05.

Most trends in zero scores – where a student could not answer a single item correctly in a subtask – showed that a statistically significantly lower proportion of students had zero scores in 2021 compared to 2020, further supporting the assertion that students in grade 4 improved their learning. The sole exception was letter sounds. In 2021, 6.1% of students received a zero score on the letter sound identification subtask, significantly higher than the 3.2% in 2020. This may be because grade 4 students were far enough advanced that a simple task such as identifying letter sounds was confusing, as it was more basic than what they were used to doing.

By sex. A statistically significantly higher proportion of boys and girls in grade 4 met the reading proficiency benchmark than grade 3. In grade 4, 48.6% of boys met the benchmark (compared to 13.1% in grade 3). Amongst girls in grade 4, 55.2% met the reading benchmark, compared to 28.5% of grade 3 girls. The same trend was seen for accuracy scores (see Figure). Boys and girls in grade 4 had significantly higher fluency and accuracy scores in all subtasks than boys and girls in grade 3.

Figure 5. Reading Accuracy Scores by Subtask, Grade, and Sex



Note: Underlined scores are statistically significant between grades at p<0.05.

By province. Grade 4 students in both provinces had higher scores in 2021 compared to 2020. In Province B, 46.9% of grade 4 students met the reading benchmark compared to 10.8% of grade 3 students in 2020. In Province A, 53.2% of grade 4 students met the reading benchmark compared to 23.8% in 2020.

Students in Province A generally had higher scores in 2020. In 2020, students in Province A had statistically significantly higher fluency scores than Province B in letter sound fluency (35.1 CLSPM in Province A compared to 30.7 in Province B) and ORF (19.3 CLWPM in Province B compared to 9.9 in Province A). They also had statistically significantly higher accuracy scores in every EGRA subtask except nonword reading and listening comprehension.

In 2021, students in Province B had slightly higher scores than students in Province A, although the difference was only statistically significant for accuracy scores in letter sound identification (77.4% compared to 69.0%) and nonword reading fluency (36.2% compared to 24.3%). In Province B, students identified 9.4 CNWPM, statistically significantly higher than 6.4 CNWPM in Province A. In other subtasks, students in

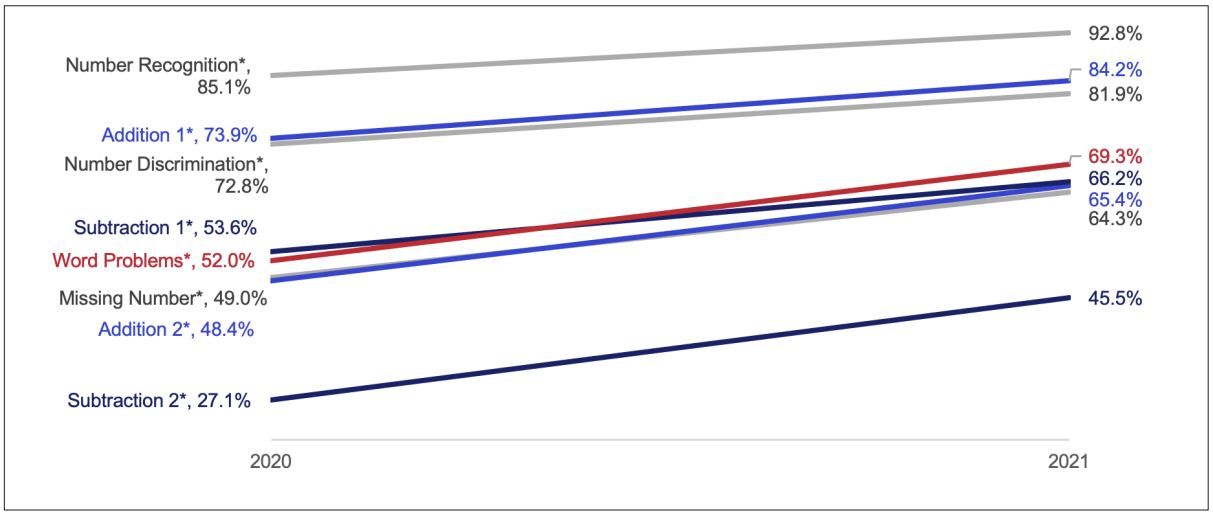
Province A identified 42.7 CLSPM (compared to 45.6 CLSPM in Province B) and read 31.5 CWPM (compared to 30.6 CWPM in Province B). These differences in fluency scores were not statistically significant.

MATHEMATICS OUTCOMES

Grade 4 students had significantly higher mathematics scores compared to grade 3 students. In EGMA fluency scores, students in 2021 had a number recognition fluency score of 39.6 correct numbers recognised per minute (CNRPM); 11.3 correct addition problems per minute (CADDPM); and 7.4 correct subtraction problems per minute (CSUBPM). In 2020, students had 28.7 CNRPM, 8.7 CADDPM, and 5.6 CSUBPM.

In accuracy scores, students in 2021 again had statistically significantly higher scores than students in 2020 in every EGMA subtask, as shown in Figure 6, with notable gains in the more advanced subtasks of addition, subtraction, and word problems. On average, 2021 students had an average score of 84.2% in addition 1; 65.4% in addition 2; 66.2% in subtraction 1; 45.5% in subtraction 2; and 69.3% in word problems. In 2020, students had an average score of 73.9% in addition 1; 48.4% in addition 2; 53.6% in subtraction 1; 27.1% in subtraction 2; and 52.0% in word problems. For addition 2, students moved from an average of just over two questions correct to just over three questions correct (out of five). For subtraction 2, students moved from just over one question correct to just over two questions correct (out of five).

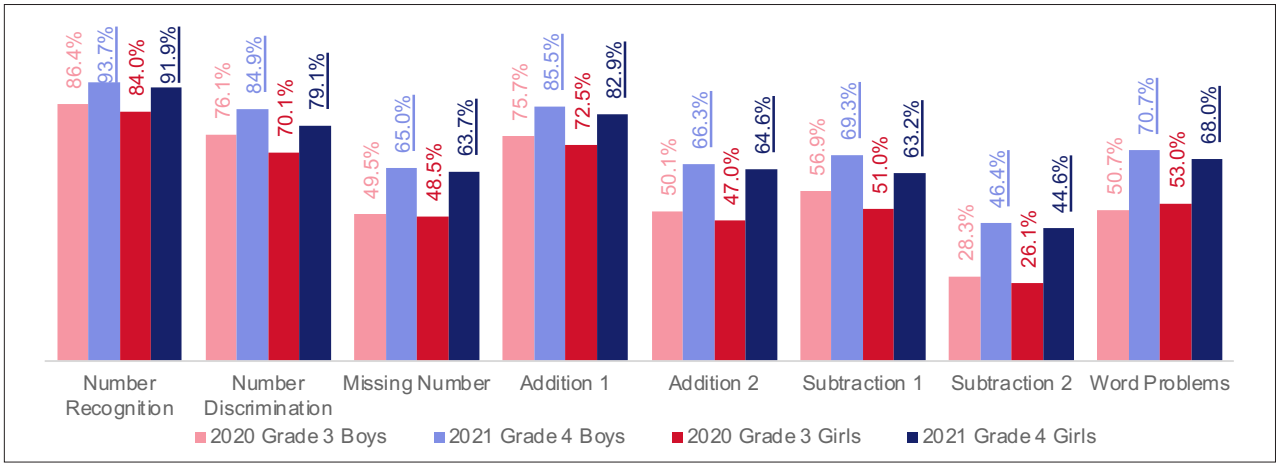
Figure 6. Mathematics Accuracy Scores by Subtask and Year



Note: Differences for subtasks with an asterisk (*) are statistically significant at p<0.05.

By sex. A statistically significantly higher proportion of boys and girls in grade 4 achieved higher fluency and accuracy scores than boys and girls in grade 3 (see Figure 6). This further supports the trend that students in grade 4 had improved their learning since grade 3.

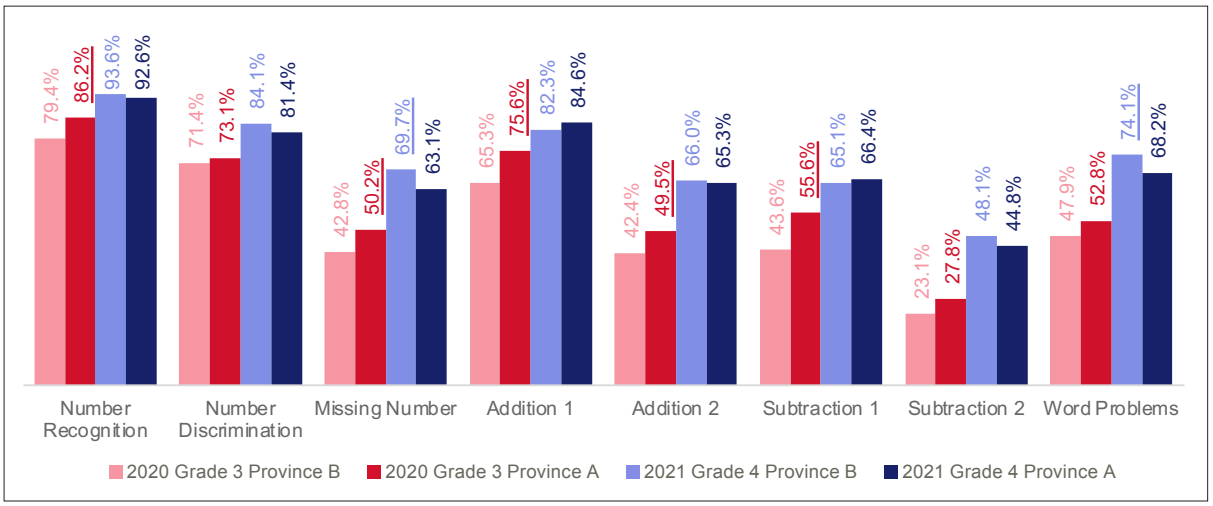
Figure 7. Mathematics Accuracy Scores by Subtask, Year, and Sex



Note: Underlined scores are statistically significant between grades at p<0.05.

By province. In 2020, students in Province A generally had higher mathematics scores than students in Province B, but scores by province were mostly comparable in 2021. In 2020, students in Province A had statistically significantly higher scores compared to Province B in number recognition fluency (30.4 CNRPM in Province A compared to 19.9 CNRPM in Province B), addition fluency (9.0 CADDPM compared to 7.1 CADDPM in Province B), and subtraction fluency (5.9 CSUMPM compared to 4.5 CSUBPM in Province B). They also had statistically significantly higher accuracy scores in every subtask except number discrimination, subtraction 2, and word problems (see Figure 8.) In 2021, students in Province A and Province B performed comparably in mathematics subtasks. However, students in Province B had statistically significantly higher accuracy scores in missing number identification and word problems.

Figure 8. Mathematics Accuracy Scores by Subtask, Year, and Province



Note: Underlined scores are statistically significant between provinces at p<0.05.

Research Question 3:
Comparison between Grade 3 and Grade 4 Students

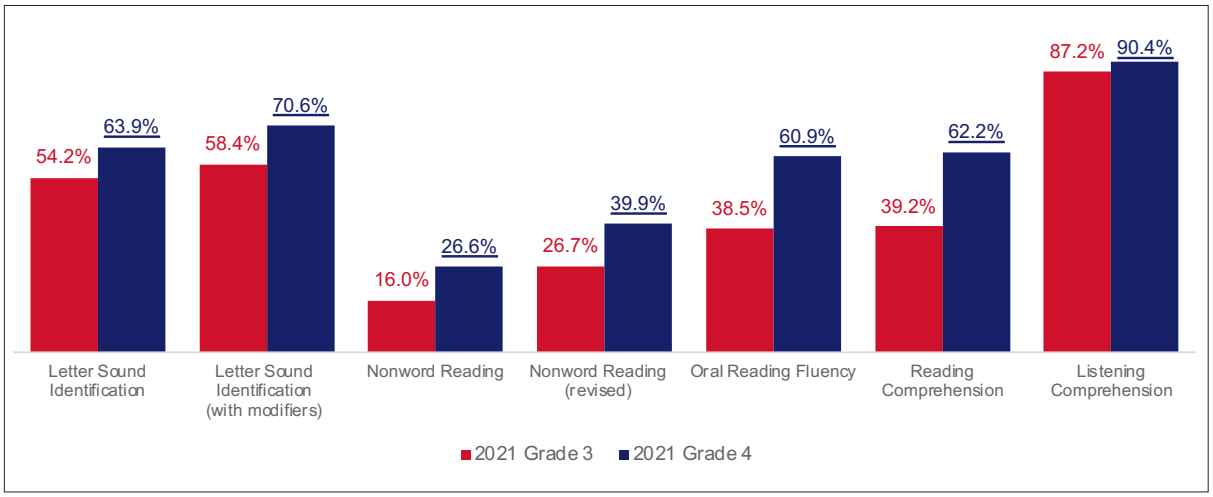
Research Question 3: How do this year's grade 4 students compare to this year's grade 3 students?

This research question serves as a proxy comparison group, exploring 2 different cohorts at the same time point (beginning of the academic year 2021) using the same tests but with the different levels of exposure to interventions (two years for grade 3; three years for grade 4). For comparability, this question assumes that students have been enrolled in schools with three years of exposure to Manahel interventions and that cohorts are comparable on external factors, such as exposure to conflict and COVID-related closures. However, these assumptions are tenuous given the protracted conflict in Syria and high rates of student displacement, in addition to the extra year of schooling received by students in grade 4. Additionally, as discussed in the Limitations section, no data were collected on students' exposure to or dosage of Manahel interventions. Therefore, results should be interpreted with caution.

READING OUTCOMES

Overall, grade 4 students significantly outperformed grade 3 students in every EGRA subtask (see Figure 9 .) This was true of boys in grade 4 compared to grade 3 and girls in grade 4 compared to grade 3. While all differences in scores between grades were statistically significant, grade 4 students had notably higher accuracy scores in ORF and reading comprehension – more advanced reading skills. Grade 4 students achieved an average accuracy score of 60.9% on ORF (compared to 38.5% for grade 3) and 62.2% for reading comprehension (compared to 39.2% for grade 3). A significantly lower proportion of grade 4 students received zero scores in all subtasks except listening comprehension (fewer than 10 students in each grade received a zero score in this subtask).

Figure 9. 2021 Reading Accuracy Scores by Grade



Note: Underlined scores are statistically significant between grades at p<0.05.

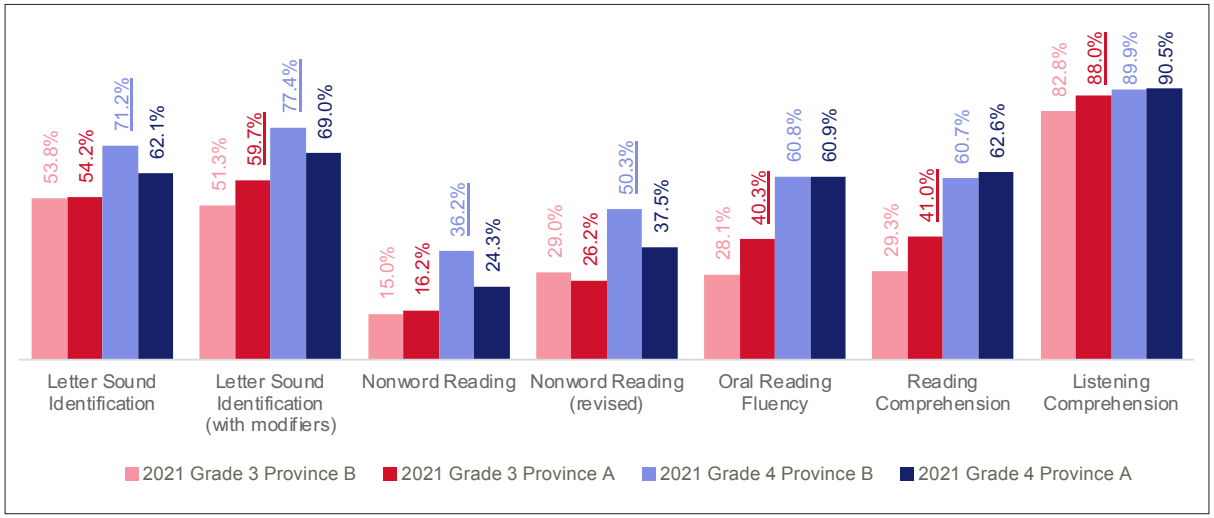
Nonwords. Nonwords. The 2021 assessment administered an additional revised nonwords subtask borrowed from the Qitabi programme in Lebanon.[15] The original subtask was also included for continuity and comparisons with previous assessment points. Students in 2021 were assessed using both nonword subtasks but were randomly assigned which subtask they received first to control for bias introduced by test fatigue.

Results show that students of both sexes and in all grades and provinces had higher scores in the revised nonword subtask than the original. For example, students in grade 3 read 7.3 CNWPM and received an average of 26.7% correct on the revised nonword subtask, but only 4.2 CNWPM and 16.0% correct on the original subtask. In addition, 44.5% of all students received a zero score on the original nonwords subtask, whilst only 28.9% received a zero score for the revised subtask. The revised subtask's fluency scores also had a slightly higher correlation coefficient with ORF (.69, compared to .65).

However, whilst students did perform better on the revised nonwords subtask, **scores were still relatively low for both grades**. As previously mentioned, students in grade 3 read 7.3 CNWPM and received an average of 26.7% correct on the revised nonword subtask, and students in grade 4 read 11.2 CNPWM and received 39.9% of items correct on average. These results do not match with letter sound identification and ORF scores. Thus, the degree to which decoding is an integral part of measuring students' reading skills remains inconclusive.

By province. As with the overall trend, grade 4 students in Province A and Province B outperformed their grade 3 counterparts in every EGRA task, including fluency scores. As shown in Figure 10, grade 4 students in Province B had significantly higher accuracy scores in both letter naming and nonword subtasks than grade 4 students in Province A. However, grade 3 students in Province A had higher accuracy scores in letter naming with modifiers, ORF, reading, and listening comprehension.

Figure 10. 2021 Reading Accuracy Scores by Subtask, Grade, and Province



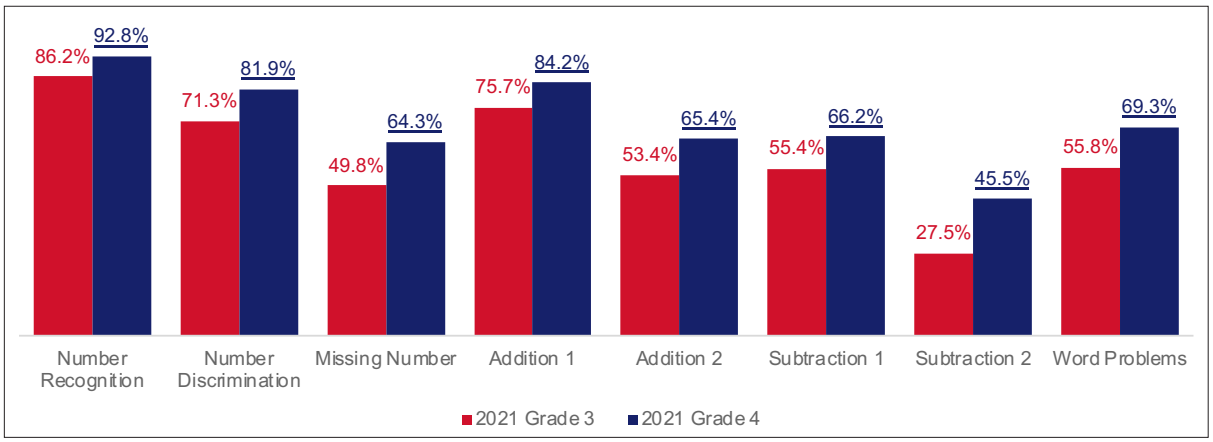
Note: Underlined scores are statistically significant between provinces at p<0.05.

[15] The original subtask used in the 2017 Idarah assessment as well as the 2019 and 2020 Manahel assessments includes several nonwords with sound clusters that are not possible in the Arabic language, and thus was thought to provide inaccurate measures of students' decoding abilities. The new revised subtask was added to test if students performed better on a revised subtask with appropriate items.

MATHEMATICS OUTCOMES

Overall, grade 4 students significantly outperformed grade 3 students in every EGMA subtask. Results are presented in Figure 11 below. Grade 4 boys significantly outperformed their grade 3 peers in every EGMA subtask, including fluency scores. The same trend was seen for grade 4 girls. **Some of the largest gains were missing numbers, subtraction 2, and word problems – subtasks that measure more complex mathematics skills.**

Figure 11. 2021 Mathematics Accuracy Scores by Subtask and Grade



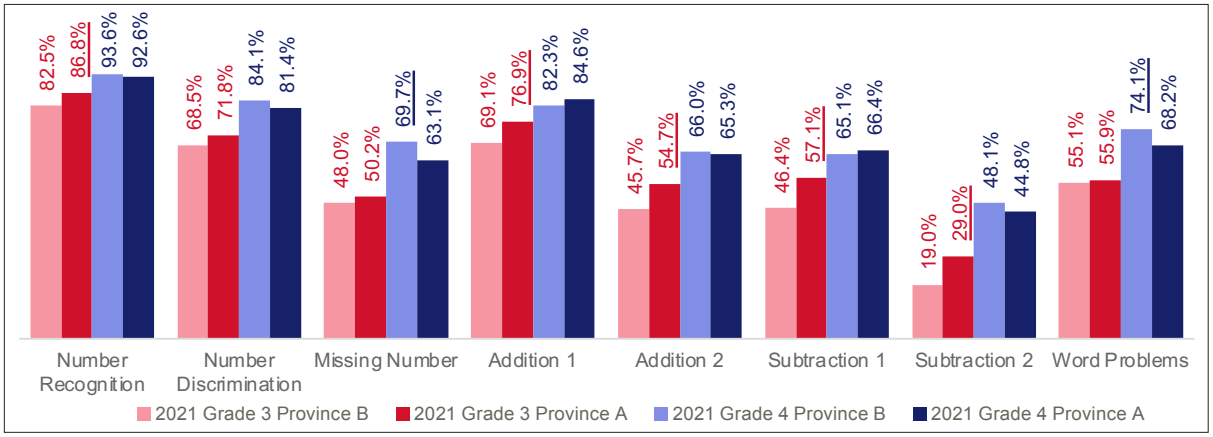
Note: Underlined scores are statistically significant at p<0.05.

On the EGMA, there was a significantly lower proportion of grade 4 students receiving zero scores in all subtasks, except number recognition and number discrimination. Only a single student from each grade received a zero score in number recognition. Two grade 3 students received a zero score in number discrimination, and three grade 4 students received a zero score. Results are presented in

Annex E: Disaggregated Results.

By province. In Province A and Province B, grade 4 students outperformed their grade 3 counterparts in every EGMA task, including fluency scores (see Figure 12 .) **In grade 3, students in Province A outperformed their counterparts in Province B, but it was the reverse in grade 4.** In grade 3, students in Province A had significantly higher accuracy scores than grade 3 students in Province B in number recognition (86.8% compared to 82.8%), addition 1 (76.9% compared to 69.1%), addition 2 (54.7% compared to 45.7%), subtraction 1 (57.1% compared to 46.4%), and subtraction 2 (29.0% compared to 19.0%). In grade 4, students in Province B had significantly higher accuracy scores than grade 4 students in Province A in missing number identification (69.7% compared to 63.1%) and word problems (74.1% compared to 68.2%).

Figure 12. 2021 Mathematics Accuracy Scores by Subtask, Grade, and Province



Note: Underlined scores are statistically significant between provinces at p<0.05.

Research Question 4: Gender Gap

Research Question 4: To what extent is there a gender gap in reading and mathematics performance amongst this year's grade 3 and grade 4 students, respectively? Does the gender gap widen or narrow from grade 3 to grade 4?

This research question aims to understand the differences in boys' and girls' performance for students in grades 3 and 4 in 2021.

READING OUTCOMES

In 2021, grade 3 girls significantly outperformed boys in almost all EGRA skills and continued to do so in grade 4. Differences between boys' and girls' accuracy scores were statistically significant for all subtasks in grade 3 (except listening comprehension) and in grade 4 (except for the original nonword subtask and listening comprehension). Table 5 shows boys' and girls' accuracy scores by grade and the difference between boys' and girls' scores.

For the majority of subtasks, there was a more considerable difference between boys' and girls' scores in grade 4. This indicates a gap in reading skills along the entire skill spectrum (from most basic skills to most advanced) that persists as students progress from grade 3 to grade 4. While the gap in letter sound and nonword subtasks remains relatively constant between grades 3 and 4, the gap between boys' and girls' scores in ORF and reading comprehension grows. In grade 3, the difference between girls' and boys' ORF scores was 5.8%, while it grew to 9.1% in grade 4. Similarly, the difference between girls' and boys' reading comprehension scores was 6.3% in grade 3 and 9.1% in grade 4.

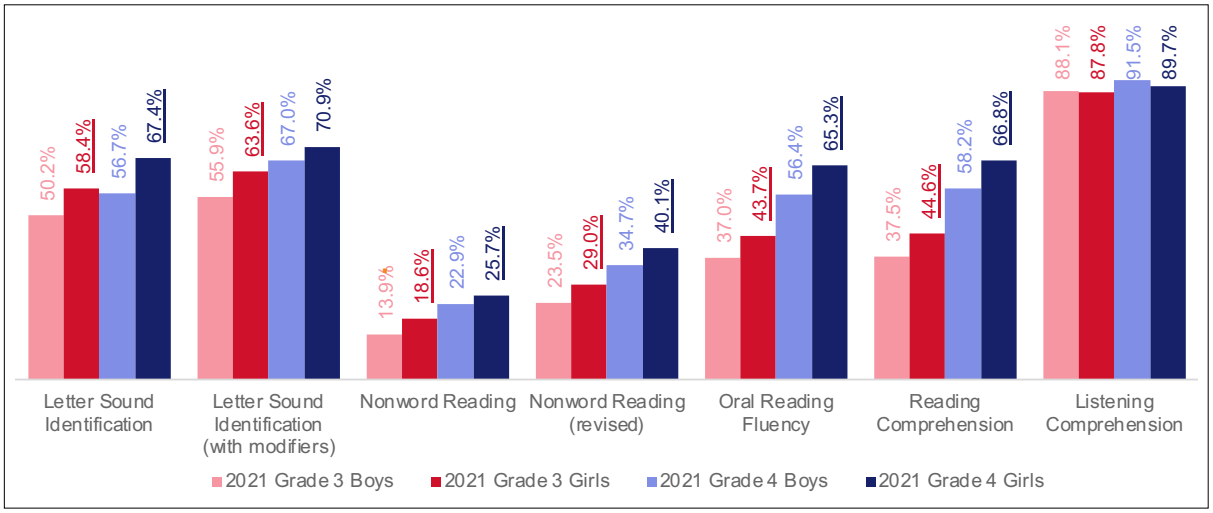
Table 5. 2021 Reading Accuracy Scores by Grade and Sex

Subtask	Grade 3			Grade 4			Bigger Gender Gap in...
	Boys	Girls	Difference (Girls - Boys)	Boys	Girls	Difference (Girls - Boys)	
Letter Sound Identification*†	50.1%	58.3%	8.2%	58.8%	68.7%	9.9%	Grade 4
Letter Sound Identification*† (with modifiers)	54.8%	62.1%	7.4%	68.1%	73.0%	4.9%	Grade 3
Nonword Reading*	14.0%	18.1%	4.2%	24.9%	28.2%	3.2%	Grade 3
Nonword Reading*† (revised)	24.4%	29.0%	4.6%	36.6%	43.0%	6.4%	Grade 4
Oral Reading Fluency*†	35.6%	41.4%	5.8%	56.2%	65.3%	9.1%	Grade 4
Reading Comprehension*†	36.1%	42.4%	6.3%	57.6%	66.7%	9.1%	Grade 4
Listening Comprehension	87.5%	86.9%	-0.6%	91.2%	89.6%	-1.6%	Grade 4

Note: An asterisk (*) indicates the difference in grade 3 boys' and girls' scores was significant at p<0.05.

By province. As with the overall trend of girls outperforming boys, **results point to a gender gap in Province A that exists in grades 3 and 4 in reading**. Girls in Province A in grade 3 had significantly higher accuracy scores in every subtask (see Figure 13) and significantly higher fluency in both letter sound subtasks and the original nonwords subtask. Grade 3 girls in Province A identified 33.3 CLNPM (compared to 292.3 for boys), 36.4 CLNPM with modifiers (compared to 31.7 for boys), and 4.8 CNWPM (compared to 3.7 for boys). In grade 4, girls had higher accuracy scores in letter sounds (without modifiers), nonwords (revised), ORF, and reading comprehension.

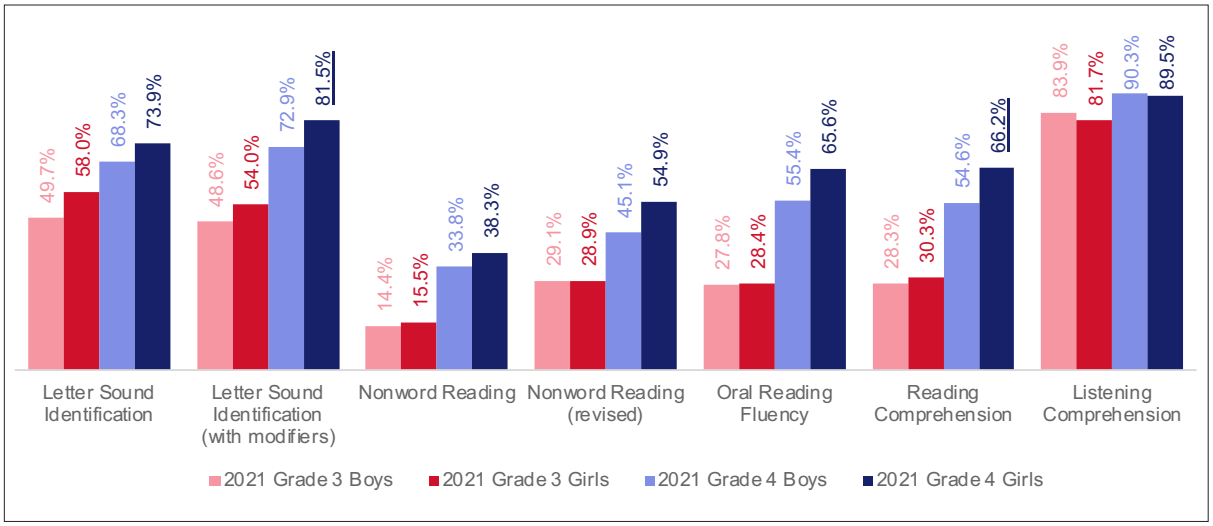
Figure 13. 2021 Reading Accuracy Scores in Province A by Sex



Note: Underlined scores are statistically significant at p<0.05.

In Province B, there were no statistically significant differences between boys' and girls' fluency or accuracy scores for any EGRA subtasks in grade 3, as shown in Figure . **Thus, both sexes performed at comparable reading levels in grade 3. However, grade 4 girls had significantly higher accuracy scores than boys in letter sounds with modifiers and reading comprehension.** They also had significantly higher fluency scores in letter sounds with modifiers (48.6 CLSPM compared to 42.2 for boys), nonword fluency (revised; 16.0 CNWPM compared to 12.3 for boys), and ORF (34.3 CWPM compared to 26.4 for boys).

Figure 14. 2021 Reading Accuracy Scores in Province B by Sex



Note: Underlined scores are statistically significant at p<0.05.

MATHEMATICS OUTCOMES

Boys had higher scores than girls in almost all mathematics subtasks in grades 3 and 4. In grade 3, differences were statistically significant for fluency and accuracy scores in all subtasks except for addition 1 accuracy, as outlined in Table 6. Additionally, a significantly higher proportion of girls received zero scores in subtraction 2. **However, in grade 4, fewer of the differences in boys' and girls' mathematics scores were statistically significant.** Grade 4 boys and girls had comparable scores in number recognition fluency and accuracy, missing number identification accuracy, addition 2 accuracy, subtraction 1 and 2 accuracies, and word problem accuracy. There were no significant differences in zero scores between grade 4 boys and girls in any subtasks.

The gaps between boys' and girls' scores were greater in grade 3 for every single subtask, indicating that girls in grade 4 were catching up to their male peers in mathematics.

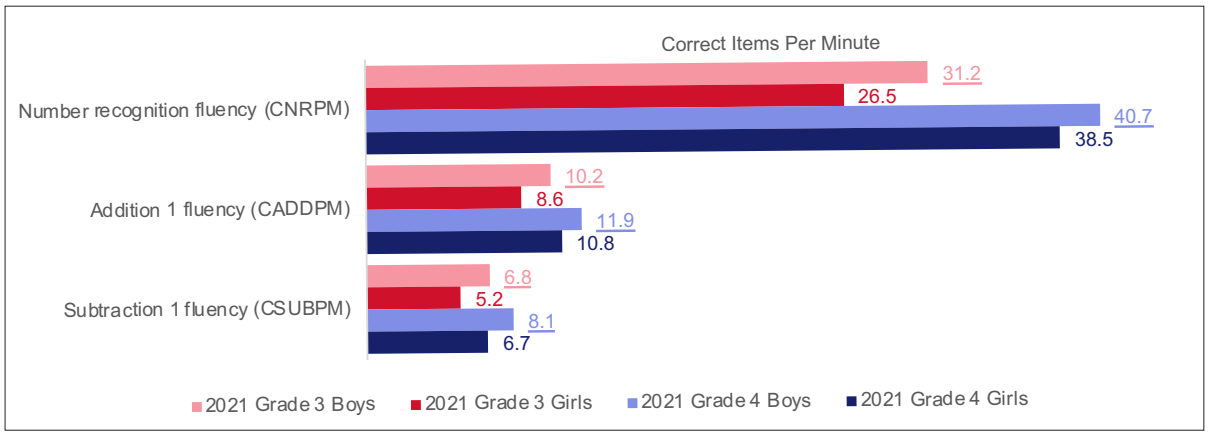
Table 6. 2021 Mathematics Accuracy Scores by Grade and Sex

Subtask	Grade 3			Grade 4			Bigger Gender Gap in...
	Boys	Girls	Difference (Boys - Girls)	Boys	Girls	Difference (Boys - Girls)	
Number Recognition*	88.0%	84.3%	3.7%	93.7%	91.9%	1.8%	Grade 3
Number Discrimination*†	75.9%	66.6%	9.3%	84.9%	79.1%	5.8%	Grade 3
Missing Number*	51.7%	48.0%	3.7%	65.0%	63.7%	1.3%	Grade 3
Addition 1†	77.4%	74.1%	3.3%	85.5%	82.9%	2.6%	Grade 3
Addition 2*	56.0%	50.7%	5.3%	66.3%	64.6%	1.7%	Grade 3
Subtraction 1*†	59.7%	51.1%	8.6%	69.3%	63.2%	6.1%	Grade 3
Subtraction 2*	33.5%	21.3%	12.2%	46.4%	44.6%	1.8%	Grade 3
Word Problems*	58.7%	52.8%	5.9%	70.7%	68.0%	2.7%	Grade 3

Note: An asterisk (*) indicates the differences in grade 3 boys' and girls' scores were significant at p<0.05. An obelisk (†) indicates the differences in grade 4 boys' and girls' scores were significant at p<0.05. .

By province. Overall, results show that in Province A, the gender gap in mathematics persists between grades 3 and 4. Grade 3 boys had significantly higher fluency scores than girls in all subtasks, as shown in Figure . Boys continued to have statistically significantly higher fluency scores in all subtasks in grade 4. Additionally, in grade 3 boys had significantly higher accuracy scores in number recognition (88.2% compared to 85.4% for girls), number discrimination (74.9% compared to 67.5% for girls), subtraction 1 (61.1% compared to 52.9% for girls), subtraction 2 (35.0% compared to 22.9% for girls), and word problems (58.9% compared to 52.8% for girls). As with fluency, boys in grade 4 had significantly higher accuracy scores in number discrimination (84.2% compared to 78.7% for girls), subtraction1 (69.5% compared to 63.4% for girls), and word problems (70.5% compared to 65.9% for girls).

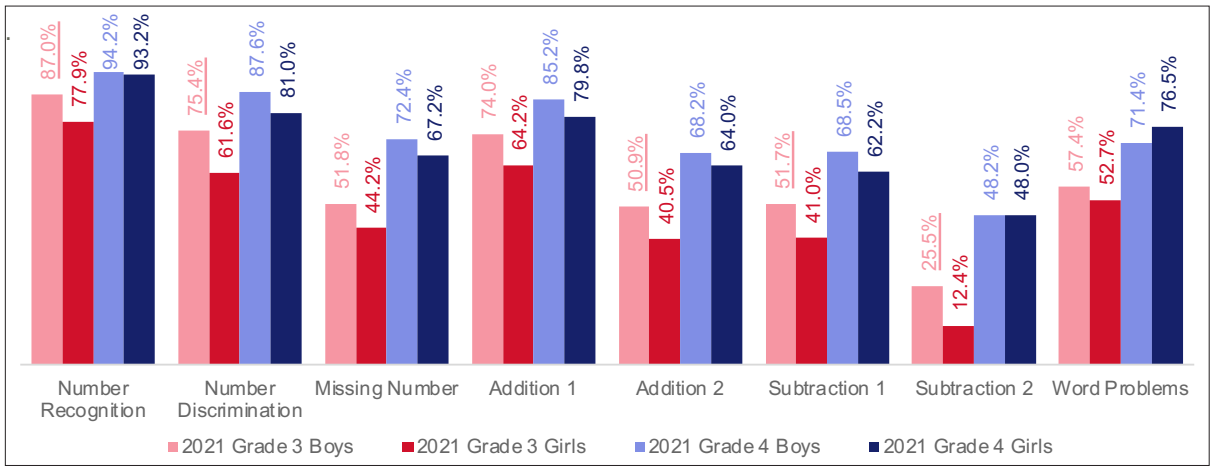
Figure 15. 2021 Mathematics Fluency Scores in Province A, by Sex and Grade



Note: Underlined scores are statistically significant between sex at p<0.05.

In Province B, the gender gap in mathematics was closing as students progressed in their learning. Boys in grade 3 had significantly higher accuracy scores than girls in grade 3 in number recognition, number discrimination, addition 2, and subtraction 1 and 2 (see Figure 16.) However, these differences did not appear in grade 4, where girls and boys had comparable fluency and accuracy scores in all subtasks.

Figure 16. 2021 Mathematics Accuracy Scores in Province B, by Sex and Grade



Discussion

This section presents key findings from the results of all four research questions and discusses hypotheses around these findings.

READING

The critical reading subtest results (oral reading fluency and reading comprehension) showed steady improvement in the proportion of students who were attaining reading proficiency from 2019 to 2020 to 2021. It is also significant that the trajectories for student performance in these two subtests track each other very closely.

As in the 2020 Manahel assessment, results showed girls outperforming boys in all reading subtasks (except listening comprehension) in 2021 in grades 3 and 4. Results also point to the difference between boys' and girls' literacy outcomes expanding as students transition from grade 3 to grade 4. Previous assessments point to external societal factors causing these results, including factors pulling younger boys out of school to work and keeping older girls from progressing through their education. The Manahel programme has made several efforts to address these factors, though such societal factors cannot be mitigated entirely by a single intervention.

Data indicates a gap in reading skills along the entire skill spectrum (from most basic skills to most advanced) that persists as students progress from grade 3 to grade 4. However, the fact that boys and girls perform at similar levels in grade 3 may also be a testament to the efforts to further support Province B students in the 2020/21 school year. The differences between boys and girls in grade 4 indicate that the gender gap may be returning in Province B in higher grades.

There were no significant differences between the reading proficiency classification of students in grade 3 in 2020 and students in grade 3 in 2021. In 2020, 21.7% of students met the reading proficiency benchmark of scoring 80% or higher on reading comprehension compared to 25.4% in 2021. This difference was not statistically significant.

Results indicate that grade 4 students performed at a significantly higher reading levels than grade 3 students indicating that students in higher grades have greater literacy competencies, as would be expected. Especially notable were the differences in scores in ORF and reading comprehension. These improvements were especially pronounced in

ORF and reading comprehension, where accuracy scores improved from 38.2% in 2020 to 62.2% in 2021 in ORF; and 40.0% correct to 62.2% correct in reading comprehension. The grade 4 progress in these areas points to the effectiveness of the assistance that Manahel interventions may have had in supporting learning gains during a school year of continued disruptions. The questions posed in the findings around the challenges of nonword tests in Arabic and the lack of predictive nature for the nonword tests talks to the value of including nonwords in future EGRA tests.

Comparing results by the province in each grade yields interesting insights regarding differences between grade 3 and grade 4 students in 2021. As in previous years, grade 3 students in Province A had higher scores in all reading subtasks (except the revised nonwords subtask) compared to grade 3 students in Province B – where there were no significant differences between boys' and girls' reading scores in grade 3, but girls began to outperform boys in grade 4. By that stage, grade 4 students in Province B had higher scores than their peers in Province A. As all other variables (operational school days, student attendance and drop out, teacher attendance and turnover) were similar between Province B and Province A it is safe to argue that these findings point to the effectiveness of the targeted aid provided to schools in Province B province during the 2020/21 school year and summer of 2021. In Province B, this may have been compounded by the change in the partners that Manahel was working with in that province.

In summary, in both 2021 and 2020, proficient readers were more likely to come from Province A than Province B; however, in 2021, reading scores in Province B improved significantly. In 2020, Province B had a significantly higher proportion of non-readers than Province A (32.9% compared to 19.4%, respectively). However, in 2021, there were comparable proportions of non-readers in Province A and Province B, but a significantly higher proportion of beginning readers in Province B (54.9% in Province B compared to 40.5% in Province A).

In addition to the gains from an extra year of schooling, several programme-related factors may have contributed to the large increases in ORF and reading comprehension from grade 3 students in 2020 to grade 4 students in 2021. First, following the 2020 Learning Assessment Manahel instituted summer clubs and after school activities. This additional support during periods of school closure likely contributed to outcome gains. Second, reading comprehension may be easier to teach remotely – assuming that students have gained the basic principles of decoding. Discussions with the Manahel team during the 2020 assessment indicated that teachers felt comprehension skills were more likely to improve during remote learning compared to other reading skills. Again, this is based on students who can already decode and have a growing familiarity with the skill of reading text as they progress from grade 3 to grade 4, especially during school closures where teachers might rely on such activities more.

The effort and concentration in encouraging students in Province B after the worrying performance in the 2020 assessment seems to have had some success in bringing their performance on par with students in Province A. In 2020 grade 3 students in Province A had higher scores than their peers in Province B: these differences disappeared amongst grade 4 students in 2021. These results indicate that the activities Manahel implemented to reach students falling behind in 2020 helped them grow in their learning. This also seems to be the case with struggling readers and particularly boys, illustrated by a drop in zero scores among students in Province B.

The programme's adaptations to teaching pedagogy for remote learning continue to appear to be successfully implemented by Manahel teachers in support of their classroom lessons. Results remain comparable between the 2020 grade 3 cohort and the 2021 grade 3 cohort. Grade 3 students in 2021 have had a notable amount of time without in-person schooling. They would have started grade 1 in 2019; had schools close in the spring of 2020 (grade 1); had several disruptions of in-person schooling in grade 2 starting in the autumn of 2020; and had schools closed for two weeks at the start of the 2021/22 school year. By contrast, grade 3 students in 2020 would have experienced nearly two years of uninterrupted schooling before the COVID-19 pandemic. Results from the 2020 assessment showed that Manahel's efforts to adapt to remote learning seemed to mitigate learning loss successfully. Efforts to encourage struggling readers, such as boys and students in Province B, seem to have succeeded in improving students' proficiency to some degree.

MATHEMATICS

Maths results between grade 3 and 4 students in Provinces A and B showed similar trends as reading with steady improvement from grade 3 to 4 especially in the higher-level skills of addition, subtraction, and word problems. This indicates that the extra year of learning and support students in grade 4 received helped them better grasp these skills. Targeted assistance by the project in higher-level mathematics skills may have contributed to these improvements. However, it is impossible to tell if students in grade 4 are performing at the level they should be given that the grade 4 students sit a grade 3 level EGMA.

Also, similar to the situation with literacy, Province B students trailed those of Province A in grade 3 but had caught up by grade 4. Grade 3 students in Province A achieved higher scores in all mathematics subtasks, significantly so for number recognition, addition 1, addition 2, and subtraction 1 compared to grade 3 students in Province B. However, grade 4 scores were comparable between the two provinces, with students in Province B significantly outperforming students in Province A in missing number and word problems.

Data from the 2021 assessment show that boys in grade 3 outperformed girls in all mathematics subtasks (except addition 1), but grade 4 boys and girls performed number recognition, missing numbers, addition 2, subtraction 2, and word problems comparably. The gaps between boys' and girls' scores were greater in grade 3 than grade 4 for every single subtask. Thus, the gender gap in mathematics is narrowing as students progress through the grades: girls in grade 4 are catching up to their male peers in mathematics but need extra support to achieve parity in learning outcomes. The narrowing of the gap may be due to programme efforts to encourage girls in mathematics.

The gender gap in mathematics is especially prevalent in Province A, where boys had statistically significantly higher scores in most subtasks into grade 4 than girls and significantly higher accuracy scores in number discrimination, subtraction 1, and word problems. Unlike Province A, grade 4 girls in Province B performed on par with their male peers with no significant differences in accuracy scores on any mathematics subtask. This trend was also seen in the 2020 assessment, where girls in Province B had a strong performance in mathematics.[16] Improvements in mathematics scores in Province B between grades 3 and 4 indicate that efforts made through the 2020/21 school year have successfully supported students in this province to catch up with their peers in Province A.

[16] See 'Syria Education Programme: 2020 Manahel Learning Assessment Report', February 2021.

Conclusion and Recommendations

This section uses the analysis from the last section to draw conclusions and make recommendations and to provide next steps based on this evidence.

The key conclusion is that despite COVID-19 disrupting schooling regularly for students along with learning time loss through the COVID-19 related implementation of double and triple shifts – and particularly for those students who reached grade 3 in 2021 – they have continued to progress and have not fallen behind where their peers were in 2019. This is a significant achievement as school closures across the world during the pandemic have impacted negatively on student performance. This progress will be testament to a number of factors and will include the remote learning interventions that Manahel put in place early on in the pandemic in 2020 and which are now used to supplement in-class teaching.

The report also concludes that although reading and maths scores have held up reasonably well and generally seen improvement over the assessments in 2021, there is still a lot of work to be done to bring the literacy and numeracy scores up to a level that will provide the majority of NWS students with the foundations that they require for their future education and work lives.

The recommendations are structured around the level at which they should be implemented: school, system and by Manahel.

SCHOOL RELATED RECOMMENDATIONS

1. Reader profiles were relatively similar between 2020 and 2021 (but a significant improvement on those of 2019), indicating that teachers are effectively implementing remote learning techniques developed over the past two years, as well as implementing agreed classroom teaching practices. However, the marginal improvements seen year on year indicate that as schools face a post-covid future

with reduced conflict (in most areas) the teaching of literacy and numeracy in the early grades -from grade 1 – needs to be a stronger focus and teachers need to use continuous assessment more strategically to analyse the status of their students and then implement targeted remedial actions to remedy.

2. Teachers should track and support non-readers more closely both in class and if and when schools revert to remote learning. This will support students most at risk of not gaining basic and higher-level reading skills that eventually enable reading fluency and comprehension.

3. Through to the end of the programme, teachers should work with boys to build reading fluency and their foundation towards reading proficiency. In grade 3, this focus should be on building the basic skills of letter sound knowledge and decoding skills needed to attain fluency and comprehension to ensure students have solid foundations on which to build later. In grade 4, support should focus on more advanced fluency and comprehension to ensure that students are prepared for the transition to higher grades, where they are more vulnerable to drop out.

4. Teachers in grades 1 and 2 should make sure that all learners have understood the basic mathematical functions (addition/subtraction) whilst teachers in grade 3 should focus on more complex mathematics skills to ensure that students master mathematics operations and real-world thinking, and so are better prepared for the more complex maths taught in grade 4.

SYSTEM RELATED RECOMMENDATIONS

5. All of the above school-based recommendations will be more successful if supported from within the system. In particular, the system actors should assist schools in interpreting and analysing their continuous assessment results in the early grades to inform their teaching and to build remediation measures around the areas of weakness.

MANAHEL RELATED RECOMMENDATIONS

6. Manahel should work with schools following assessments (both internal continuous assessments and EGRA/EGMA) to assist the teachers in using the test data to inform remediation efforts. This will require Manahel to train teachers in how to analyse their students' assessment results to inform remediation and how to use the summer school clubs and after school lessons to maximum effect.

7. Manahel should extend targeted services to grade 3 students in Province B through the end of the 2021/22 school year and consider providing after school literacy clubs to students in grade 2. The Manahel programme might also consider working with teachers to emphasise foundational reading skills such as letter sound identification and decoding in Platform A, where students in grade 4 received significantly lower scores than their peers in Platform B. Equally, it appears that grade 3 students in Platform B need extra assistance in reading.

8. Manahel should remove nonword subtasks in future EGRAs. Research on measuring reading in Arabic indicates that decoding may not contribute to reading comprehension because of the nature of the language.[17, 18] Thus, it is unsurprising that nonword fluency and accuracy scores remained relatively low. This confirms other research that shows that nonword tests are not a good predictor of learner reading performance, unlike letter sound identification, ORF and reading comprehension.

9. Results clearly show gains in reading between grades 3 and 4 for boys and girls in Provinces A and B. To sustain these gains through the end of the programme, the Manahel team should continue to help teachers increase the amount of time spent on reading with the help of online tools and targeted interventions for non-readers.

10. Manahel intends to create girl-focused after-school centres and will measure grade 5 and 6 girls' reading and mathematics outcomes. Based on these results, the programme could monitor grade 4 girls' mathematics performance, especially in Province A. This would catch indications of the mathematics gender gap early on and allow the programme and teachers time to work with girls who struggle to match boys in their numeracy outcomes.

11. Manahel should observe male grade 2 and 3 teachers in maths classes – particularly in those Province A schools where girls' performance in maths is particularly poor – to analyse their interaction with both boys and girls to see if the actions and bias of the teachers are related to female under-performance. This should focus on who is being asked questions, who is speaking in small group work, who is coming up to the board, whose work is being celebrated etc.

12. Concerns about the performance of boys, particularly in grades 3 and 4, has led to Manahel focusing on tracking and monitoring their attendance at school. This initiative could be extended to monitor dropouts from school, although this is complicated by the mobility of students between schools and regular student absenteeism.

13. As it seems students in Province B are pulling ahead in mathematics results, Manahel needs to analyse what elements of the intervention in Province B had impact on learner performance and replicate these in the regions of Province A where students appear to be falling behind in relative terms, such as District 1, where there has been a notably high level of conflict over the past year.

---

[17] Arabic is a diglossia language, meaning it has 2 variants for different situations. In this case, the first variant is Modern Standard Arabic (MSA), used for reading and writing, while another variant is the spoken colloquial dialect which can differ significantly from MSA. The simple view of reading (SVR) model, on which the EGRA is based, explains reading comprehension as the product of decoding (the ability to apply knowledge of letter-sound relationships, including knowledge of letter patterns, to correctly pronounce written words) and listening comprehension. However, the validity of SVR for Arabic has not been tested.

[18] Asadi, Ibrahim A., Asaid Khateb, and Michal Shany. 'How simple is reading in Arabic? A cross-sectional investigation of reading comprehension from first to sixth grade.' Journal of Research in Reading S1, no. 40 (2017): S1–S22. https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-9817.12093

Recommendations for Future Research and Learning

While EGRA and EGMA testing will continue through the time that remains for the project, the research and learning agenda should focus on better understanding of what has worked at school and classroom level. The findings from these deep dives should be shared with EDs, school head teachers, education NGOs and other stakeholders so that the lessons can be learned and, where possible, changes made.

LEARNING ASSESSMENT (EGRA/EGMA)

These will be conducted at two points in time. The first will be in November 2022 and the second one will be in May 2023 which will also serve as an endline measurement for the life of the project. Manahel proposes to administer the EGRA/EGMA in a random sample of schools which Manahel has supported but in which teachers are not being paid by the project and a randomised sample of schools where teachers are being paid. This would provide some measure of understanding of how providing teacher pay impacts on learner performance.

Beyond the scope of this project a scientific process is required to develop grade-specific annual and semester-based reading and numeracy benchmarks and cutscores. This would allow a much more accurate understanding of the proportion of students who are achieving at the expected level for their grade and age and to better track student progress towards those benchmarks.

IMPORTANCE OF TEACHER PAY

This research has been completed and is being further analysed to better answer questions about the relationship between payment of teachers and learner performance.

SCHOOL CASE STUDIES

Manahel intends to develop case studies of schools where teachers are being paid and those where they are not being paid as part of the intervention. This would allow a deep dive to better understand the triangular relationship between learners' performance, teachers' payment, and teachers' commitment.

POSITIVE DEVIANCE STUDIES

Manahel intends to supplement the research and understanding with detailed case studies of individual project schools which have seen a robust improvement of learner results and/or are maintaining high levels of learner performance in EGRA and EGMA in conditions where other schools are failing to do the same, to better understand the conditions that lead to improved and sustained learner performance.

STUDENT GENDER AND VULNERABILITY RESEARCH

Three pieces of work are proposed with a focus on gender and disability:

1. A study to explore attendance by girls in the early grades and/or attendance of children with disabilities in the early grades over time in unsupported schools (to test the assumption that the weight of supporting the payment and support of teachers by parents falls disproportionately on parents of girls and children with disability) and compare that to attendance of these two groups in supported schools using a case study approach.

2. Small-scale research to understand if girls' well-being is comparable to boys in the later years and widen the time-on-task/lesson observation work to a small number of upper primary teachers (approximately 30) to see if there is a discernible difference in teaching. Manahel will prioritise introducing learning circles to support teachers to create gender-responsive pedagogy and a growth mindset and measure how these interventions are perceived.

3. Manahel is planning a GESI review during the extension period and reflecting on improved GESI focused activities that can be applied during the extension period as well as making recommendations for future programming.

Annex A: Study Terms of Reference

The subcontractor shall use the same EGRA and EGMA tools administered under the 2019 and 2020 assessments, including both the letter identification EGRA subtasks: the first without complex Arabic modifiers and the second with a scaled-down use of Arabic modifiers. Pending further discussions with Manahel, the pseudowords subtask may undergo limited, targeted changes made by Manahel staff with advice from the subcontractor. Retaining these tools allows comparability across timepoints. No changes are anticipated for the EGMA tool.

Supporting surveys will also be reviewed and adapted or created to address the final research questions. These surveys may include a student survey and a brief head teacher survey. The student survey will ask general questions about time out of school, absenteeism, and will also collect some household information as well as information on student levels of displacement in addition to student knowledge of and participation in Manahel activities. The head teacher survey will collect information about student enrolment and attendance for weighting.

In academic year 2021/22, Manahel is supporting 435 schools. The study design will again include 2 grades.

	Number of schools sampled	Proportion of total schools supported	Number of students sampled	Margin of error for grade 3[19]	Margin of error for grade 4[20]
2021 Sample	75	17.2%	1500 total (750 in grade 3, 750 in grade 4)	±2.9	±5.6

A two-stage sampling approach will be used, beginning with a sample of schools drawn from a full list of intervention schools (sampling frame) and then a sample of 10 students selected from those schools on the day of data collection.

A sampling frame that includes these data on all Manahel intervention schools will be required to determine a) if all strata can be accomplished in the sampling strategy and b) to determine the appropriate sample size for the study.

The sample for the study will be determined based on the parameters above, as well as a consideration of the level to which results will be generalised. Additionally, logistic details and challenges to data administration in a conflict-affected context may also affect the sample. To mitigate some of the expected challenges in a conflict-affected

[19] Grade 3 Margin of Error considers the midline grade 3 ICC of 0.21 and 21.7% of the population attaining the reading benchmark.

[20] Grade 4 Margin of Error considers the midline grade 4 ICC of 0.15 and 46.0% of the population attaining the reading benchmark.

context such as Syria, the sampling approach will include a thorough replacement strategy for both schools and student-level sampling.

The final sample of schools, classes, and students will be determined based on consultation with the Manahel team. The subcontractor and Manahel will train data collectors on the use of appropriate sampling strategies to reach the desired number of respondents according to the selected sampling plan.

The subcontractor shall be responsible for training Manahel staff to conduct the operational EGRA and EGMA data collection in Provinces A and B. The subcontractor will create training materials and will facilitate a Master Trainer training in October 2021. Trainings will be provided in English. Manahel staff who have been trained in EGRA/EGMA will be available to translate and supplement technical input. The training will be conducted remotely by WebEx or similar: the two Manahel Master Trainers are Syria-based and not able to travel to Turkey. The Master Trainers will in turn train Manahel field staff in Syria to serve as enumerators for a November/December 2021 data collection. The Master Trainers will train 20 enumerators for three days. The subcontractor will support the Master Trainers throughout the enumerator training, answering questions and troubleshooting as necessary.

Manahel staff will also be online and available to answer questions. Topics for the enumerator training will include:

- An orientation to the EGRA, EGMA, and their subtasks
- Protocols for administering the EGRA, EGMA, and surveys
- Electronic data collection techniques
- Data quality assurance measures
- Child safeguarding considerations and accountability to the affected population
- COVID-19 health and safety precautions

Enumerators will be assessed on their accuracy to ensure reliability of the results.

Manahel encourages, but does not require, STS to pre-record any key elements where fidelity might be lost in step-down training (or to work with the relevant Manahel staff so they can pre-record in Arabic).

During the data collection process, the subcontractor shall remotely supervise data collection in conjunction with Manahel staff. The subcontractor and Manahel will maintain detailed documentation of all issues encountered in tracker which will be used in the data cleaning process. Additionally, electronic data capture via tablets will contribute to data quality, consistency, and collection efficiency by streamlining fieldwork and reducing measurement and data entry errors.

Each day, data will be uploaded from the tablets via Wi-Fi to the Tangerine® server and then downloaded and stored securely on a password-protected server for cleaning, review, and analysis using Excel and SPSS. Using a data collector tracker and school visit forms, data will be cleaned based on pre-determined criteria: time and date inconsistencies, consent checks, and survey sessions timing. As with the 2019 and 2020 assessments, the subcontractor will also conduct daily enumerator scoring checks and provide feedback to enumerators as needed. These checks consist of comparing the scoring of two enumerators assessing the same child. This comparison ensures that enumerators are scoring students consistently and is key to quality data collection.

The subcontractor will be responsible for the data cleaning and analysis of the primary EGRA, EGMA, and survey data sets. Three levels of data cleaning will be conducted in December 2021 to ensure that the data is complete, accurate, and internally consistent. The subcontractor will follow standard best practices for cleaning and finalising data, including developing and providing a master codebook, as well as merging or appending data files where possible for easier use and manipulation. Disposition codes will be applied to categorise the various issues or problems that emerged in the data collection process as well as in the datasets. These disposition codes will be used to determine cleaning rules which will be incorporated into the database using the syntax to clean the data accordingly.

The subcontractor will produce a brief final report that answers the stated research questions. As part of this reporting, mean differences by subtask, subgroups, and student factors will be explored to describe trends in the results. The subcontractor will present preliminary findings in a Data Dive in early to mid-December. In this Data Dive, the subcontractor will also gather feedback on further directions for analysis, contextual information, and preliminary recommendations from the Manahel team as well as FCDO. A draft outline of the report will be shared with Manahel for initial feedback in early December 2021. The full draft report will be submitted in January 2021, allowing for Manahel's review and feedback before the submission of the final report by early February 2021.

Annex B: Evaluation framework

In 2021-2, School-to-School International will conduct a study to measure student learning outcomes in reading and mathematics in a sample of Manahel's intervention schools to answer the research questions outlined below. The results of the 2021 study will be compared to the 2019 and 2020 results, as well as serve as a point of reference for comparison to future student learning outcomes in reading and mathematics.

1. What proportion of G3 students in 2020 and in 2021 are classified as 'progressing' and 'proficient' readers?

   The purpose of this RQ is to measure Manahel progress against the Impact Indicator and to see the percentage of students who can read in comparison to previous years to respond to the log frame indicator.

2. How have last year's cohort progressed in reading and maths outcomes (Grade 3 in 2020, now in Grade 4)[21]?

   The purpose of this RQ is mainly to track growth within a cohort across an additional year of Manahel interventions and to measure students' progress in reading and math and to provide Manahel with insights to ensure the project meets the needs of the schools and students it serves.

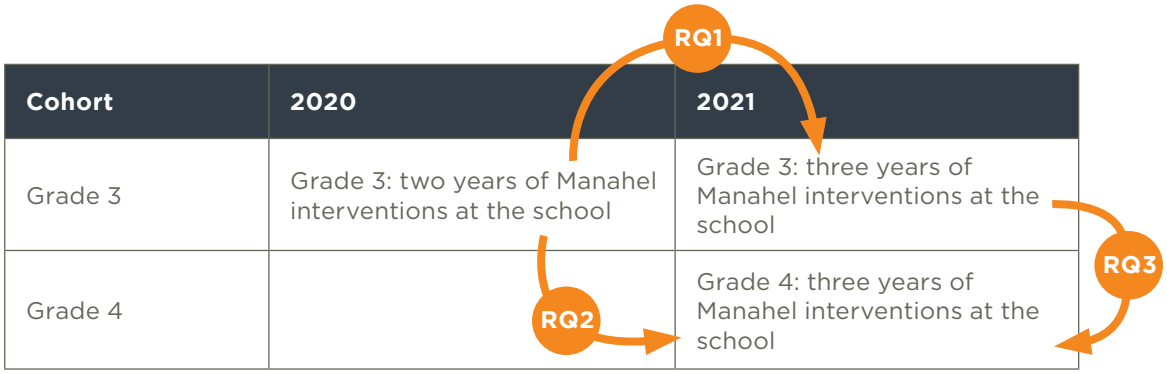
3. How do this year's Grade 4 students compare to this year's Grade 3 students? This RQ is to serve as a proxy comparison group. It is also useful to see the additional learning in another year of Manahel intervention.

4. To what extent is there a gender gap in reading and math performance among this year's Grade 3 and Grade 4 students respectively? Does the gender gap widen or narrow from Grade 3 to Grade 4?

   The purpose of this RQ is to understand differences in performance based on gender and to examine gender-based differences in performance and identify any gaps in performance related to gender.

The 2021 data can be compared to 2020 and 2019 data since the EGRA/EGMA tools are the same (with the addition of a new subtask in 2020), and students are assessed at the same timepoint in the school year (i.e., beginning of Grade 3). However, assumptions regarding comparability of cohorts on other factors, such as conflict or impact of COVID-related closures in 2020 are outside the scope of the study. Any comparison of EGRA/EGMA scores across years implies comparability on these factors. The graphic below illustrates the utility of comparisons across time and grades.

[21] Note that this study will not track individual students from previous studies.



Cohort	2020	2021
Grade 3	Grade 3: two years of Manahel interventions at the school	Grade 3: three years of Manahel interventions at the school
Grade 4		Grade 4: three years of Manahel interventions at the school

The 2021 learning assessment will use the Arabic-language EGRA and EGMA instruments previously administered with Grade 3 and Grade 4 students under the 2019 and 2020 Manahel studies. Table B.1 below illustrates this proposed study design and Table B.2 its alignment with Manahel's required indicator reporting.

Table B. 1. Proportion of Grade 3 Students Per Proficiency Band, by Assessment Year

Project intervention grades and years	2017 June (end of school year)	2018	2019 October (beginning of school year)	2020 October (beginning of school year)	2021 October (beginning of school year)
Grade 1					
Grade 2					
Grade 3	Idarah EGRA		Manahel EGRA and EGMA	Manahel EGRA and EGMA	Manahel EGRA and EGMA
Grade 4				Manahel EGRA and EGMA	Manahel EGRA and EGMA

Table B. 2. Proportion of Grade 3 Students Per Proficiency Band, by Assessment Year

IMPACT	Impact Indicator 1	Baseline 2017	Y1 (2018)	Y2 (2019)	Y3 (2020)	Y4 (2021)	Target*
Student resilience and learning outcomes through provision of quality and inclusive formal, and informal learning opportunities	Percentage of students in the top two categories in progressing reader, and proficient reader of early grades students as measured by EGRA results	53%	N/A	25.7%	34%	TBD	44%
		Source					
		Idarah EGRA June 2017, end of Grade 3	Not reported	Manahel EGRA October 2019, start of Grade 3	Manahel EGRA October 2020, start of Grade 3	Manahel EGRA October 2021, start of Grade 3	

The 2017 baseline value is derived from the Idarah report, page 35. The definitions, as outlined in the Idarah report and applied to the 2017, 2019 and 2020 studies, are:

- Non-readers: students who were unable to read a single word of the story reading passage.

- Beginning readers: students who read between 1 and 22 Correct Words per Minute (CWPM) but scored less than 80% on the comprehension subtask.

- Progressing readers: students who read 23 CWPM or more but scored less than 80% on the comprehension subtask.

- Proficient readers: students who scored 80% or more on the reading comprehension subtask.

Equal numbers of boys and girls will be sampled, thus allowing the study to compare learning assessment results by gender. No other disaggregates will be used in the analysis for the 2021 study.

We will compare the learning outcomes of grade 3 students and grade 4 students. This will help us to understand the learning trajectory of students in Manahel-supported schools. While attribution to Manahel will not be possible with this approach, we will be able to demonstrate the amount of learning that occurred in a year.

In academic year 2019/20, Manahel supported 450 schools. One third of these schools took part in the 2019 Learning Assessment for a total of 1,479 students.

In academic year 2020/21, Manahel supported 516 schools (454 in Northwest Syria) with 75 of these schools taking part in the 2020 learning assessment. Because the 2020 study design included two grades but resources allowed for the same number of students (1500), the number of schools in the sample was lower than in the 2019 assessment.

In academic year 2021/22, Manahel is supporting 435 schools[22]. The study design will again include two grades. STS will use the sample outlined below in Table B.3.

Table B. 3. Proportion of Grade 3 Students Per Proficiency Band, by Assessment Year

	Number of schools sampled	Proportion of total schools supported	Number of students sampled	Margin of error for grade 3[23]	Margin of error for grade 4[24]
2021 Sample (same sample size as midline)	75	17.2%	1500 total (750 in Grade 3,750 in Grade 4)	±2.9	±5.6

A two-stage sampling approach will be used, beginning with a sample of schools drawn from a full list of intervention schools and then a sample of 10 students per grade selected from those schools on the day of data collection, sampling boys and girls equally.

[22] As of the time of writing (September 2021).

23 Grade 3 Margin of Error considers the midline Grade 3 ICC of 0.21 and 21.7% of the population attaining the reading benchmark.

24 Grade 4 Margin of Error considers the midline Grade 4 ICC of 0.15 and 46.0% of the population attaining the reading benchmark.

A sampling frame that includes these data on all Manahel intervention schools will be required to determine a) if all strata can be accomplished in the sampling strategy and b) to determine the appropriate sample size for the study.

The sample for the study will be determined based on the parameters above, as well as a consideration of the level to which results will be generalized. Additionally, logistic details and challenges to data administration in a conflict-affected context may also affect the sample. To mitigate some of the expected challenges in a conflict-affected context such as Syria, the sampling approach will include a thorough replacement strategy for both schools and student-level sampling.

The final sample of schools, classes, and students will be determined based on consultation with the Manahel team. STS and Manahel will train data collectors on the use of appropriate sampling strategies to reach the desired number of respondents according to the selected sampling plan.

Annex C: Use and Influence Plan

In order to ensure that findings and lessons learned from the study will be applied to programme implementation and the broader field, Manahel will adhere to the following evaluation use and influence plan.

Table C.1. Evaluation Use and Influence Plan

Deliverable	Timeline	Intended Audience	Anticipated Use
Data Dive	December 22, 2021	Targeted leadership from Chemonics' implementation team	Present initial findings from analysis; programme can make small pivots and adjustments based on results while providing context and feedback to the analysis.
Brief Technical Report	February 2022	Chemonics staff, FCDO staff	Accountability to funder and programme implementer. Reference for methodology used and technical details around findings.
Graphic Summary	February 2022	Chemonics staff, FCDO staff	Present key findings in a digestible way with actionable recommendations. Programme staff and funder can use this deliverable to adapt programming approach as needed.

The FCDO will have unlimited access to all final deliverables produced by STS.

Annex D: Assessment and Survey Tools

Table D.1. Summary of EGRA and EGMA Subtask

Tool	Subtask	Thematic Skill	Purpose	Administration	Scoring
EGRA	Letter sound identification (without Arabic modifiers)	Mechanics of Reading	Alphabet knowledge	Timed – 2 minutes; autostop after first 10 items	Accuracy (% correct) and fluency (Correct letter sounds per minute (CLSPM); 100 items total
	Letter sound identification (with Arabic modifiers)	Mechanics of Reading	Alphabet knowledge	Timed – 2 minutes; autostop after first 10 items	Accuracy (% correct) and fluency (Correct letter sounds per minute (CLSPM); 100 items total
	Nonword reading (original)	Mechanics of Reading	Decoding	Timed – 2 minutes; autostop after first 5 items	Accuracy (% correct) and fluency (Correct nonwords per minute (CNWPM); 50 items total
	Nonword reading (revised)	Mechanics of Reading	Decoding	Timed – 2 minutes; autostop after first 5 items	Accuracy (% correct) and fluency (Correct nonwords per minute (CNWPM); 50 items total
	Oral reading fluency	Mechanics of Reading	Decoding and reading fluency	Timed – 2 minutes; autostop after first 11 items	Accuracy (% correct) and fluency (Correct words per minute (CWPM); 82 items total
	Reading comprehension	Comprehension	Reading comprehension	Untimed; number of questions asked corresponds to how many words read in oral reading fluency passage	Accuracy (% correct); 5 items total
	Listening comprehension	Understanding	Oral language comprehension and vocabulary	Untimed; all questions asked of all respondents	Accuracy (% correct); 6 items total

Tool	Subtask	Thematic Skill	Purpose	Administration	Scoring
EGMA	Number recognition	Whole numbers	Numerals and numericities identification	Timed – 2 minutes; no autostop	Accuracy (% correct) and fluency (correct numbers recognised per minute (CNRPM); 20 items total
	Quantity discrimination	Whole numbers	Numerical magnitudes comparisons	Untimed; autostop after 4 consecutive incorrect items	Accuracy (% correct); 10 items total
	Missing numbers	Whole numbers	Number patterns identification	Untimed; autostop after 4 consecutive incorrect items	Accuracy (% correct); 10 items total
	Addition (level 1)	Operations	Arithmetic skills	Timed – 2 minutes; no autostop[25]	Accuracy (% correct) and fluency (correct addition problems per minute (CADDPM); 20 items total
	Addition (level 2)	Operations	Arithmetic skills	Untimed; no autostop; only administered if respondent correctly answered at least 1 item correct on Addition level 1	Accuracy (% correct); 5 items total
	Subtraction (level 1)	Operations	Arithmetic skills	Timed – 2 minutes; no autostop	Accuracy (% correct) and fluency (correct subtraction problems per minute (CSUBPM); 20 items total
	Subtraction (level 2)	Operations	Arithmetic skills	Untimed; no autostop; only administered if respondent correctly answered at least one item on Subtraction level 1	Accuracy (% correct); 5 items total
	Word problems	Real world problems	Conceptual and real-word mathematics understanding	Untimed; autostop after four consecutive incorrect items	Accuracy (% correct); 6 items total

[25] Additionally, students who did not correctly answer any items on the addition or subtraction level 1 subtasks were not asked items from the corresponding level 2 subtask.

EGRA Tools

2017 Letter Sound Identification (without Arabic Modifiers) – Student Stimulus

			ـي		أمثلة:		ـهـ		عـ	
	١٠	٩	٨	٧	٦	٥	٤	٣	٢	١
١٠	ط	ف	ـه	ـت	ب	ـب	أ	ـن	آ	ق
٢٠	- -ِ -	ـص	ذ	- -ُ -	ـم	ه	- -ُ -	ـس	حـ	أ
٣٠	أ	ع	ك	ؤ	ـن	ـث	إ	ـف	ـل	- -ُ -
٤٠	ـي	- -َ -	ـش	عـ	س	و	آ	ـم	إ	ن
٥٠	ظ	خـ	و	- -ِ -	ـف	- -ِ -	ص	- -َ -	ظ	ذ
٦٠	- -ِ -	- -ُ -	ـن	ل	ـف	ـم	- -َ -	ـعـ	ـن	- -َ -
٧٠	- -َ -	ي	آ	م	ـغ	د	- -َ -	ر	ظ	إ
٨٠	و	ـل	ـف	س	- -َ -	أ	ز	ـف	ق	ـه
٩٠	أ	ـهـ	ـي	حـ	- -ُ -	خـ	جـ	ز	- -ِ -	ع
١٠٠	ء	- -ِ -	آ	- -ُ -	ـك	- -َ -	ض	ـل	- -ِ -	- -َ -

2019 Letter Sound Identification (with Arabic Modifiers) – Student Stimulus

	ـلَـ	لِ		كُ		أمثلة:				
	١٠	٩	٨	٧	٦	٥	٤	٣	٢	١
١٠	ـبَ	أَ	طَ	ـهَ	ـتَ	بَ	قَ	آ	فَ	ـنَ
٢٠	أَ	حَـ	- -ِ -	ذَ	- -َ -	ـصَ	هَ	ـمَ	- -ُ -	ـسَ
٣٠	ـضَ	فَ	إِ	أَ	ؤ	ـلَ	عَ	ـثَ	كَ	ـنَ
٤٠	عَـ	نَ	ـأُ	آ	ـيَ	ـمَ	وَ	ـشَ	سَ	إِ
٥٠	وَ	خَـ	فَ	صَ	ظَ	ـجَ	ـصِ	ـتِ	ظَ	ذَ
٦٠	لَ	ـدَ	ـعَـ	ـهَ	ـمَ	ـقُ	ـنَ	فَ	ـنَ	ـبُ
٧٠	إِ	مَ	ظَ	دَ	ـظَ	آ	ـغَ	يَ	رَ	ـثَ
٨٠	ـفَ	زَ	سَ	أَ	ـهَ	قَ	وَ	وِ	ـفَ	ـلَ
٩٠	زَ	عَ	ـخَ	ـيَ	ـدُ	ـحَ	أَ	ـجَ	ـلِكـ	ـهَ
١٠٠	ـضَ	ـلَ	ـتَـ	ءَ	آ	ـطَ	ـكَ	ـشِ	ـحَ	ـمُ

Nonword Reading (Original) – Student Stimulus

بَسِان			شَلَامِيذُ	الفَلَّطُ	أمثلة:
	٥	٤	٣	٢	١
٥	مَعَجُ	سَدُعَرْ	أُفَالٍ	كَوْرَ	قِمَالِي
١٠	مَجِكْنَان	أُضِيخُ	كَنْتُو	أَسْرَأُ	سَنَّغَ
١٥	سَطِيفُلْ	الْقُسُنِ	أُشَرِّقُ	كَنْثُ	بِالْوَجِيبِ
٢٠	عَقَص	يَاشَتْنُو	نُقَاخُ	هَوْزَن	زَعَكَ
٢٥	هَصَمَهُ	عَدْرُ	الزَّفْلَقَة	أَصْنَاق	أَسُوهِشُّ
٣٠	شَعِيَ	نَمُقَ	خَفِجَ	يَلَبُ	لَغَادِيدُ
٣٥	النَّدِيقْ	زَيْدَبُ	مَرَبْ	غَجَتُهْم	أَضِيَ
٤٠	بَيْفُنان	سُلْمَى	لِجَادَا	عُبِيرًا	الكِرْفُ
٤٥	سُلْدَنِي	كِسَامٌ	مَجْثَمْ	عُبِيرَجٍ	قَشَمَ
٥٠	كَبَع	كَأَدَ	مَثَعُ	يَقْلِزُ	يَهْرِدُهُ

Nonword Reading (Revised) – Student Stimulus

بَخُط			صَلَعَ	طَنَلَ	أمثلة
	٥	٤	٣	٢	١
٥	نَمَشَ	لَسِقْ	عَوَمَ	مَفَدَ	تَّسَقَ
١٠	زَرَتَ	لَرَقَ	لَشَتَ	كُّصَحَ	ضَلِعُ
١٥	كُسِبْاَ	ضُحْوَ	شَمِجَ	رُيِف	حَطَمَ
٢٠	خَرَبَ	عُلِاثَ	كَوم	هُفَسَ	كُشْدَ
٢٥	طَجَصَ	عُرِاه	قَظَوَ	كَعِتَ	جَزِقَ
٣٠	دُجِال	لُفَنْمَ	عَمِفَ	رَالَفَ	طَزَرَ
٣٥	عُولَشَ	صُلَذَ	رُّجَكَ	فُّارعُ	صَرَشَ
٤٠	عُفِال	لَفَثَ	قَوق	رَحَنْسَ	دَقصَ
٤٥	كُّوذ	تَشَوف	لَار	صَلَفَ	اضيج
٥٠	لُّيب	جَشَرَ	فَنَثَ	طْسَيب	سُّحَمَ

Oral Reading Fluency and Reading Comprehension

Reading Passage Student Stimulus

سارَة بِنْتٌ مُوَذَّبَةٌ تَعِيشُ سَعِيدَةً مَعَ أَبَوَيْها، وَهِيَ تُحِبُّ الْفَاكِهَةَ وَالْبَرْتُقَالَ.

وَتُحَافِظُ عَلَى دُرُوسِها تَسَاعِدُ وَالِدَتَها فِي الْمَطْبَخِ.

وَفِي يَوْمٍ شاهَدَتْ سارَة بائِعاً مُتَجَوِّلاً يَبِيعُ الْمَأكُولاتِ عَلَى مَقْرَبَةٍ مِنَ الْمَدْرَسَةِ، فَفَكَّرَتْ أَنْ تَأْكُلَ مِنْهُ.
رَجَعَتْ إِلَى الْبَيْتِ وَحِينَ دَخَلَتْ إِلَى الْمَطْبَخِ صاحَتْ وَوَضَعَتْ يَدَها عَلَى بَطْنِها، أَسْرَعَتْ بِها أُمُّها إِلَى الْمَرْكَزِ الصِّحِّيِّ. وَأَخْبَرَتْ سارَة الطَّبِيبَ بِما أَكَلَتْ مِنَ الْبائِعِ الْمُتَجَوِّلِ.
أَجْرَى الطَّبِيبُ فُحُوصاتِهِ وَوَصَفَ لَها الدَّواءَ وَنَصَحَها بِعَدَمِ تَناوُلِ الْأَطْعِمَةِ الْمَكْشُوفَةِ؛ لِأَنَّها مُلَوَّثَةٌ وَغَيْرُ صِحِّيَّةٍ، قالَتْ سارَة: لَنْ أُكَرِّرَ هَذا الْفِعْلَ أَبَداً.

Answer	Question	Text
سارة	حول من تدور القصة؟	سارَة بِنْتٌ مُوَذَّبَةٌ تَعِيشُ سَعِيدَةً مَعَ أَبَوَيْها، وَهِيَ تُحِبُّ الْفَاكِهَةَ وَالْبَرْتُقَالَ. وَتُحَافِظُ عَلَى دُرُوسِها
في المطبخ	أين تحاول سارة مساعدة والدتها؟	تُساعِدُ والِدَتَها فِي الْمَطْبَخِ
بيبيع المأكولات	ماذا يعمل البائع المتجول على مقربة من المدرسة؟	وَفِي يَوْمٍ شاهَدَتْ سارَة بائِعاً مُتَجَوِّلاً يَبِيعُ الْمَأكُولاتِ عَلَى مَقْرَبَةٍ مِنَ الْمَدْرَسَةِ، فَفَكَّرَتْ أَنْ تَأْكُلَ مِنْهُ
بما أكلت من البائع المتجول	ماذا أخبرت سارة الطبيب عندما ذهبت إليه؟	رَجَعَتْ إِلَى الْبَيْتِ وَحِينَ دَخَلَتْ إِلَى الْمَطْبَخِ صاحَتْ وَوَضَعَتْ يَدَها عَلَى بَطْنِها، أَسْرَعَتْ بِها أُمُّها إِلَى الْمَرْكَزِ الصِّحِّيِّ. وَأَخْبَرَتْ سارَة الطَّبِيبَ بِما أَكَلَتْ مِنَ الْبائِعِ الْمُتَجَوِّلِ.
لأنها ملوثة وغير صحية	لماذا نهتنع عن تناول الطعام من البائع المتجولين؟	أَجْرَى الطَّبِيبُ فُحُوصاتِهِ وَوَصَفَ لَها الدَّواءَ وَنَصَحَها بِعَدَمِ تَناوُلِ الْأَطْعِمَةِ الْمَكْشُوفَةِ؛ لِأَنَّها مُلَوَّثَةٌ وَغَيْرُ صِحِّيَّةٍ، قالَتْ سارَة: لَنْ أُكَرِّرَ هَذا الْفِعْلَ أَبَداً.

Listening Comprehension

Listening Passage – No Stimulus

سأقوم الآن بقراءة قصة قصيرة لك و بصوت عالي لمرة واحدة فقط، ثم سأطرح بعض الأسئلة عن القصة. حاول أن تستمع بدقة و أن تجيب عن الأسئلة بأفضل ما يمكنك. يمكنك أن تستخدم أية لهجة عربية تريدها خلال الاجابة عن الاسئلة.

هل أنت جاهز؟ لنبدأ.

كُلَّ صَبَاحٍ ، تَحْلُبُ أُمّي بَقَرَتَها، وَتُحَضّرُ لَنَا الفَطُور. ذَاتَ يَوْمٍ، لَمْ تَجِدِ البَقَرَةَ فِي الإسْطَبْلِ. اسْتَيْقَظْتُ فَلَمْ أَرَ فَطُوراً عَلى الطّاوِلَةِ. كانَتْ أُمّي قَدْ خَرَجَتْ لِلبَحْثِ عَنْهَا عِنْدَ الجِيرانِ، وفِي الحَقْلِ، وَقُرْبَ النَّهْرِ. غَيْرَ أَنَّهَا لَمْ تَعْثُرْ عَلَيْهَا. فَبَكَتْ حُزْناً لأَنَّهَا لَمْ تَعْثُرْ عَلَيْهَا. فَبَكَتْ حُزْناً لأَنَّها كَانَتْ تُحِبُّها كَثِيراً. فَلَمَّا عَادَتْ سَمِعَتْ ضَجِيجاً فِي المَطْبَخِ. إنَّهَا البَقَرَةُ تَأْكُلُ جَزَراً مِنْ سَلّةِ الخُضَارِ.

Listening Comprehension Questions – No Stimulus

#	Question	Answer
1	اذا تفعل الأم كل صباح؟	تحلب الأم بقرتها وتحضر لنا الفطور
2	اذا حدث للأم في الاسطبل؟	لم تجد البقرة
3	لماذا لم يجد الطفل فطوراً على الطاولة؟	لأن الأم لم تحلب البقرة
4	أين بحثت الأم عن البقرة؟	بحثت عنها عند الجيران وفي الحقل وقرب النهر
5	لم بكت الأم؟	أنها تحب بقرتها , لم تجد البقرة
6	ما الذي أحدث الضجيج في المطبخ؟	البقرة

EGMA Tools

Number Identification – Student Stimulus

A

٣٠	١٢	٠	٩	٢
٤٨	٢٣	٣٩	٤٥	٢٢
٦٥	٨٧	٧٤	٣٣	٩١
٩٨٩	٧٣١	٥٨٠	٢٤٥	١٠٨

Number Discrimination – Student Stimulus

B1 (Examples)

٨	٤
٢٢	١٢

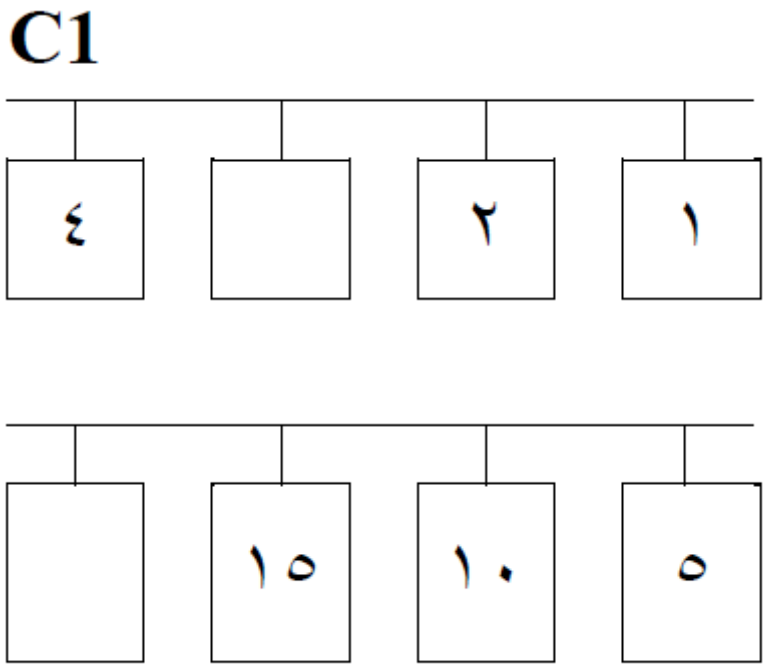
B2

٥	٧
٢٥	١٢
٢٩	٣٤
٤٨	٥٨
٦٧	٦٥

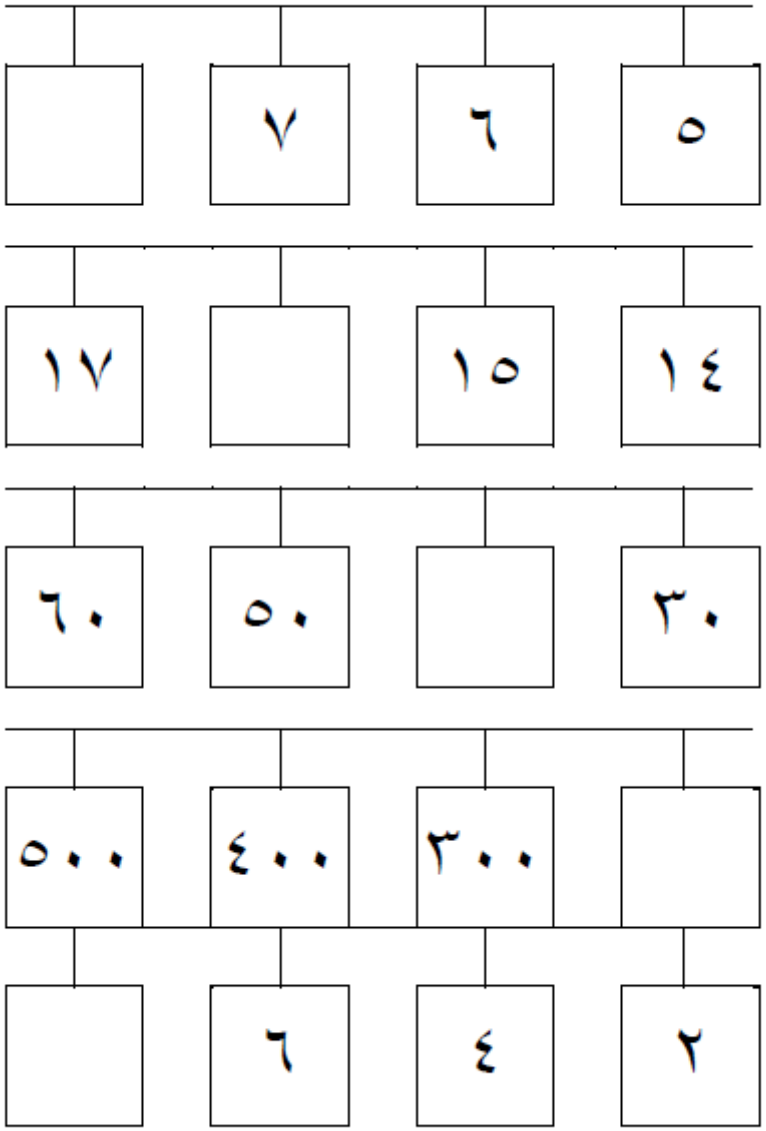
B3

٧٨	٩٤
١٥٣	١٤٦
٥٣٧	٢٨٧
٦٠٥	٦٥٠
٩٦٧	٩٦٥

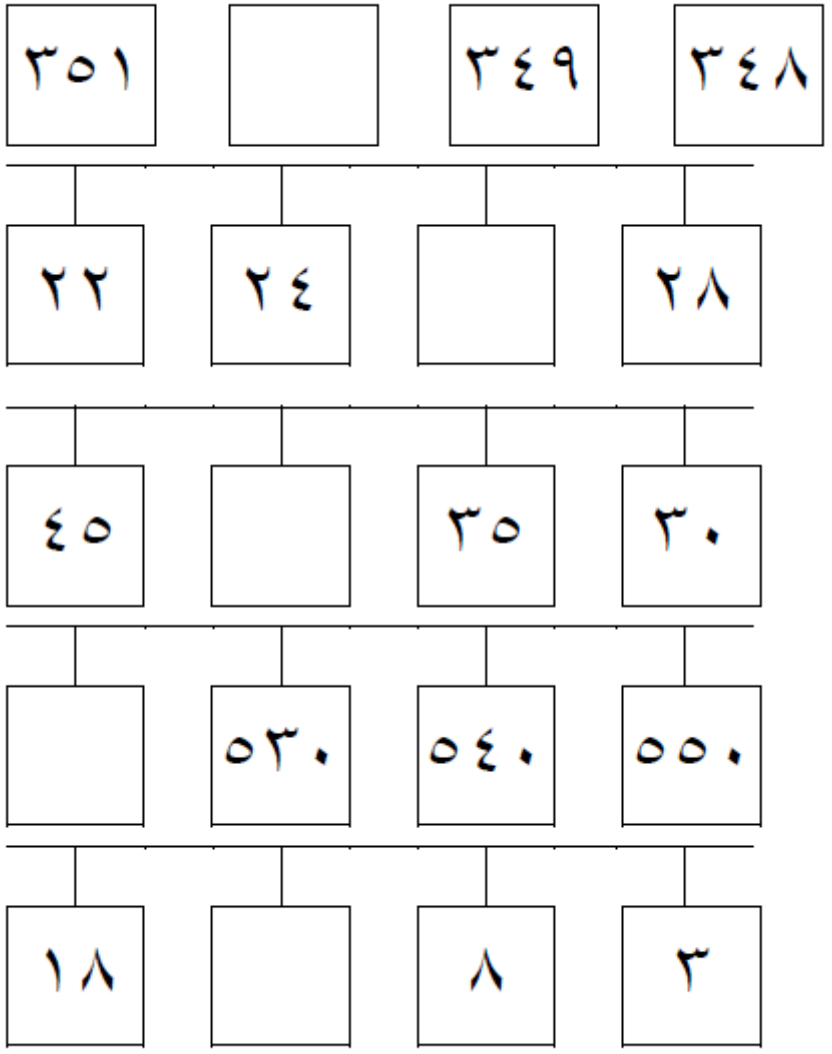
Missing Number – Student Stimulus



C2


٧
٦
٥
١٧
١٥
١٤
٦٠
٥٠
٣٠
٥٠٠
٤٠٠
٣٠٠
٦
٤
٢

C3


٣٥١
٣٤٩
٣٤٨
٢٢
٢٤
٢٨
٤٥
٣٥
٣٠
٥٣٠
٥٤٠
٥٥٠
١٨
٨
٣

Addition Level 1 – Student Stimulus

D1

□ = ٣ + ١
□ = ٢ + ٣
□ = ٢ + ٦
□ = ٥ + ٤
□ = ٣ + ٣
□ = ٨ + ٠
□ = ٣ + ٧
□ = ٧ + ٢
□ = ٥ + ٥
□ = ٨ + ٢

D2

□ = ٣ + ١١
□ = ٤ + ١٣
□ = ٣ + ١٦
□ = ٦ + ٨
□ = ٨ + ٧
□ = ٧ + ٩
□ = ٨ + ٨
□ = ١١ + ٢
□ = ٢ + ١٠
□ = ١٠ + ٨

Addition Level 2 – Student Stimulus

D3

□ = ٣ + ١٦
□ = ٧ + ١٨
□ = ١٢ + ٢٤
□ = ٣٧ + ٢٢
□ = ٢٦ + ٣٨

E1

□ = ١ – ٤
□ = ٢ – ٥
□ = ٢ – ٨
□ = ٥ – ٩
□ = ٣ – ٦
□ = ٠ – ٨
□ = ٣ – ١٠
□ = ٧ – ٩
□ = ٥ – ١٠
□ = ٨ – ١٠

E2

□ = ٣ _ ١٤
□ = ٤ _ ١٧
□ = ٣ _ ١٩
□ = ٦ _ ١٤
□ = ٧ _ ١٥
□ = ٩ _ ١٦
□ = ٨ _ ١٦
□ = ١١ _ ١٣
□ = ٢ _ ١٢
□ = ١٠ _ ١٨

Subtraction Level 2 – Student Stimulus

E3

□ = ٣ _ ١٩
□ = ٧ _ ٢٥
□ = ١٢ _ ٣٦
□ = ٣٧ _ ٥٩
□ = ٢٦ _ ٦٤

Word Problems – No Stimulus

#	Variable	Item	Answer
1	1	**باص فيه ستة أطفال** [توقف وتحقق من أن الطالب متابع معك] **اثنان منهم صبيان والباقي بنات** [توقف وتحقق من أن الطالب متابع معك] كم عدد البنات؟	٤
2	2	باص فيه عدد من الأطفال عند بدء الرحلة [توقف وتحقق من أن الطالب متابع معك] ثم ركب فيه طفلان آخران لاحقاً [توقف وتحقق من أن الطالب متابع معك] فأصبح مجموع عدد الأطفال في الباص ٩ أطفال [توقف وتحقق من أن الطالب متابع معك] كم عدد الأطفال الذين كانوا في الباص عند بدء الرحلة؟	٧
3	3	وزعت ١٢ قطعة حلوى على ٤ أطفال بالتساوي. [توقف وتحقق من أن الطالب متابع معك] كم عدد قطع الحلوى التي يحصل عليها كل طفل؟	٣

HEAD TEACHER SURVEY

Variable	Questions	Options
SCHOOL_STATUS	هل المدرسة تعقد دروسا مباشرة مع الطلاب في المدرسة ام دروسا عن بعد في هذا الأسبوع؟	0 "دروس شخصية فقط", 1 "كل من الدروس الشخصية والدروس عن بعد", 2 "دروس عن بعد فقط"
REMOTE_DAYS	كم عدد أيام الدراسة في الأسبوع والتي تعتمد التعليم عن بعد؟	1 "١", 2 "٢", 3 "٣", 4 "٤"
LIBRARY	هل يوجد غرفة مستقلة للمكتبة في مدرستك؟	1 "نعم", 0 "لا"
RESOURCE_ROOM	هل يوجد غرفة المصادر ذوي الاحتياجات الاضافية في مدرستك؟	1 "نعم", 0 "لا"
GRADE_ASSESSED	ما هي الصفوف التي يتم تقييمها اليوم ؟	3 "الصف الثالث", 4 "الصف الرابع", 7 "الصفين الثالث والرابع"
G3_SECTIONS	ما هو عدد شعب الصف الثالث؟	number
G4_SECTIONS	ما هو عدد شعب الصف الرابع؟	number
S2_1	ما هي الشعبة المختارة من الصف الثالث للقيام بالتقييم؟	number
S2_2	ما هو اسم المعلم؟	text
S2_3	ما هو عدد الأطفال الذكور الحاضرين في الصف الثالث المختار للقيام بالتقييم خلال تحديد الطلاب؟	number
S2_4	ما هو عدد الأطفال الإناث الحاضرات في الصف الثالث المختار للقيام بالتقييم خلال تحديد الطلاب؟	number
S2_5	ما هي الشعبة المختارة من الصف الرابع للقيام بالتقييم ؟	number
S2_6	ما هو اسم المعلم؟	text
S2_7	ما هو عدد الأطفال الذكور الحاضرين في الصف الرابع المختار للقيام بالتقييم خلال تحديد الطلاب؟	number
S2_8	ما هو عدد الأطفال الإناث الحاضرات في الصف الرابع المختار للقيام بالتقييم خلال تحديد الطلاب؟	number
S3_1	ما هو عدد الأطفال الذكور المسجلين في الصف الثالث؟	number
S3_2	ما هو عدد الأطفال الإناث المسجلات في الصف 3؟	number
S3_3	ما هو عدد الأطفال الذكور المسجلين في قسم (الشعبة) الصف الثالث الذي سيخضع للتقييم ؟	number
S3_4	ما هو عدد الأطفال الإناث المسجلات في قسم (الشعبة) الصف الثالث الذي سيخضع للتقييم؟	number
S3_5	كم عدد الأولاد المسجلين في الصف الرابع؟	number
S3_6	كم عدد الفتيات المسجلات في الصف الرابع؟	number
S3_7	كم عدد الطلاب الذكور المسجلين بشعبة الصف الرابع الذين تم أخذ عينات منهم في التقييم؟	number
S3_8	كم عدد الطالبات الاناث المسجلات بشعبة الصف الرابع الذين تم أخذ عينات منهم في التقييم؟	number

Annex E: Disaggregated Results

RESEARCH QUESTION 1: PROGRESSING AND PROFICIENT READERS

PROFICIENCY BANDS

Table E. 4. Proportion of Grade 3 Students Per Proficiency Band, by Assessment Year

	2020		2021	
	n	%	n	%
Non-reader	168	21.6%	196	22.7%
Beginning reader	365	45.4%	342	42.7%
Progressing reader	74	11.3%	64	9.1%
Proficient reader	123	21.7%	149	25.4%

Table E. 5. Proportion of Grade 3 Students Per Proficiency Band by Year across Gender

	2020				2021			
	Boys		Girls		Boys		Girls	
	n	%	n	%	n	%	n	%
Non-reader	78	23.3%	90	20.3%	78	23.3%*	90	20.3%
Beginning reader	176	48.9%	189	42.6%	176	48.9%	189	42.6%
Progressing reader	33	14.7%*	41	8.6%	33	14.7%	41	8.6%
Proficient reader	34	13.1%	89	28.5%*	34	13.1%	89	28.5%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05 level.								

Table E. 6. Proportion of Grade 3 Students Per Proficiency Band by Year across Province

	2020				2021			
	Province B		Province A		Province B		Province A	
	n	%	n	%	n	%	n	%
Non-reader	63	32.9%*	105	19.4%	53	25.8%	143	22.2%
Beginning reader	117	49.6%	248	44.6%	101	54.9%*	241	40.5%
Progressing reader	17	6.7%	57	12.2%	14	8.0%	50	9.3%
Proficient reader	22	10.8%	101	23.8%*	21	11.3%	128	28.0%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05 level.								

Table E. 7. Proportion of Grade 3 Students Per Proficiency Band by Province across Years

Proficiency Band	Province B				Province A			
	2020		2021		2020		2021	
	n	%	n	%	n	%	n	%
Non-reader	63	32.9%	53	25.8%	105	19.4%	143	22.2%
Beginning reader	117	49.6%	101	54.9%	248	44.6%	241	40.5%
Progressing reader	17	6.7%	14	8.0%	57	12.2%	50	9.3%
Proficient reader	22	10.8%	21	11.3%	101	23.8%	128	28.0%

RESEARCH QUESTION 2: STUDENT PROGRESSION FROM GRADE 3 TO GRADE 4

PROFICIENCY BANDS

Table E. 8. Proportion of 2020 G3 and 2021 G4 Students Per Proficiency Band by Year

	2020		2021	
	n	%	n	%
Non-reader	168	21.6%*	105	12.6%
Beginning reader	365	45.4%*	201	23.3%
Progressing reader	74	11.3%	92	12.1%
Proficient reader	123	21.7%	341	52.0%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column proportions.				

Table E. 9. Proportion of 2020 G3 and 2021 G4 Students Per Proficiency Band by Gender across Years

Proficiency Band	Boys				Girls			
	2020		2021		2020		2021	
	n	%	n	%	n	%	n	%
Non-reader	78	23.3%	70	17.4%	90	20.3%*	35	7.9%
Beginning reader	176	48.9%*	104	22.8%	189	42.6%*	97	23.8%
Progressing reader	33	14.7%	40	11.2%	41	8.6%	52	13.1%*
Proficient reader	34	13.1%	148	48.6%*	89	28.5%	193	55.2%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column proportions.								

Table E. 10. Proportion of 2020 G3 and 2021 G4 Students Per Proficiency Band by Province across Years

Proficiency Band	Province B				Province A			
	2020		2021		2020		2021	
	n	%	n	%	n	%	n	%
Non-reader	63	32.9%*	17	10.8%	105	19.4%*	88	13.0%
Beginning reader	117	49.6%*	49	22.4%	248	44.6%*	152	23.6%
Progressing reader	17	6.7%	30	19.9%*	57	12.2%	62	10.3%
Proficient reader	22	10.8%	88	46.9%*	101	23.8%	253	53.2%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column proportions.								

EGRA AND EGMA MEAN FLUENCY RATES AND ACCURACY SCORES

Table E. 11. EGRA Fluency Rates and Accuracy Scores, by Assessment Year

	2020		2021	
	n	Mean	n	Mean
Fluency Rates				
2019 Letter sound identification fluency (CLSPM)	730	34.4	739	43.2*
Nonword reading fluency (CNWPM)	730	4.0	739	7.0*
Oral reading fluency (CWPM)	730	17.8	739	31.3*
Accuracy Scores				
2019 Letter sound identification % Correct Out of 100 Total Items	730	62.5%	739	70.6%*
Nonword % Correct Out of 50 Total Items	730	15.6%	739	26.6%*
ORF % Correct Out of 82 Total Items	730	38.2%	739	60.9%*
Reading Comp % Correct Out of five Total Items	730	40.0%	739	62.2%*
Listening Comp % Correct Out of six Total Items	730	82.2%	739	90.4%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.				

Table E. 12. EGMA Fluency Rates and Accuracy Scores, by Assessment Year

	2020		2021	
	n	Mean	n	Mean
Fluency Rates				
Number recognition fluency (CNPM)	730	28.7	739	39.6*
Addition 1 fluency (CADDPM)	730	8.7	739	11.3*
Subtraction 1 fluency (CSUBPM)	730	5.6	739	7.4*
Accuracy Scores				
Number recognition: % correct out of 20 total items	730	85.1%	739	92.8%*
Number discrimination: % correct out of 10 total items	730	72.8%	739	81.9%*
Missing number: % correct out of 10 total items	730	49.0%	739	64.3%*
Addition 1: % correct out of 20 total items	730	73.9%	739	84.2%*
Addition 2: % correct out of five total items	730	48.4%	739	65.4%*
Subtraction 1: % correct out of 20 total items	730	53.6%	739	66.2%*
Subtraction 2: % correct out of five total items	730	27.1%	739	45.5%*
Word Problems: % correct out of three total items	730	52.0%	739	69.3%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.				

Table E. 13. EGRA Fluency Rates and Accuracy Scores by Gender across Years

Subtask	Boys				Girls			
	2020		2021		2020		2021	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
2019 Letter sound identification fluency (CLSPM)	321	31.6	362	39.4*	409	36.6	377	47.0*
Nonword reading fluency (CNWPM)	321	3.3	362	6.5*	409	4.5	377	7.5*
Oral reading fluency (CWPM)	321	14.6	362	28.5*	409	20.3	377	34.1*
Accuracy Scores								
2019 Letter sound identification % Correct Out of 100 Total Items	321	57.9%	362	68.1%*	409	66.1%	377	73.0%*
Nonword % Correct Out of 50 Total Items	321	12.9%	362	24.9%*	409	17.7%	377	28.2%*
ORF % Correct Out of 82 Total Items	321	32.9%	362	56.2%*	409	42.3%	377	65.3%*
Reading Comp % Correct Out of five Total Items	321	33.8%	362	57.6%*	409	44.9%	377	66.7%*
Listening Comp % Correct Out of six Total Items	321	78.6%	362	91.2%*	409	85.1%	377	89.6%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

Table E. 14. EGMA Fluency Rates and Accuracy Scores by Gender across Years

Subtask	Boys				Girls			
	2020		2021		2020		2021	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
Number recognition fluency (CNRPM)	321	29.9	362	40.7*	409	27.8	377	38.5*
Addition 1 fluency (CADDPM)	321	9.1	362	11.9*	409	8.3	377	10.8*
Subtraction 1 fluency (CSUBPM)	321	6.1	362	8.1*	409	5.3	377	6.7*
Accuracy Scores								
Number recognition: % correct out of 20 total items	321	86.4%	362	93.7%*	409	84.0%	377	91.9%*
Number discrimination: % correct out of 10 total items	321	76.1%	362	84.9%*	409	70.1%	377	79.1%*
Missing number: % correct out of 10 total items	321	49.5%	362	65.0%*	409	48.5%	377	63.7%*
Addition 1: % correct out of 20 total items	321	75.7%	362	85.5%*	409	72.5%	377	82.9%*
Addition 2: % correct out of five total items	321	50.1%	362	66.3%*	409	47.0%	377	64.6%*

Subtask	Boys				Girls			
	2020		2021		2020		2021	
	n	Mean	n	Mean	n	Mean	n	Mean
Subtraction 1: % correct out of 20 total items	321	56.9%	362	69.3%*	409	51.0%	377	63.2%*
Subtraction 2: % correct out of five total items	321	28.3%	362	46.4%*	409	26.1%	377	44.6%*
Word Problems: % correct out of three total items	321	50.7%	362	70.7%*	409	53.0%	377	68.0%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

Table E. 15. EGRA Fluency Rates and Accuracy Scores by Year across Province

	2020				2021			
	Province B		Province A		Province B		Province A	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
2019 Letter sound identification fluency (CLSPM)	219	30.7	511	35.1*	184	45.6	555	42.7
Nonword reading fluency (CNWPM)	219	3.2	511	4.1	184	9.4*	555	6.4
Oral reading fluency (CWPM)	219	9.9	511	19.3*	184	30.6	555	31.5
Accuracy Scores								
2019 Letter sound identification % Correct Out of 100 Total Items	219	57.1%	511	63.5%*	184	77.4%*	555	69.0%
Nonword % Correct Out of 50 Total Items	219	12.7%	511	16.1%	184	36.2%*	555	24.3%
ORF % Correct Out of 82 Total Items	219	22.7%	511	41.2%*	184	60.8%	555	60.9%
Reading Comp % Correct Out of five Total Items	219	26.0%	511	42.7%*	184	60.7%	555	62.6%
Listening Comp % Correct Out of six Total Items	219	80.1%	511	82.7%	184	89.9%	555	90.5%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

Table E. 16. EGMA Fluency Rates and Accuracy Scores, by Year across Province

	2020				2021			
	Province B		Province A		Province B		Province A	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
Number recognition fluency (CNRPM)	219	19.9	511	V	184	39.2	555	39.7
Addition 1 fluency (CADDPM)	219	7.1	511	9.0*	184	10.7	555	11.5
Subtraction 1 fluency (CSUBPM)	219	4.5	511	5.9*	184	7.2	555	7.5

	2020				2021			
	Province B		Province A		Province B		Province A	
	n	Mean	n	Mean	n	Mean	n	Mean
Accuracy Scores								
Number recognition: % correct out of 20 total items	219	79.4%	511	86.2%*	184	93.6%	555	92.6%
Number discrimination: % correct out of 10 total items	219	71.4%	511	73.1%	184	84.1%	555	81.4%
Missing number: % correct out of 10 total items	219	42.8%	511	50.2%*	184	69.7%*	555	63.1%
Addition 1: % correct out of 20 total items	219	65.3%	511	75.6%*	184	82.3%	555	84.6%
Addition 2: % correct out of five total items	219	42.4%	511	49.5%*	184	66.0%	555	65.3%
Subtraction 1: % correct out of 20 total items	219	43.6%	511	55.6%*	184	65.1%	555	66.4%
Subtraction 2: % correct out of five total items	219	23.1%	511	27.8%	184	48.1%	555	44.8%
Word Problems: % correct out of three total items	219	47.9%	511	52.8%	184	74.1%*	555	68.2%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

Table E. 17. EGRA Fluency Rates and Accuracy Scores, by Province across Years

Subtask	Province B				Province A			
	2020		2021		2020		2021	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
2019 Letter sound identification fluency (CLSPM)	219	30.7	184	45.6*	511	35.1	555	42.7*
Nonword reading fluency (CNWPM)	219	3.2	184	9.4*	511	4.1	555	6.4*
Oral reading fluency (CWPM)	219	9.9	184	30.6*	511	19.3	555	31.5*
Accuracy Scores								
2019 Letter sound identification % Correct Out of 100 Total Items	219	57.1%	184	77.4%*	511	63.5%	555	69.0%*
Nonword % Correct Out of 50 Total Items	219	12.7%	184	36.2%*	511	16.1%	555	24.3%*
ORF % Correct Out of 82 Total Items	219	22.7%	184	60.8%*	511	41.2%	555	60.9%*
Reading Comp % Correct Out of five Total Items	219	26.0%	184	60.7%*	511	42.7%	555	62.6%*
Listening Comp % Correct Out of six Total Items	219	80.1%	184	89.9%*	511	82.7%	555	90.5%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

Table E. 18. EGMA Fluency Rates and Accuracy Scores, by Province across Years

Subtask	Province B				Province A			
	2020		2021		2020		2021	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
Number recognition fluency (CNRPM)	219	19.9	184	39.2*	511	30.4	555	39.7*
Addition 1 fluency (CADDPM)	219	7.1	184	10.7*	511	9.0	555	11.5*
Subtraction 1 fluency (CSUBPM)	219	4.5	184	7.2*	511	5.9	555	7.5*
Accuracy Scores								
Number recognition: % correct out of 20 total items	219	79.4%	184	93.6%*	511	86.2%	555	92.6%*
Number discrimination: % correct out of 10 total items	219	71.4%	184	84.1%*	511	73.1%	555	81.4%*
Missing number: % correct out of 10 total items	219	42.8%	184	69.7%*	511	50.2%	555	63.1%*
Addition 1: % correct out of 20 total items	219	65.3%	184	82.3%*	511	75.6%	555	84.6%*
Addition 2: % correct out of five total items	219	42.4%	184	66.0%*	511	49.5%	555	65.3%*
Subtraction 1: % correct out of 20 total items	219	43.6%	184	65.1%*	511	55.6%	555	66.4%*
Subtraction 2: % correct out of five total items	219	23.1%	184	48.1%*	511	27.8%	555	44.8%*
Word Problems: % correct out of three total items	219	47.9%	184	74.1%*	511	52.8%	555	68.2%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

EGRA AND EGMA ZERO SCORES

Table E. 19. Proportion of EGRA and EGMA Zero Scores, by Assessment Year

Subtask		2020 data		2021 data	
		n	%	n	%
EGRA Zero scores					
2019 Letter sound zero score	1+ correct	705	96.8%*	691	93.9%
	Zero correct	25	3.2%	48	6.1%*
Non-word zero score	1+ correct	335	47.7%	424	62.2%*
	Zero correct	395	52.3%*	315	37.8%*
Oral reading fluency zero score	1+ correct	562	78.4%	634	87.4%*
	Zero correct	168	21.6%*	105	12.6%
Reading comprehension zero score	1+ correct	479	69.7%	607	84.7%*
	Zero correct	251	30.3%*	132	15.3%
Listening comprehension zero score	1+ correct	727	99.9%	734	99.7%
	Zero correct	3	0.1%	5	0.3%
EGMA Zero scores					

Number Recognition zero score	1+ correct	729	100.0%	738	100.0%
	Zero correct	1	0.0%	1	0.0%
Number Discrimination zero score	1+ correct	729	99.8%	736	99.9%
	Zero correct	1	0.2%	3	0.1%
Missing Number zero score	1+ correct	709	96.9%	733	99.6%*
	Zero correct	21	3.1%*	6	0.4%
Addition 1 zero score	1+ correct	709	96.6%	729	99.2%*
	Zero correct	21	3.4%*	10	0.8%
Addition 2 zero score	1+ correct	631	86.3%	716	97.1%*
	Zero correct	99	13.7%*	23	2.9%
Subtraction 1 zero score	1+ correct	665	91.4%	715	98.3%*
	Zero correct	65	8.6%*	24	1.7%
Subtraction 2 zero score	1+ correct	457	62.3%	592	79.5%*
	Zero correct	273	37.7%*	147	20.5%
Word Problem zero score	1+ correct	627	84.7%	701	95.5%*
	Zero correct	103	15.3%*	38	4.5%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.					

Table E. 20. Proportion of EGRA and EGMA Zero Scores, by Gender across Years

Subtask	Boys				Girls			
	2020		2021		2020		2021	
	n	%	n	%	n	%	n	%
EGRA								
2019 Letter Sound Identification	15	4.7%	29	7.6%	10	2.1%	19	4.8%
Nonword Reading	181	55.7%*	177	42.7%	214	49.7%*	138	33.1%
Oral Reading Fluency	78	23.3%	70	17.4%	90	20.3%*	35	7.9%
Reading Comprehension	124	34.3%*	89	22.0%	127	27.1%*	43	8.8%
Listening Comprehension	2	0.2%	2	0.2%	1	0.0%	3	0.5%
EGMA								
Number Recognition	1	0.1%	1	0.1%	0	0.0%	0	0.0%
Number Discrimination	0	0.0%	3	0.3%	1	0.3%	0	0.0%
Missing Number	15	6.0%*	3	0.3%	6	0.8%	3	0.6%
Addition 1	9	2.9%*	4	0.4%	12	3.8%*	6	1.1%
Addition 2	36	11.2%*	10	2.4%	63	15.7%*	13	3.5%
Subtraction 1	23	6.8%*	12	1.5%	42	10.0%*	12	1.9%
Subtraction 2	97	30.5%*	63	18.7%	176	43.4%*	84	22.2%
Word Problems	49	18.0%*	17	4.2%	54	13.1%*	21	4.8%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

Table E. 21. Proportion of EGRA and EGMA Zero Scores, by Year across Province

	2020				2021			
	Province B		Province A		Province B		Province A	
	n	%	n	%	n	%	n	%
EGRA								
2019 Letter Sound Identification	5	2.6%	20	3.3%	4	1.6%	44	7.2%*
Nonword Reading	114	55.0%	281	51.8%	43	22.4%	272	41.5%*
Oral Reading Fluency	63	32.9%*	105	19.4%	17	10.8%	88	13.0%
Reading Comprehension	100	46.5%*	151	27.1%	25	13.7%	107	15.6%
Listening Comprehension	0	0.0%	3	0.1%	0	0.0%	5	0.4%
EGMA								
Number Recognition	0	0.0%	1	0.0%	0	0.0%	1	0.0%
Number Discrimination	0	0.0%	1	0.2%	0	0.0%	3	0.2%
Missing Number	5	1.7%	16	3.4%	0	0.0%	6	0.5%
Addition 1	10	5.1%	11	3.1%	2	1.6%	8	0.6%
Addition 2	29	17.0%	70	13.1%	2	1.6%	21	3.2%
Subtraction 1	26	14.8%*	39	7.4%	5	1.3%	19	1.8%
Subtraction 2	85	45.4%	188	36.2%	31	14.5%	116	21.9%*
Word Problems	34	18.3%	69	14.7%	6	3.1%	32	4.8%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

Table E. 22. Proportion of EGRA and EGMA Zero Scores, by Province across Years

Subtask	Province B				Province A			
	2020		2021		2020		2021	
	n	%	n	%	n	%	n	%
EGRA								
20219 Letter Sound Identification	5	2.6%	4	1.6%	20	3.3%	44	7.2%*
Nonword Reading	114	55.0%*	43	22.4%	281	51.8%*	272	41.5%
Oral Reading Fluency	63	32.9%*	17	10.8%	105	19.4%*	88	13.0%
Reading Comprehension	100	46.5%*	25	13.7%	151	27.1%*	107	15.6%
Listening Comprehension	0	0.0%	0	0.0%	3	0.1%	5	0.4%
EGMA								
Number Recognition	0	0.0%	0	0.0%	1	0.0%	1	0.0%
Number Discrimination	0	0.0%	0	0.0%	1	0.2%	3	0.2%
Missing Number	5	1.7%	0	0.0%	16	3.4%*	6	0.5%
Addition 1	10	5.1%	2	1.6%	11	3.1%*	8	0.6%
Addition 2	29	17.0%*	2	1.6%	70	13.1%*	21	3.2%
Subtraction 1	26	14.8%*	5	1.3%	39	7.4%*	19	1.8%
Subtraction 2	85	45.4%*	31	14.5%	188	36.2%*	116	21.9%
Word Problems	34	18.3%*	6	3.1%	69	14.7%*	32	4.8%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

RESEARCH QUESTION 3: COMPARISON BETWEEN GRADE 3 AND GRADE 4 STUDENTS

PROFICIENCY BANDS

Table E. 23. Proportion of Students Per Proficiency Band, by Grade

	Grade 3		Grade 4	
	n	%	n	%
Non-reader	196	22.7%*	105	12.6%
Beginning reader	342	42.7%*	201	23.3%
Progressing reader	64	9.1%	92	12.1%
Proficient reader	149	25.4%	341	52.0%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column proportions.				

Table E. 24. Proportion of Students Per Proficiency Band by Gender across Grade

Proficiency Band	Boys				Girls			
	Grade 3		Grade 4		Grade 3		Grade 4	
	n	%	n	%	n	%	n	%
Non-reader	110	27.6%*	70	17.4%	86	17.8%*	35	7.9%
Beginning reader	173	40.8%*	104	22.8%	169	44.7%*	97	23.8%
Progressing reader	28	8.2%	40	11.2%	36	9.9%	52	13.1%
Proficient reader	67	23.4%	148	48.6%*	82	27.5%	193	55.2%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column proportions.								

Table E. 25. Proportion of Students Per Proficiency Band by Grade across Province

	Grade 3				Grade 4			
	Province B		Province A		Province B		Province A	
	n	%	n	%	n	%	n	%
Non-reader	53	25.8%	143	22.2%	17	10.8%	88	13.0%
Beginning reader	101	54.9%*	241	40.5%	49	22.4%	152	23.6%
Progressing reader	14	8.0%	50	9.3%	30	19.9%*	62	10.3%
Proficient reader	21	11.3%	128	28.0%*	88	46.9%	253	53.2%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column proportions.								

Table E. 26. Proportion of Students Per Proficiency Band by Province across Grade

Proficiency Band	Province B				Province A			
	Grade 3		Grade 4		Grade 3		Grade 4	
	n	%	n	%	n	%	n	%
Non-reader	53	25.8%*	17	10.8%	143	22.2%*	88	13.0%
Beginning reader	101	54.9%*	49	22.4%	241	40.5%*	152	23.6%
Progressing reader	14	8.0%	30	19.9%*	50	9.3%	62	10.3%*
Proficient reader	21	11.3%	88	46.9%*	128	28.0%	253	53.2%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column proportions.								

EGRA AND EGMA MEAN FLUENCY RATES AND ACCURACY SCORES

Table E. 27. EGRA Mean Fluency Rates and Accuracy Scores, by Grade

	Grade 3		Grade 4	
	n	Mean	n	Mean
Fluency Rates				
2017 Letter Sound identification fluency (CLSPM)	751	31.3	739	38.2*
2019 Letter Sound identification fluency (CLSPM)	751	33.6	739	43.2*
Nonword reading fluency (CNWPM)	751	4.2	739	7.0*
Nonword (revised) reading fluency (CNWPM)	751	7.3	739	11.2*
Oral reading fluency (CWPM)	751	18.6	739	31.3*
Accuracy Scores				
2017 Letter Sounds % Correct Out of 100 Total Items	751	54.2%	739	63.9%*
2019 Letter Sounds % Correct Out of 100 Total Items	751	58.4%	739	70.6%*
Nonword % Correct Out of 50 Total Items	751	16.0%	739	26.6%*
Nonword (Revised) % Correct Out of 50 Total Items	751	26.7%	739	39.9%*
ORF % Correct Out of 82 Total Items	751	38.5%	739	60.9%*
Reading Comp % Correct Out of Five Total Items	751	39.2%	739	62.2%*
Listening Comp % Correct Out of Six Total Items	751	87.2%	739	90.4%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.				

Table E. 28. EGMA Fluency Rates and Accuracy Scores, by Grade

	Grade 3		Grade 4	
	n	Mean	n	Mean
Fluency Rates				
Number recognition fluency (CNRPM)	751	28.9	739	39.6*
Addition 1 fluency (CADDPM)	751	9.4	739	11.3*
Subtraction 1 fluency (CSUBPM)	751	6.0	739	7.4*
Accuracy Scores				
Number recognition: % correct out of 20 total items	751	86.2%	739	92.8%*
Number discrimination: % correct out of 10 total items	751	71.3%	739	81.9%*
Missing number: % correct out of 10 total items	751	49.8%	739	64.3%*
Addition 1: % correct out of 20 total items	751	75.7%	739	84.2%*
Addition 2: % correct out of five total items	751	53.4%	739	65.4%*
Subtraction 1: % correct out of 20 total items	751	55.4%	739	66.2%*
Subtraction 2: % correct out of five total items	751	27.5%	739	45.5%*
Word Problems: % correct out of three total items	751	55.8%	739	69.3%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.				

Table E. 29. EGRA Fluency Rates and Accuracy Scores, by Gender across Grade

Subtask	Boys				Girls			
	Grade 3		Grade 4		Grade 3		Grade 4	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
2017 Letter Sound identification fluency (CLSPM)	378	29.4	362	35.1*	373	33.2	377	41.2*
2019 Letter Sound identification fluency (CLSPM)	378	31.5	362	39.4*	373	35.8	377	47.0*
Nonword reading fluency (CNWPM)	378	3.7	362	6.5*	373	4.7	377	7.5*
Nonword (revised) reading fluency (CNWPM)	378	6.9	362	10.3*	373	7.7	377	12.1*
Oral reading fluency (CWPM)	378	17.7	362	28.5*	373	19.6	377	34.1*
Accuracy Scores								
2017 Letter Sounds % Correct Out of 100 Total Items	378	50.1%	362	58.8%*	373	58.3%	377	68.7%*
2019 Letter Sounds % Correct Out of 100 Total Items	378	54.8%	362	68.1%*	373	62.1%	377	73.0%*
Nonword % Correct Out of 50 Total Items	378	14.0%	362	24.9%*	373	18.1%	377	28.2%*
Nonword (Revised) % Correct Out of 50 Total Items	378	24.4%	362	36.6%*	373	29.0%	377	43.0%*
ORF % Correct Out of 82 Total Items	378	35.6%	362	56.2%*	373	41.4%	377	65.3%*
Reading Comp % Correct Out of Five Total Items	378	36.1%	362	57.6%*	373	42.4%	377	66.7%*
Listening Comp % Correct Out of Six Total Items	378	87.5%	362	91.2%*	373	86.9%	377	89.6%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

Table E. 30. EGMA Fluency Rates and Accuracy Scores, by Gender across Grade

Subtask	Boys				Girls			
	Grade 3		Grade 4		Grade 3		Grade 4	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
Number recognition fluency (CNRPM)	378	31.2	362	40.7*	373	26.5	377	38.5*
Addition 1 fluency (CADDPM)	378	10.2	362	11.9*	373	8.6	377	10.8*
Subtraction 1 fluency (CSUBPM)	378	6.8	362	8.1*	373	5.2	377	6.7*
Accuracy Scores								
Number recognition: % correct out of 20 total items	378	88.0%	362	93.7%*	373	84.3%	377	91.9%*
Number discrimination: % correct out of 10 total items	378	75.9%	362	84.9%*	373	66.6%	377	79.1%*

Subtask	Boys				Girls			
	Grade 3		Grade 4		Grade 3		Grade 4	
	n	Mean	n	Mean	n	Mean	n	Mean
Missing number: % correct out of 10 total items	378	51.7%	362	65.0%*	373	48.0%	377	63.7%*
Addition 1: % correct out of 20 total items	378	77.4%	362	85.5%*	373	74.1%	377	82.9%*
Addition 2: % correct out of five total items	378	56.0%	362	66.3%*	373	50.7%	377	64.6%*
Subtraction 1: % correct out of 20 total items	378	59.7%	362	69.3%*	373	51.1%	377	63.2%*
Subtraction 2: % correct out of five total items	378	33.5%	362	46.4%*	373	21.3%	377	44.6%*
Word Problems: % correct out of three total items	378	58.7%	362	70.7%*	373	52.8%	377	68.0%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

Table E. 31. EGRA Fluency Rates and Accuracy Scores, by Grade across Province

	Grade 3				Grade 4			
	Province B		Province A		Province B		Province A	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
2017 Letter Sound identification fluency (CLSPM)	189	31.4	562	31.3	184	41.8*	555	37.4
2019 Letter Sound identification fluency (CLSPM)	189	31.5	562	34.0	184	45.6	555	42.7
Nonword reading fluency (CNWPM)	189	3.9	562	4.2	184	9.4*	555	6.4
Nonword (revised) reading fluency (CNWPM)	189	7.9	562	7.2	184	14.2*	555	10.5
Oral reading fluency (CWPM)	189	13.1	562	19.6*	184	30.6	555	31.5
Accuracy Scores								
2017 Letter Sounds % Correct Out of 100 Total Items	189	53.8%	562	54.2%	184	71.2%*	555	62.1%
2019 Letter Sounds % Correct Out of 100 Total Items	189	51.3%	562	59.7%	184	77.4%*	555	69.0%
Nonword % Correct Out of 50 Total Items	189	15.0%	562	16.2%	184	36.2%*	555	24.3%
Nonword (Revised) % Correct Out of 50 Total Items	189	29.0%	562	26.2%	184	50.3%*	555	37.5%
ORF % Correct Out of 82 Total Items	189	28.1%	562	40.3%	184	60.8%	555	60.9%
Reading Comp % Correct Out of Five Total Items	189	29.3%	562	41.0%	184	60.7%	555	62.6%
Listening Comp % Correct Out of Six Total Items	189	82.8%	562	88.0%	184	89.9%	555	90.5%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

Table E. 32. EGMA Fluency Rates and Accuracy Scores, by Grade across Province

	Grade 3				Grade 4			
	Province B		Province A		Province B		Province A	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
Number recognition fluency (CNRPM)	189	25.3	562	29.5*	184	39.2	555	39.7
Addition 1 fluency (CADDPM)	189	8.3	562	9.6*	184	10.7	555	11.5
Subtraction 1 fluency (CSUBPM)	189	4.8	562	6.2*	184	7.2	555	7.5
Accuracy Scores								
Number recognition: % correct out of 20 total items	189	82.5%	562	86.8%	184	93.6%	555	92.6%
Number discrimination: % correct out of 10 total items	189	68.5%	562	71.8%	184	84.1%	555	81.4%
Missing number: % correct out of 10 total items	189	48.0%	562	50.2%	184	69.7%*	555	63.1%
Addition 1: % correct out of 20 total items	189	69.1%	562	76.9%*	184	82.3%	555	84.6%
Addition 2: % correct out of five total items	189	45.7%	562	54.7%*	184	66.0%	555	65.3%
Subtraction 1: % correct out of 20 total items	189	46.4%	562	57.1%*	184	65.1%	555	66.4%
Subtraction 2: % correct out of five total items	189	19.0%	562	29.0%*	184	48.1%	555	44.8%
Word Problems: % correct out of three total items	189	55.1%	562	55.9%	184	74.1%*	555	68.2%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

Table E. 33. EGRA Fluency Rates and Accuracy Scores, by Province across Grade

Subtask	Province B				Province A			
	Grade 3		Grade 4		Grade 3		Grade 4	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
2017 Letter Sound identification fluency (CLSPM)	189	31.4	184	41.8*	562	31.3	555	37.4*
2019 Letter Sound identification fluency (CLSPM)	189	31.5	184	45.6*	562	34.0	555	42.7*
Nonword reading fluency (CNWPM)	189	3.9	184	9.4*	562	4.2	555	6.4*
Nonword (revised) reading fluency (CNWPM)	189	7.9	184	14.2*	562	7.2	555	10.5*
Oral reading fluency (CWPM)	189	13.1	184	30.6*	562	19.6	555	31.5*
Accuracy Scores								
2017 Letter Sounds % Correct Out of 100 Total Items	189	53.8%	184	71.2%*	562	54.2%	555	62.1%*

Subtask	Province B				Province A			
	Grade 3		Grade 4		Grade 3		Grade 4	
	n	Mean	n	Mean	n	Mean	n	Mean
2019 Letter Sounds % Correct Out of 100 Total Items	189	51.3%	184	77.4%*	562	59.7%	555	69.0%*
Nonword % Correct Out of 50 Total Items	189	15.0%	184	36.2%*	562	16.2%	555	24.3%*
Nonword (Revised) % Correct Out of 50 Total Items	189	29.0%	184	50.3%*	562	26.2%	555	37.5%*
ORF % Correct Out of 82 Total Items	189	28.1%	184	60.8%*	562	40.3%	555	60.9%*
Reading Comp % Correct Out of Five Total Items	189	29.3%	184	60.7%*	562	41.0%	555	62.6%*
Listening Comp % Correct Out of Six Total Items	189	82.8%	184	89.9%*	562	88.0%	555	90.5%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

RESEARCH QUESTION 4: GENDER GAP BETWEEN GRADE 3 AND 4 STUDENTS IN 2021

EGRA AND EGMA MEAN FLUENCY RATES AND ACCURACY SCORES

Table E. 34. EGMA Fluency Rates and Accuracy Scores, by Province across Grade

Subtask	Province B				Province A			
	Grade 3		Grade 4		Grade 3		Grade 4	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
Number recognition fluency (CNRPM)	189	25.3	184	39.2*	562	29.5	555	39.7*
Addition 1 fluency (CADDPM)	189	8.3	184	10.7*	562	9.6	555	11.5*
Subtraction 1 fluency (CSUBPM)	189	4.8	184	7.2*	562	6.2	555	7.5*
Accuracy Scores								
Number recognition: % correct out of 20 total items	189	82.5%	184	93.6%*	562	86.8%	555	92.6%*
Number discrimination: % correct out of 10 total items	189	68.5%	184	84.1%*	562	71.8%	555	81.4%*
Missing number: % correct out of 10 total items	189	48.0%	184	69.7%*	562	50.2%	555	63.1%*
Addition 1: % correct out of 20 total items	189	69.1%	184	82.3%*	562	76.9%	555	84.6%*
Addition 2: % correct out of five total items	189	45.7%	184	66.0%*	562	54.7%	555	65.3%*

Subtask	Province B				Province A			
	Grade 3		Grade 4		Grade 3		Grade 4	
	n	Mean	n	Mean	n	Mean	n	Mean
Subtraction 1: % correct out of 20 total items	189	46.4%	184	65.1%*	562	57.1%	555	66.4%*
Subtraction 2: % correct out of five total items	189	19.0%	184	48.1%*	562	29.0%	555	44.8%*
Word Problems: % correct out of three total items	189	55.1%	184	74.1%*	562	55.9%	555	68.2%*
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

EGRA AND EGMA ZERO SCORES

Table E. 35. Proportion of EGRA and EGMA Zero Scores, by Grade

	Grade 3		Grade 4	
	n	%	n	%
EGRA				
2017 Letter Sound Identification	92	11.7%*	58	6.2%
2019 Letter Sound Identification	77	9.8%*	48	6.1%
Nonword Reading	394	51.1%*	315	37.8%
Nonword (revised)	260	32.8%*	195	25.0%
Oral Reading Fluency	196	22.8%*	105	12.6%
Reading Comprehension	278	33.8%*	132	15.3%
Listening Comprehension	7	0.6%	5	0.3%
EGMA				
Number Recognition	1	0.2%	1	0.0%
Number Discrimination	2	0.1%	3	0.1%
Missing Number	22	2.8%*	6	0.4%
Addition 1	25	2.8%*	10	0.8%
Addition 2	81	9.0%*	23	2.9%
Subtraction 1	71	7.2%*	24	1.7%
Subtraction 2	314	37.2%*	147	20.5%
Word Problems	89	11.9%*	38	4.5%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column proportions.				

Table E. 36. Proportion of EGRA and EGMA Zero Scores, by Gender across Grade

Subtask	Boys				Girls			
	Grade 3		Grade 4		Grade 3		Grade 4	
	n	%	n	%	n	%	n	%
EGRA								
2017 Letter Sound Identification	56	14.7%*	36	9.3%	36	8.7%*	22	3.3%
2019 Letter Sound Identification	47	12.4%*	29	7.6%	30	7.2%*	19	4.8%
Nonword Reading	211	57.6%*	177	42.7%	183	44.5%*	138	33.1%
Nonword Reading (revised)	139	35.4%*	104	27.3%	121	30.2%*	91	22.8%
Oral Reading Fluency	110	27.7%*	70	17.4%	86	17.8%*	35	7.9%
Reading Comprehension	155	38.8%*	89	22.0%	123	28.7%*	43	8.8%
Listening Comprehension	3	0.5%	2	0.2%	4	0.7%	3	0.5%
EGMA								
Number Recognition	0	0.0%	1	0.1%	1	0.4%	0	0.0%
Number Discrimination	1	0.1%	3	0.3%	1	0.1%	0	0.0%
Missing Number	13	4.0%*	3	0.3%	9	1.7%	3	0.6%
Addition 1	15	3.6%*	4	0.4%	10	1.9%	6	1.1%
Addition 2	34	8.2%*	10	2.4%	47	9.8%*	13	3.5%
Subtraction 1	37	8.9%*	12	1.5%	34	5.5%*	12	1.9%
Subtraction 2	125	28.6%*	63	18.7%	189	45.9%*	84	22.2%
Word Problems	47	11.7%*	17	4.2%	42	12.2%*	21	4.8%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column proportions.								

Table E. 37. Proportion of EGRA and EGMA Zero Scores, by Grade across Province

Subtask	Grade 3				Grade 4			
	Province B		Province A		Province B		Province A	
	n	%	n	%	n	%	n	%
EGRA								
2017 Letter Sound Identification	15	6.6%	77	12.7%	4	2.5%	54	7.1%*
2019 Letter Sound Identification	23	14.9%*	54	8.9%	4	1.6%	44	7.2%*
Nonword Reading	80	41.4%	314	52.8%*	43	22.4%	272	41.5%*
Nonword Reading (revised)	52	21.5%	208	34.8%*	21	13.2%	174	27.8%*
Oral Reading Fluency	53	25.9%	143	22.2%	17	10.8%	88	13.0%
Reading Comprehension	79	37.9%	199	33.1%	25	13.7%	107	15.6%
Listening Comprehension	2	1.4%	5	0.4%	0	0.0%	5	0.4%
EGMA								
Number Recognition	1	1.2%	0	0.0%	0	0.0%	1	0.0%
Number Discrimination	2	0.4%	0	0.0%	0	0.0%	3	0.2%
Missing Number	4	2.6%	18	2.9%	0	0.0%	6	0.5%
Addition 1	13	6.8%*	12	2.0%	2	1.6%	8	0.6%
Addition 2	29	13.1%	52	8.3%	2	1.6%	21	3.2%
Subtraction 1	34	17.3%*	37	5.4%	5	1.3%	19	1.8%
Subtraction 2	92	49.7%*	222	34.9%	31	14.5%	116	21.9%*
Word Problems	21	12.8%	68	11.8%	6	3.1%	32	4.8%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column proportions.								

Table E. 38. Proportion of EGRA and EGMA Zero Scores, by Province across Grade

Subtask	Province B				Province A			
	Grade 3		Grade 4		Grade 3		Grade 4	
	n	%	n	%	n	%	n	%
EGRA								
2017 Letter Sound Identification	15	6.6%	4	2.5%	77	12.7%*	54	7.1%
2019 Letter Sound Identification	23	14.9%*	4	1.6%	54	8.9%	44	7.2%
Nonword Reading	80	41.4%*	43	22.4%	314	52.8%*	272	41.5%
Nonword Reading (Revised)	52	21.5%	21	13.2%	208	34.8%*	174	27.8%
Oral Reading Fluency	53	25.9%*	17	10.8%	143	22.2%*	88	13.0%
Reading Comprehension	79	37.9%*	25	13.7%	199	33.1%*	107	15.6%
Listening Comprehension	2	1.4%	0	0.0%	5	0.4%	5	0.4%
EGMA								
Number Recognition	1	1.2%	0	0.0%	0	0.0%	1	0.0%
Number Discrimination	2	0.4%	0	0.0%	0	0.0%	3	0.2%
Missing Number	4	2.6%	0	0.0%	18	2.9%*	6	0.5%
Addition 1	13	6.8%*	2	1.6%	12	2.0%*	8	0.6%
Addition 2	29	13.1%*	2	1.6%	52	8.3%*	21	3.2%
Subtraction 1	34	17.3%*	5	1.3%	37	5.4%*	19	1.8%
Subtraction 2	92	49.7%*	31	14.5%	222	34.9%*	116	21.9%
Word Problems	21	12.8%*	6	3.1%	68	11.8%*	32	4.8%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column proportions.								

RESEARCH QUESTION 4: GENDER GAP BETWEEN GRADE 3 AND 4 STUDENTS IN 2021

EGRA AND EGMA MEAN FLUENCY RATES AND ACCURACY SCORES

Table E. 39. EGRA Fluency Rates and Accuracy Scores, by Gender across Grade

Subtask	Grade 3				Grade 4			
	Boys		Girls		Boys		Girls	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
2017 Letter Sound identification fluency (CLSPM)	378	29.4	373	33.2*	362	35.1	377	41.2*
2019 Letter Sound identification fluency (CLSPM)	378	31.5	373	35.8*	362	39.4	377	47.0*
Nonword reading fluency (CNWPM)	378	3.7	373	4.7*	362	6.5	377	7.5
Nonword (revised) reading fluency (CNWPM)	378	6.9	373	7.7	362	10.3	377	12.1*
Oral reading fluency (CWPM)	378	17.7	373	19.6	362	28.5	377	34.1*
Accuracy Scores								
2017 Letter Sounds % Correct Out of 100 Total Items	378	50.1%	373	58.3%*	362	58.8%	377	68.7%*

Subtask	Grade 3				Grade 4			
	Boys		Girls		Boys		Girls	
	n	Mean	n	Mean	n	Mean	n	Mean
2019 Letter Sounds % Correct Out of 100 Total Items	378	54.8%	373	62.1%*	362	68.1%	377	73.0%*
Nonword % Correct Out of 50 Total Items	378	14.0%	373	18.1%*	362	24.9%	377	28.2%
Nonword (Revised) % Correct Out of 50 Total Items	378	24.4%	373	29.0%*	362	36.6%	377	43.0%*
ORF % Correct Out of 82 Total Items	378	35.6%	373	41.4%*	362	56.2%	377	65.3%*
Reading Comp % Correct Out of Five Total Items	378	36.1%	373	42.4%*	362	57.6%	377	66.7%*
Listening Comp % Correct Out of Six Total Items	378	87.5%	373	86.9%	362	91.2%	377	89.6%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

Table E. 40. EGMA Fluency Rates and Accuracy Scores, by Gender across Grade

Subtask	Grade 3				Grade 4			
	Boys		Girls		Boys		Girls	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
Number recognition fluency (CNRPM)	378	31.2*	373	26.5	362	40.7	377	38.5
Addition 1 fluency (CADDPM)	378	10.2*	373	8.6	362	11.9*	377	10.8
Subtraction 1 fluency (CSUBPM)	378	6.8*	373	5.2	362	8.1*	377	6.7
Accuracy Scores								
Number recognition: % correct out of 20 total items	378	88.0%*	373	84.3%	362	93.7%	377	91.9%
Number discrimination: % correct out of 10 total items	378	75.9%*	373	66.6%	362	84.9%*	377	79.1%
Missing number: % correct out of 10 total items	378	51.7%*	373	48.0%	362	65.0%	377	63.7%
Addition 1: % correct out of 20 total items	378	77.4%	373	74.1%	362	85.5%*	377	82.9%
Addition 2: % correct out of five total items	378	56.0%*	373	50.7%	362	66.3%	377	64.6%
Subtraction 1: % correct out of 20 total items	378	59.7%*	373	51.1%	362	69.3%*	377	63.2%
Subtraction 2: % correct out of five total items	378	33.5%*	373	21.3%	362	46.4%	377	44.6%
Word Problems: % correct out of three total items	378	58.7%*	373	52.8%	362	70.7%	377	68.0%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

Table E. 41. EGRA Fluency Rates and Accuracy Scores, by Gender across Grade, Province B

Subtask	Grade 3				Grade 4			
	Boys		Girls		Boys		Girls	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
2017 Letter Sound identification fluency (CLSPM)	98	30.1	91	32.8	91	40.1	93	43.3
2019 Letter Sound identification fluency (CLSPM)	98	30.5	91	32.7	91	42.2	93	48.6*
Nonword reading fluency (CNWPM)	98	3.7	91	4.1	91	8.6	93	10.2
Nonword (revised) reading fluency (CNWPM)	98	8.1	91	7.7	91	12.3	93	16.0*
Oral reading fluency (CWPM)	98	14.3	91	11.8	91	26.4	93	34.3*
Accuracy Scores								
2017 Letter Sounds % Correct Out of 100 Total Items	98	49.7%	91	58.0%	91	68.3%	93	73.9%
2019 Letter Sounds % Correct Out of 100 Total Items	98	48.6%	91	54.0%	91	72.9%	93	81.5%*
Nonword % Correct Out of 50 Total Items	98	14.4%	91	15.5%	91	33.8%	93	38.3%
Nonword (Revised) % Correct Out of 50 Total Items	98	29.1%	91	28.9%	91	45.1%	93	54.9%
ORF % Correct Out of 82 Total Items	98	27.8%	91	28.4%	91	55.4%	93	65.6%
Reading Comp % Correct Out of Five Total Items	98	28.3%	91	30.3%	91	54.6%	93	66.2%*
Listening Comp % Correct Out of Six Total Items	98	83.9%	91	81.7%	91	90.3%	93	89.5%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

Table E. 42. EGMA Fluency Rates and Accuracy Scores, by Gender across Grade, Province B

Subtask	Grade 3				Grade 4			
	Boys		Girls		Boys		Girls	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
Number recognition fluency (CNRPM)	98	27.2	91	23.4	91	38.5	93	39.7
Addition 1 fluency (CADDPM)	98	9.0	91	7.7	91	11.5	93	10.1
Subtraction 1 fluency (CSUBPM)	98	5.4	91	4.3	91	7.7	93	6.7
Accuracy Scores								
Number recognition: % correct out of 20 total items	98	87.0%*	91	77.9%	91	94.2%	93	93.2%
Number discrimination: % correct out of 10 total items	98	75.4%*	91	61.6%	91	87.6%	93	81.0%
Missing number: % correct out of 10 total items	98	51.8%	91	44.2%	91	72.4%	93	67.2%
Addition 1: % correct out of 20 total items	98	74.0%	91	64.2%	91	85.2%	93	79.8%

Subtask	Grade 3				Grade 4			
	Boys		Girls		Boys		Girls	
	n	Mean	n	Mean	n	Mean	n	Mean
Addition 2: % correct out of five total items	98	50.9%*	91	40.5%	91	68.2%	93	64.0%
Subtraction 1: % correct out of 20 total items	98	51.7%*	91	41.0%	91	68.5%	93	62.2%
Subtraction 2: % correct out of five total items	98	25.5%*	91	12.4%	91	48.2%	93	48.0%
Word Problems: % correct out of three total items	98	57.4%	91	52.7%	91	71.4%	93	76.5%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

Table E. 43. EGRA Fluency Rates and Accuracy Scores, by Gender across Grade, Province A

Subtask	Grade 3				Grade 4			
	Boys		Girls		Boys		Girls	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
2017 Letter Sound identification fluency (CLSPM)	280	29.3	282	33.3*	271	34.0	284	40.7*
2019 Letter Sound identification fluency (CLSPM)	280	31.7	282	36.4*	271	38.7	284	46.6*
Nonword reading fluency (CNWPM)	280	3.7	282	4.8*	271	6.1	284	6.8
Nonword (revised) reading fluency (CNWPM)	280	6.7	282	7.8	271	9.9	284	11.1
Oral reading fluency (CWPM)	280	18.3	282	21.0	271	29.0	284	34.0*
Accuracy Scores								
2017 Letter Sounds % Correct Out of 100 Total Items	280	50.2%	282	58.4%*	271	56.7%	284	67.4%*
2019 Letter Sounds % Correct Out of 100 Total Items	280	55.9%	282	63.6%*	271	67.0%	284	70.9%
Nonword % Correct Out of 50 Total Items	280	13.9%	282	18.6%*	271	22.9%	284	25.7%
Nonword (Revised) % Correct Out of 50 Total Items	280	23.5%	282	29.0%*	271	34.7%	284	40.1%*
ORF % Correct Out of 82 Total Items	280	37.0%	282	43.7%*	271	56.4%	284	65.3%*
Reading Comp % Correct Out of Five Total Items	280	37.5%	282	44.6%*	271	58.2%	284	66.8%*
Listening Comp % Correct Out of Six Total Items	280	88.1%	282	87.8%	271	91.5%	284	89.7%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

Table E. 44. EGMA Fluency Rates and Accuracy Scores, by Gender across Grade, Province A

Subtask	Grade 3				Grade 4			
	Boys		Girls		Boys		Girls	
	n	Mean	n	Mean	n	Mean	n	Mean
Fluency Rates								
Number recognition fluency (CNRPM)	280	31.9*	282	27.1	271	41.2*	284	38.2
Addition 1 fluency (CADDPM)	280	10.4*	282	8.8	271	12.0*	284	11.0
Subtraction 1 fluency (CSUBPM)	280	7.0*	282	5.4	271	8.2*	284	6.7
Accuracy Scores								
Number recognition: % correct out of 20 total items	280	88.2%*	282	85.4%	271	93.5%	284	91.6%
Number discrimination: % correct out of 10 total items	280	75.9%*	282	67.5%	271	84.2%*	284	78.7%
Missing number: % correct out of 10 total items	280	51.6%	282	48.7%	271	63.3%	284	62.8%
Addition 1: % correct out of 20 total items	280	78.0%	282	75.9%	271	85.6%	284	83.7%
Addition 2: % correct out of five total items	280	56.9%	282	52.5%	271	65.8%	284	64.7%
Subtraction 1: % correct out of 20 total items	280	61.1%*	282	52.9%	271	69.5%*	284	63.4%
Subtraction 2: % correct out of five total items	280	35.0%*	282	22.9%	271	45.9%	284	43.8%
Word Problems: % correct out of three total items	280	58.9%*	282	52.8%	271	70.5%*	284	65.9%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column means.								

EGRA AND EGMA ZERO SCORES

Table E. 45. Proportion of EGRA and EGMA Zero Scores, by Gender across Grade

Subtask	Grade 3				Grade 4			
	Boys		Girls		Boys		Girls	
	n	%	n	%	n	%	n	%
EGRA								
2017 Letter Sound Identification	56	14.7%	36	8.7%*	36	9.3%	22	3.3%*
2019 Letter Sound Identification	47	12.4%	30	7.2%*	29	7.6%	19	4.8%
Nonword Reading	211	57.6%	183	44.5%*	177	42.7%	138	33.1%*
Nonword Reading (revised)	139	35.4%	121	30.2%	104	27.3%	91	22.8%
Oral Reading Fluency	110	27.7%	86	17.8%*	70	17.4%	35	7.9%*
Reading Comprehension	155	38.8%	123	28.7%*	89	22.0%	43	8.8%*
Listening Comprehension	3	0.5%	4	0.7%	2	0.2%	3	0.5%
EGMA								
Number Recognition	0	0.0%	1	0.4%	1	0.1%	0	0.0%

Subtask	Grade 3				Grade 4			
	Boys		Girls		Boys		Girls	
	n	%	n	%	n	%	n	%
Number Discrimination	1	0.1%	1	0.1%	3	0.3%	0	0.0%
Missing Number	13	4.0%	9	1.7%	3	0.3%	3	0.6%
Addition 1	15	3.6%	10	1.9%	4	0.4%	6	1.1%
Addition 2	34	8.2%	47	9.8%	10	2.4%	13	3.5%
Subtraction 1	37	8.9%	34	5.5%	12	1.5%	12	1.9%
Subtraction 2	125	28.6%	189	45.9%*	63	18.7%	84	22.2%
Word Problems	47	11.7%	42	12.2%	17	4.2%	21	4.8%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level. for column proportions.								

Table E. 46. Proportion of EGRA and EGMA Zero Scores, by Gender across Grade, Province B

Subtask	Grade 3				Grade 4			
	Boys		Girls		Boys		Girls	
	n	%	n	%	n	%	n	%
EGRA								
2017 Letter Sound Identification	8	5.1%	7	8.1%	2	2.1%	2	2.9%
2019 Letter Sound Identification	12	14.9%	11	14.8%	3	2.8%	1	0.4%
Nonword Reading	39	39.9%	41	42.9%	25	27.9%	18	17.5%
Nonword Reading (revised)	25	18.6%	27	24.4%	11	16.4%	10	10.3%
Oral Reading Fluency	29	26.8%	24	24.9%	11	14.8%	6	7.2%
Reading Comprehension	43	40.3%	36	35.4%	15	18.7%	10	9.2%
Listening Comprehension	0	0.0%	2	2.8%	0	0.0%	0	0.0%
EGMA								
Number Recognition	0	0.0%	1	2.4%	0	0.0%	0	0.0%
Number Discrimination	1	0.4%	1	0.4%	0	0.0%	0	0.0%
Missing Number	1	0.8%	3	4.4%	0	0.0%	0	0.0%
Addition 1	7	4.4%	6	9.2%	1	0.6%	1	2.5%
Addition 2	12	8.2%	17	18.1%	1	0.6%	1	2.5%
Subtraction 1	15	14.8%	19	19.8%	2	1.2%	3	1.3%
Subtraction 2	38	34.8%	54	64.8%*	10	9.4%	21	19.0%
Word Problems	10	9.5%	11	16.1%	3	4.3%	3	2.0%
Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level.for column proportions.								

Table E. 47. Proportion of EGRA and EGMA Zero Scores, by Gender across Grade, Province A

Subtask	Grade 3				Grade 4			
	Boys		Girls		Boys		Girls	
	n	%	n	%	n	%	n	%
EGRA								
2017 Letter Sound Identification	48	16.4%	29	8.8%*	34	10.9%	20	3.4%*
2019 Letter Sound Identification	35	12.0%	19	5.8%*	26	8.6%	18	5.8%
Nonword Reading	172	60.8%	142	44.8%*	152	46.1%	120	36.9%*
Nonword Reading (revised)	114	38.4%	94	31.2%	93	29.8%	81	25.9%
Oral Reading Fluency	81	27.8%	62	16.6%*	59	18.0%	29	8.1%*
Reading Comprehension	112	38.5%	87	27.6%*	74	22.8%	33	8.6%*
Listening Comprehension	3	0.6%	2	0.3%	2	0.2%	3	0.6%
EGMA								
Number Recognition	0	0.0%	0	0.0%	1	0.1%	0	0.0%
Number Discrimination	0	0.0%	0	0.0%	3	0.3%	0	0.0%
Missing Number	12	4.5%	6	1.2%*	3	0.3%	3	0.7%
Addition 1	8	3.5%	4	0.5%*	3	0.3%	5	0.8%
Addition 2	22	8.2%	30	8.3%	9	2.8%	12	3.7%
Subtraction 1	22	7.9%	15	3.0%*	10	1.6%	9	2.1%
Subtraction 2	87	27.4%	135	42.5%*	53	20.8%	63	22.9%
Word Problems	37	12.1%	31	11.5%	14	4.2%	18	5.4%

Note: An asterisk (*) indicates differences within the category that are statistically significant at the p<0.05-level. for column proportions.

Annex F: Correlation Between Subtasks

Correlation between Subtasks

Table F.1. Correlation Between Reading Subtasks

		Letter Sound fluency (without modifiers) (CLNPM)	Letter Sound fluency (with modifiers) (CLNPM)	Nonword reading fluency (CNWPM)	Nonword (revised) reading fluency (CNWPM)	Oral reading fluency (CWPM)	Letter Sound Id. (without modifiers) % Correct Out of Total Items	Letter Sound Id. (with modifiers) % Correct Out of Total Items	Nonword % Correct Out of Total Items	Nonword (revised) % Correct Out of Total Items	ORF % Correct Out of Total Items	Reading Comp % Correct Out of Total Items	Listening Comp % Correct Out of Total Items
Letter Sound fluency (without modifiers) (CLNPM)	Pearson Correlation	1	.595**	.496**	.543**	.504**	.910**	.702**	.495**	.537**	.562**	.522**	.212**
	Sig. (2-tailed)		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1490	1489	1489	1489	1489	1490	1490	1490	1490	1489	1490	1490
Letter Sound fluency (with modifiers) (CLNPM)	Pearson Correlation	.595**	1	.502**	.541**	.517**	.558**	.707**	.508**	.555**	.569**	.526**	.155**
	Sig. (2-tailed)	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1489	1489	1489	1489	1487	1489	1489	1489	1489	1487	1489	1489
Nonword reading fluency (CNWPM)	Pearson Correlation	.496**	.502**	1	.827**	.652**	.483**	.575**	.989**	.836**	.689**	.629**	.184**
	Sig. (2-tailed)	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1489	1489	1489	1489	1487	1489	1489	1489	1489	1487	1489	1489

		Letter Sound fluency (without modifiers) (CLNPM)	Letter Sound fluency (with modifiers) (CLNPM)	Nonword reading fluency (CNWPM)	Nonword (revised) reading fluency (CNWPM)	Oral reading fluency (CWPM)	Letter Sound Id. (without modifiers) % Correct Out of Total Items	Letter Sound Id. (with modifiers) % Correct Out of Total Items	Nonword % Correct Out of Total Items	Nonword (revised) % Correct Out of Total Items	ORF % Correct Out of Total Items	Reading Comp % Correct Out of Total Items	Listening Comp % Correct Out of Total Items
Nonword (revised) reading fluency (CNWPM)	Pearson Correlation	.543**	.541**	.827**	1	.699**	.505**	.621**	.826**	.961**	.734**	.678**	.169**
	Sig. (2-tailed)	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1489	1489	1489	1489	1487	1489	1489	1489	1489	1487	1489	1489
Oral reading fluency (CWPM)	Pearson Correlation	.504**	.517**	.652**	.699**	1	.431**	.597**	.641**	.680**	.918**	.858**	.236**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1489	1487	1487	1487	1489	1489	1489	1489	1489	1489	1489	1489
Letter Sound Identification (without modifiers) % Correct Out of Total Items	Pearson Correlation	.910**	.558**	.483**	.505**	.431**	1	.724**	.490**	.530**	.512**	.478**	.235**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000
	N	1490	1489	1489	1489	1489	1490	1490	1490	1490	1489	1490	1490
Letter Sound Identification (with modifiers) % Correct Out of Total Items	Pearson Correlation	.702**	.707**	.575**	.621**	.597**	.724**	1	.588**	.657**	.680**	.642**	.226**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000
	N	1490	1489	1489	1489	1489	1490	1490	1490	1490	1489	1490	1490

		Letter Sound fluency (without modifiers) (CLNPM)	Letter Sound fluency (with modifiers) (CLNPM)	Nonword reading fluency (CNWPM)	Nonword (revised) reading fluency (CNWPM)	Oral reading fluency (CWPM)	Letter Sound Id. (without modifiers) % Correct Out of Total Items	Letter Sound Id. (with modifiers) % Correct Out of Total Items	Nonword % Correct Out of Total Items	Nonword (revised) % Correct Out of Total Items	ORF % Correct Out of Total Items	Reading Comp % Correct Out of Total Items	Listening Comp % Correct Out of Total Items
Nonword % Correct Out of Total Items	Pearson Correlation	.495**	.508**	.989**	.826**	.641**	.490**	.588**	1	.849**	.694**	.631**	.187**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000
	N	1490	1489	1489	1489	1489	1490	1490	1490	1490	1489	1490	1490
Nonword (revised) % Correct Out of Total Items	Pearson Correlation	.537**	.555**	.836**	.961**	.680**	.530**	.657**	.849**	1	.751**	.693**	.196**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000
	N	1490	1489	1489	1489	1489	1490	1490	1490	1490	1489	1490	1490
ORF % Correct Out of Total Items	Pearson Correlation	.562**	.569**	.689**	.734**	.918**	.512**	.680**	.694**	.751**	1	.926**	.250**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000
	N	1489	1487	1487	1487	1489	1489	1489	1489	1489	1489	1489	1489
Reading Comp % Correct Out of Total Items	Pearson Correlation	.522**	.526**	.629**	.678**	.858**	.478**	.642**	.631**	.693**	.926**	1	.294**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000
	N	1490	1489	1489	1489	1489	1490	1490	1490	1490	1489	1490	1490

		Letter Sound fluency (without modifiers) (CLNPM)	Letter Sound fluency (with modifiers) (CLNPM)	Nonword reading fluency (CNWPM)	Nonword (revised) reading fluency (CNWPM)	Oral reading fluency (CWPM)	Letter Sound Id. (without modifiers) % Correct Out of Total Items	Letter Sound Id. (with modifiers) % Correct Out of Total Items	Nonword % Correct Out of Total Items	Nonword (revised) % Correct Out of Total Items	ORF % Correct Out of Total Items	Reading Comp % Correct Out of Total Items	Listening Comp % Correct Out of Total Items
Listening Comp % Correct Out of Total Items	Pearson Correlation	.212**	.155**	.184**	.169**	.236**	.235**	.226**	.187**	.196**	.250**	.294**	1
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	N	1490	1489	1489	1489	1489	1490	1490	1490	1490	1489	1490	1490

Table F.2. Correlations between Mathematics Subtasks

		Number recognition fluency (CNRPM)	Addition 1 fluency (CADDPM)	Subtraction 1 fluency (CSUBPM)	Number recognition % correct out of 20 total items	Number discrimination % correct out of 10 total items	Missing number % correct out of 10 total items	Addition 1 % correct out of 20 total items	Addition 2 % correct out of five total items	Subtraction 1 % correct out of 20 total items	Subtraction 2 % correct out of five total items	Word Problems % correct out of three total items
Number recognition fluency (CNRPM)	Pearson Correlation	1	.666**	.578**	.687**	.637**	.582**	.558**	.514**	.587**	.504**	.357**
	Sig. (2-tailed)		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490
Addition 1 fluency (CADDPM)	Pearson Correlation	.666**	1	.753**	.504**	.491**	.508**	.767**	.574**	.687**	.550**	.391**
	Sig. (2-tailed)	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490
Subtraction 1 fluency (CSUBPM)	Pearson Correlation	.578**	.753**	1	.452**	.483**	.493**	.564**	.513**	.893**	.602**	.411**
	Sig. (2-tailed)	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490

		Number recognition fluency (CNRPM)	Addition 1 fluency (CADDPM)	Subtraction 1 fluency (CSUBPM)	Number recognition % correct out of 20 total items	Number discrimination % correct out of 10 total items	Missing number % correct out of 10 total items	Addition 1 % correct out of 20 total items	Addition 2 % correct out of five total items	Subtraction 1 % correct out of 20 total items	Subtraction 2 % correct out of five total items	Word Problems % correct out of three total items
Number recognition % correct out of 20 total items	Pearson Correlation	.687**	.504**	.452**	1	.760**	.596**	.621**	.529**	.550**	.432**	.333**
	Sig. (2-tailed)	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000
	N	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490
Number discrimination % correct out of 10 total items	Pearson Correlation	.637**	.491**	.483**	.760**	1	.632**	.545**	.542**	.563**	.501**	.354**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000	0.000
	N	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490
Missing number % correct out of 10 total items	Pearson Correlation	.582**	.508**	.493**	.596**	.632**	1	.560**	.552**	.586**	.546**	.426**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000
	N	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490

		Number recognition fluency (CNRPM)	Addition 1 fluency (CADDPM)	Subtraction 1 fluency (CSUBPM)	Number recognition % correct out of 20 total items	Number discrimination % correct out of 10 total items	Missing number % correct out of 10 total items	Addition 1 % correct out of 20 total items	Addition 2 % correct out of five total items	Subtraction 1 % correct out of 20 total items	Subtraction 2 % correct out of five total items	Word Problems % correct out of three total items
Addition 1 % correct out of 20 total items	Pearson Correlation	.558**	.767**	.564**	.621**	.545**	.560**	1	.659**	.658**	.517**	.417**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000
	N	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490
Addition 2 % correct out of five total items	Pearson Correlation	.514**	.574**	.513**	.529**	.542**	.552**	.659**	1	.572**	.619**	.384**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000
	N	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490
Subtraction 1 % correct out of 20 total items	Pearson Correlation	.587**	.687**	.893**	.550**	.563**	.586**	.658**	.572**	1	.639**	.464**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000
	N	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490	w1490

		Number recognition fluency (CNRPM)	Addition 1 fluency (CADDPM)	Subtraction 1 fluency (CSUBPM)	Number recognition % correct out of 20 total items	Number discrimination % correct out of 10 total items	Missing number % correct out of 10 total items	Addition 1 % correct out of 20 total items	Addition 2 % correct out of five total items	Subtraction 1 % correct out of 20 total items	Subtraction 2 % correct out of five total items	Word Problems % correct out of three total items
Subtraction 2 % correct out of five total items	Pearson Correlation	.504**	.550**	.602**	.432**	.501**	.546**	.517**	.619**	.639**	1	.447**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000
	N	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490
Word Problems % correct out of three total items	Pearson Correlation	.357**	.391**	.411**	.333**	.354**	.426**	.417**	.384**	.464**	.447**	1
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
	N	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490	1490

Annex G: Bibliography

Ibrahim A. Asadi, Asaid Khateb, and Michal Shany, How simple is reading in Arabic? A cross-sectional investigation of reading comprehension from first to sixth grade, Journal of Research in Reading, Volume 40, Issue S1, 2017, pp S1–S22, https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-9817.12093

'Education,' United Nations Children's Fund, www.unicef.org/syria/education

'Exposure to War Stressors Questionnaire,' Children and War Foundation, https://www.childrenandwar.org/projectsresources/measures

Gijsbert Stoet and David C. Geary, Sex Differences in Mathematics and Reading Achievement Are Inversely Related: Within- and Across-Nation Assessment of 10 Years of PISA Data. 2013. https://doi.org/10.1371/journal.pone.0057988

RTI International, Early Grade Mathematics Assessment (EGMA) Toolkit, 2014.

RTI International, Early Grade Reading Assessment (EGRA) Toolkit, Second Edition. Washington, DC: United States Agency for International Development, 2015.

USAID, Mapping the Ecosystem of Education Data for Internally Displaced Persons in the Middle East and Beyond: Issues, Challenges, and Recommendations. Washington, DC: United States Agency for International Development, 2021. Retrieved from: https://www.eccnetwork.net/sites/default/files/media/file/Mapping%20the%20Ecosystem%20of%20Education%20Data%20for%20IDPs.pdf

World Bank Group, Simulating the Potential Impacts of COVID-19 School Closures on Schooling and Learning Outcomes: A Set of Global Estimates, http://pubdocs.worldbank.org/en/798061592482682799/covid-and-education-June17-r6.pdf

Annex H: List of Consultees

No additional stakeholders were consulted beyond Chemonics International programme staff, School-to-School International, and the FCDO Senior Responsible Officer.

Disclaimer

This document has been redacted to protect the individuals involved in the Syria Education Programme. All names of people and locations have either been altered or removed, as has any information that may identify people or locations.

The Syria Education Programme is funded by UK aid from the UK government.

Published in January 2022 by the Syria Education Programme.

Please consider the environment before printing this report.