Syria Education Programme
2022 Manahel Learning Assessment Report

March 2023

Chemonics Group UK Limited

School-to-School International

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

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 Idlib and Aleppo to collect data for this assessment. We also thank the hundreds of students who took time from their school days to be assessed 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 Group UK Limited.

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	2019 Novel Coronavirus
CSUBPM	Correct Subtraction Problems per Minute
CWPM	Correct Words per Minute
ED	Education Directorate
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
FCDO	United Kingdom Government's Foreign, Commonwealth and Development Office
IRR	Inter-rater Reliability test
ORF	Oral Reading Fluency
SGO	Safeguarding Officer
SEP	Syria Education Programme
STS	School-to-School International
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 the Idleb and Aleppo provinces in the Northwest of Syria. Manahel is a five-year project (July 2018 to June 2023)[1] funded by the United Kingdom Government's Foreign, Commonwealth and Development Office (FCDO) and implemented by Chemonics Group UK Limited. Manahel provides access to safe, inclusive, and quality learning opportunities to children in Northwest Syria while 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 and to inform the education directorates about learning outcomes. The study's results also serve as a point of comparison to the 2021 Manahel midline assessment.[2] The study was conducted by Manahel partner, School-to-School International (STS).

Five research questions guided the study:

1. What proportion of grade 3 students in 2020, 2021, and in 2022 are classified as 'progressing' and 'proficient' readers, and how have those proportions changed over the period?
2. How has last year's grade 3 cohort progressed in reading and maths outcomes in grade 4?
3. How do this year's grade 4 students compare to this year's grade 3 students?
4. What proportion of grade 3 students in 2022 are classified as 'non-readers' and what proportion cannot complete simple (addition) mathematical calculations (score 0 in reading and maths tasks)?
5. 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?

The study draws on data collected in December 2022 from 400 grade 3 students and 400 grade 4 students, with 53 percent being girls, attending 40 Manahel project schools in Idleb and Aleppo provinces. Data collection captured student and school data using four tools: an Early Grade Reading Assessment (EGRA), an Early Grade Mathematics Assessment (EGMA), a student survey, and a head teacher survey.

FINDINGS

RQ1: Proportions of students who are progressing and proficient readers and how have those proportions changed over the period?

The 2022 grade 3 EGRA results show a significant and consistent improvement on those of 2021 and 2020. In 2022, there is a decrease in the percentage of non-readers and beginning readers from 2020, however over a fifth of boys in grade 3 remain unable to read a single word of a grade level text. Overall, there has been significant improvement in the proportion of grade 3 students who can read with proficiency in 2022 compared to students in grade 3 in 2021 and 2020. Well over a third of grade 3 students (37.4 percent) read with understanding at the correct grade level in 2022, compared to a quarter in 2021. When those who are categorised as progressing readers are added, nearly half of grade 3 students attain this benchmark level for success in later schooling. This is up from a third of grade 3 students in 2020. Girls continue to outperform boys, by a significant margin, with 43.2 percent of girls reading fluently compared to 30.4 percent of boys.

---

[1] FCDO Syria Education Programme 1 ended in June 2022. From 1 July 2022 to 30 June 2023, Chemonics delivered the FCDO Bridging period which is technically year 1 of SEP II

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

Meanwhile 22.1 percent of grade 3 boys are unable to read a single word, hardly improved from the 22.8 percent who could not read a word in 2020. The proportion of grade 3 girls who cannot read a word is 11.3 percent down from 20.2 percent in 2020. Idleb province has fewer non-readers than Aleppo, but the latter has almost closed the gap at the proficient reader level.

Related Recommendation:

Reader profiles show a significant improvement from 2020 to 2022, indicating that teachers, supported by lead teachers and learning circles, are effectively implementing improved literacy classroom teaching practices, peer support, and continuation of the increased numbers of literacy sessions. Reading challenge campaigns are also raising the profile and enjoyment of reading among motivated students. The adoption of teacher-led early grade reading assessment (TEGRA) alongside the use of continuous assessment in grade 1 onwards, should help teachers identify struggling learners earlier, particularly boys, and guide them towards a targeted teaching and remediation approach focused on specific skills that students struggle with. This needs to be supported by deepening training on collecting, analysing and interpreting learner reading data. The potential of summer school remediation activities, with teacher-led early grade reading assessment (TEGRA) repeated after these targeted interventions, could help reduce zero scores and the number of students who repeat a grade or drop out. The after-school literacy clubs provided by Manahel could also be helpful for students to practice their reading.

RQ2: How has last year's grade 3 cohort progressed in reading and maths outcomes in grade 4?

Overall, EGRA and EGMA results show that students in grade 4 in 2022 outperformed students in grade 3 in 2021 based on administration of the same tests to both grades. This indicates that students improve their learning with an additional year of schooling, as is expected. However, these results do not indicate if students in grade 4 are performing at the expected level. While nearly 70 percent of grade 4 students were proficient readers in 2022 a quarter were proficient readers when in grade 3 a year earlier. In 2021, 22.8 percent of grade 3 students were non-readers, compared to only 6.3 percent of grade 4 students in 2022. Similarly, 42.9 percent of grade 3 students were beginning readers in 2021, compared to 16.3 percent of grade 4 students in 2022. Both genders showed similarly significant improvements in both reading and mathematics between grade 3 in 2021 and grade 4 in 2022.

Related Recommendation: To sustain these gains through the end of the programme and beyond, the delivery team should continue to help teachers increase the amount of time spent on both basic reading skills and more complex skills like fluency and comprehension. If this RQ is to be repeated FCDO should commission research into the transition from grade 3 to grade 4. Some of the improvement in performance in both reading and mathematics of the cohort in grade 4 compared to the same cohort in grade 3 a year earlier, may be due to the non-transitioning of poor performing students to grade 4, but keeping them for an extra year in grade 3.

RQ3: How do this year's grade 4 students compare to this year's grade 3 students?

Overall, grade 4 students significantly outperformed grade 3 students in every EGRA subtask. While all differences in scores between grades were statistically significant, grade 4 students had notably higher accuracy scores in ORF and reading comprehension – key more advanced reading skills. Grade 4 students had an average accuracy score of 73.3 percent on oral reading fluency compared to 49.1 percent for grade 3 students and 76.1 percent for reading comprehension compared to 52.7 percent for grade 3 students. A significantly lower proportion of grade 4 students received zero scores in all subtasks with no grade 4 students scoring zero on the listening comprehension subtask.

Related Recommendation: In grade 3, teachers' focus should be on building the skills needed to attain comprehension in reading, which will be needed 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. If the same EGRA tests are going to be used in SEP II, the reading comprehension subtest should be discontinued for grade 4

students as it has a ceiling effect. Finally, as with the last RQ, research is needed into the transition from grade 3 to grade 4 in the project schools.

RQ4: Proportion of Grade 3 students in 2022 are classified as 'non-readers' and proportion cannot complete simple (addition) mathematical calculations (zero scores in reading and maths tasks)

A significantly higher proportion of boys received zero scores in both oral reading fluency and reading comprehension in grade 3. This indicates that generally girls are performing at a higher level than boys in both these complex reading subtasks and that a higher proportion of boys have failed to master essential decoding and reading skills.

In mathematics, overall weak students find subtraction harder than addition, so a higher proportion of students scored zero on single digit subtraction than scored zero for single digit addition. However, no significant differences are observed between boys and girls on these subtasks.

> **Related Recommendation:** The project team should focus on analysing factors responsible for lower performance of boys in the complex reading tasks and design targeted remedial strategies. Teachers in grades 1 and 2 should make sure that all learners have understood the basic mathematical functions (addition/subtraction) while 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 mathematics taught in grade 4.

RQ5: Gender Gap in performance for grade 3 and 4 students and changes in the gap between grade 3 and 4

In 2022, 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 reading subtasks in grades 3 and 4. For the most basic reading skills measured by letter sound identification and listening comprehension, a bigger gender gap was observed in grade 3. For the advanced and critical reading skills of oral reading fluency and reading comprehension, the bigger gender gap was observed in grade 4.

In mathematics, the gender gap was consistently greater in grade 3 than grade 4, with boys outperforming girls on all maths subtasks in grade 3. However, by grade 4 girls had largely closed the performance gap with girls performing slightly better than boys on addition 2, subtraction 2 and word sums, all more complex algorithms. Boys still scored better on the other EGMA tests.

> **Related Recommendation:** The key recommendation that has already been made is that boys in the early grade in Manahel schools need specific and targeted, data-informed, enjoyable and incentivised approaches to learning to decode words and so read, ultimately with fluency. While girls are performing much better in mathematics by grade 4, weaknesses in their performance in earlier grades still need to be addressed so the closing of the gender gap we see in grade 4 occurs in the early grades.

Additional Recommendations include:

SCHOOL RELATED RECOMMENDATIONS

1. Early grade reading scores using EGRA show a significant improvement from 2020 to 2022, even taking into account COVID 19 related school closures, indicating that teachers, supported by the lead teachers, are effectively implementing key elements of the Manahel programme. However, a stronger focus on the use of continuous assessment, along with the introduction of teacher-led early grade reading assessment (TEGRA), to identify and track struggling students and support in developing in-school strategies of remediation will be helpful in guiding teachers towards a targeted teaching approach focused on specific skills that students struggle with. The learnings from TEGRA and continuous assessment could help teachers in differentiating instruction according to student needs. This will support students most at risk of not gaining basic and higher-level reading skills. Such

interventions and remediation are particularly required to help the fifth of all grade 3 boys who have failed to learn to read even a word after three years in school.

2. Through the end of the programme and in SEP II, early grade teachers should work with boys to build reading fluency so they can learn to read with understanding and achieve reading proficiency. The focus in grade 3 should be on interesting and stimulating methods of 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.
3. Teachers in grades 1 and 2 should make sure that all students – particularly girls, who tend to struggle more than boys in the early years – have understood the basic mathematical functions (addition/subtraction) while teachers in grade 3 should focus on more complex mathematics skills to ensure that students master mathematics operations and real-world thinking and are therefore better prepared for the more complex mathematics taught in grade 4. This talks to the need for end of grade benchmarks to be set for numeracy, as well as literacy.

PROJECT RELATED RECOMMENDATIONS

1. Teacher-led Early Grade Reading Assessment (TEGRA) could be implemented in the early stages of SEP II. This approach to assessing early grade students, which combines teachers running group and one-on-one assessments with effective planning and remediation cycles, would help identify struggling and failing students while providing appropriate measures that will help ensure that no student reaches grade 4 unable to read or do basic mathematics.
2. SEP II should work with schools following assessments (both internal continuous assessments, TEGRA and EGRA/EGMA) to assist the teachers in using the test data to inform their remediation efforts. This will require SEP II to train teachers in how to analyse their students' assessment results to inform remediation and how to use the in-school and after-school lessons to maximum effect.
3. Results clearly show gains in reading between grades 3 and 4 for boys and girls. To sustain these gains through SEP II, the project could continue to help teachers increase the amount of time spent on reading with the help of online tools and targeted interventions for non-readers. Research questions which compare student performance in grade 3 with performance in grade 4 could be dropped as Manahel field officers report that weak students in grade 3 often repeat the grade. This inevitably skews the results. If the comparison is to remain it could focus in on individual students who have progressed.
4. Tracking and monitoring attendance data, could help provide insights to the lower performance observed for boys in literacy tasks. This should be supplemented with interventions to assist vulnerable students. 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.
5. As it seems students in Aleppo are now performing at improved levels the project needs to analyse what elements of the situation and intervention in Aleppo had an impact on learner performance and attempt to replicate these in the regions where students appear to be falling behind. This could be done by conducting after-action reviews with implementing staff or an in-depth 'bright spots' study with schools and districts that showed substantial improvement.
6. A number of research related recommendations are detailed in the final sections of this report.

SYSTEM RELATED RECOMMENDATION

1. All of the above school-based recommendations will be more successful if supported from within the system. In particular, the EDs, Education Assemblies and school leaders should assist schools in implementing TEGRA so that they can interpret and analyse their continuous and termly assessment results in the early grades to inform their teaching and to build remediation measures around areas of weakness. This will help ensure that gender-based differences in performance are better understood and tackled as early as possible in the early grades with adapted teaching methodologies demonstrating UDL principles supported by targeted remediation during summer camps and other activities and school readiness to reduce zero scores and repetition.

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 public services – including schools and education management through the Education Directorates (ED) and the Ministry of Education – in the northern Aleppo and northern Idleb provinces.

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.[3]

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 schooling continued to be punctuated by temporary closures due to spikes in COVID-19 infection rates for some time and more recently schools in some districts have been impacted by cholera outbreaks. Even more recently, but after the data was collected for this report, the schools were closed again due to the impact of the 7.7 magnitude earthquake which devastated South East Turkiye and Northwest Syria.

MANAHEL PROGRAMME BACKGROUND

This learning assessment takes place in the bridging period between Syria Education Programme I (SEP I) and SEP II, and in fact forms the first year of SEP II. SEP I and II are funded by the United Kingdom's Foreign, Commonwealth and Development Office (FCDO). SEP I and the bridging year is implemented by Chemonics International as the Manahel programme. It builds on the previous Idarah project, funded by the United Kingdom and the European Union between 2014 and 2018.

Manahel is helping to provide access to safe, inclusive, and quality learning opportunities for children in conflict-affected, opposition-held areas of Syria while strengthening educational actors to manage education effectively. Manahel focuses on pedagogy, curriculum and planning, inclusion, and child protection, as well as community and caregiver awareness in relation to girls' education in particular. 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, accommodation for children with mild or moderate disability, and improved literacy and numeracy instruction. On average, teachers deliver 12 literacy and four numeracy 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, in addition to assistance to parents and guardians on supporting their children's reading development[4]. Overall Manahel has impacted some 560,000 children over the course of the project in northern Syria's Idleb and Aleppo provinces.

---

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

[4] 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 2022 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, with insights on factors influencing student performance to ensure the programme meets the shifting needs of the schools and students it serves.

The study draws on data collected in December 2022 from grade 3 and grade 4 students. Data collection included four tools: an Early Grade Reading Assessment (EGRA), an Early Grade Mathematics Assessment (EGMA), a student survey, and a head teacher survey.[5]

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

Table 1. 2022 learning assessment research questions

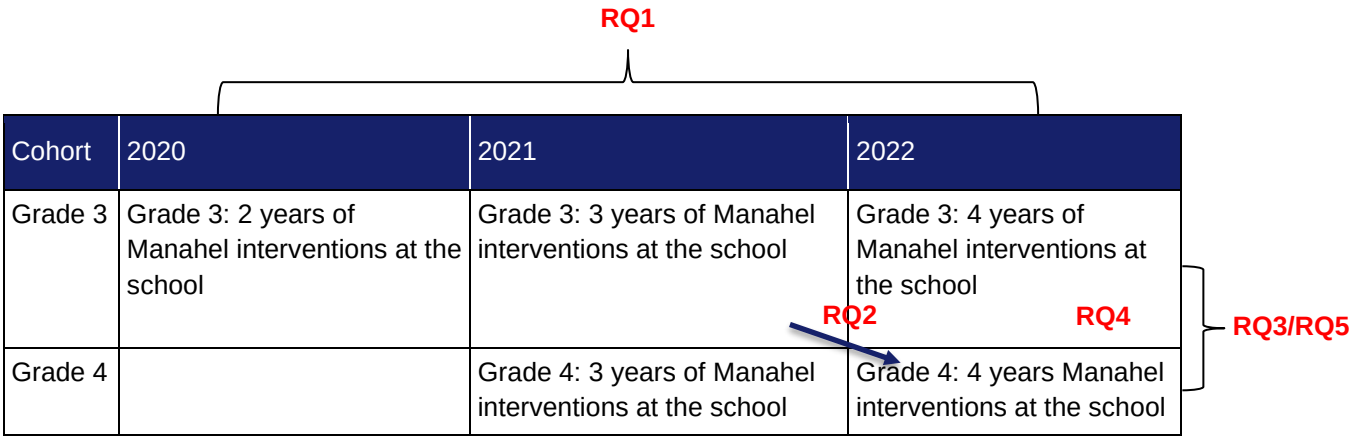
Research Question	Purpose
1. What proportion of grade 3 students in 2020, 2021 and in 2022 are classified as 'progressing' and 'proficient' readers, and how have those proportions changed over the period?	The purpose of RQ1 is to measure Manahel progress against the Impact Indicator and to see the percentage of students who can read in comparison to previous years to respond to the log frame indicator.
2. How has last year's grade 3 cohort progressed in reading and maths outcomes in grade 4?	The purpose of RQ2 is mainly to track growth within a cohort across an additional year of Manahel intervention (even though the last year has seen lower levels of intervention) and to measure students' progress in reading and maths.
3. How do this year's grade 4 students compare to this year's grade 3 students?	The purpose of RQ3 is to provide an alternate measure of growth with an additional year of Manahel intervention and instruction (even though the last year has seen lower levels of intervention) and to measure students' progress in reading and maths. It also provides Manahel with a picture of the extent to which the project is impacting on the schools and students it serves.
4. What proportion of grade 3 students in 2022 are classified as 'non-readers' and what proportion cannot complete simple (addition) mathematical calculations (score 0 in reading and maths tasks)?	The purpose of RQ4 is to get a measure of the proportion of learners who are failing to learn to read and understand the simplest of mathematical concepts. This will be an important metric for SEP II.
5. To what extent is there a gender gap in reading and maths 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 RQ5 is to understand, examine, and identify differences or gaps in performance based on gender.

It also indicates where the data for the various research questions is located. So, for instance, RQ2 is tracking 2021 grade 3 students into 2022 grade 4.

[5] Tools and their uses are included in **Error! Reference source not found.**.

Figure 1figure describes the comparisons and assumptions contained within these research questions. It also indicates where the data for the various research questions is located. So, for instance, RQ2 is tracking 2021 grade 3 students into 2022 grade 4.

Figure 1. Comparisons and assumptions in research questions



Cohort	2020	2021	2022
Grade 3	Grade 3: 2 years of Manahel interventions at the school	Grade 3: 3 years of Manahel interventions at the school	Grade 3: 4 years of Manahel interventions at the school
Grade 4		Grade 4: 3 years of Manahel interventions at the school	Grade 4: 4 years Manahel interventions at the school

The school's 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 late helps ensure that weak students are not asked to stay away on the day of the assessment, which would impact the validity of the results.

SAMPLING

Sampling was done using a two-stage approach:

First, STS randomly selected 30 schools in Idleb and 10 in Aleppo provinces, 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. This provision was not required in this data collection process but was available.

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 makes sure that the teachers do not pre-select the strongest students for the assessment. The head teacher at each school also completed a head teacher survey. Table 2 provides a summary of the target and final sample.

Table 2. Target and final sample

Province	Number of schools	Grades Assessed	Target sample		Final sample	
			Number of Girls	Number of Boys	Number of Girls	Number of Boys
Idlib	30	Grade 3	150	150	159	141
		Grade 4	150	150	166	134
Aleppo	10	Grade 3	50	50	50	50

		Grade 4	50	50	50	50
Total	40		400	400	425	375

The student sample is generalisable at the province level or by gender. Results at lower subgroup levels are associated with lower levels of confidence. While the targeted number of boys and girls to be assessed varies 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, 2021, and 2022 learning assessments to maintain consistency across studies. For 2022, the nonword reading subtasks A and B were removed from the EGRA, and the number discrimination and recognition subtasks were removed from the EGMA. These subtasks were found to have limited predictive capacity. Additionally, these removals allowed for a balance between the literacy and numeracy subtasks and a reduction in assessment time. This would allow for each student to be tested in one sitting, reducing time and administrative complexity. 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	2022 Manahel Study
Student Assessment Timepoint	END of grade 3	START of grade 3[6] 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		

[6] 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	2022 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[7]	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[8]. All the other subtasks were the same as in 2019 and 2020	2021 EGRA, nonword reading subtasks A and B were removed
EGMA Tools	EGMA tool adapted from tools from the 2012 MAHARAT project in Iraq	Same EGMA tool as Idarah			Same EGMA, number discrimination and recognition were removed
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	Included: • Student Survey • Head Teacher Survey

[7] There are two letter sounds identification subtasks and are differentiated between the years they were developed. Letter sounds identification without modifiers is sometimes referred to as letter sounds identification 2017, as this is the subtask from the 2017 Idarah study. Letter sounds identification with modifiers is sometimes referred to as letter sounds identification 2019, as this subtask was developed in 2019.

	2017 Idarah Study	2019 Manahel Study	2020 Manahel Study	2021 Manahel Study	2022 Manahel Study
Assessment Administration	Paper administration with stopwatches and timers.		Tablet administration using Tangerine®[9, 10]		

The final tools for the 2022 assessment include an updated EGRA, EGMA, a student survey, and head teacher survey. These tools are described in Table 4. A copy of all evaluation tools appears in **Error! Reference source not found.**.

Table 4. Description of the 2022 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.
Student Survey	This survey asked students about their home and education experience - including questions on family size, displacement, academic history, learning experience during COVID-19 school closures, attendance, and participation in Manahel activities.
Head Teacher Survey	The head teacher survey includes a brief survey on student enrolment and attendance. 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 28–29 November 2022 for three hours each day. The STS team in the United States trained two Manahel trainers on data collection procedures – both based in Northwest Syria and both previously learning assessment team leaders. STS facilitated the TOT in English with Arabic interpretation provided by the Manahel programme lead who attended the training to ensure the two trainees fully understood the training content, and to translate. Following the TOT, the two Manahel trainers conducted the enumerator training on 30 November - 2 December 2022 with 13 Manahel and downstream partner programme staff that would serve as enumerators. The Manahel programme lead provided virtual support. All the trainees had been enumerators on earlier EGRA/EGMA learning assessments. The training was conducted in Arabic with materials provided by STS. 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.

Enumerators visited 40 Manahel-supported schools from 3 - 20 December 2022. Enumerators were divided into three teams of four enumerators. 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 10.0 percent of the sampled

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

[10] 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*[10] and the *Early Grade Mathematics Assessment Toolkit*.

students, per standard EGRA practice.[11] 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 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, and they could opt out of the assessment at any time. Students were selected randomly on the day of the data collection and as already stated, the schools only knew the evening before that they were selected to make sure weak students were not excluded from the random sampling procedure or participating.

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, EGMA, and student survey 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 STATA version 16.1.

After applying the weighting functions, STS analysts produced descriptive statistics disaggregated by variables of interest. Descriptive results were analysed for statistically significant differences by gender and grade using chi-square tests and t-tests.[12] Weights and analysis for 2020 and 2021, previously computed in SPSS, were imported into Stata. All analyses were conducted using STATA version 16.1.[13]

CHALLENGES AND LIMITATIONS

The following limitations should be considered when reviewing the findings of the 2022 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 2020 and 2021 studies were included in the 2022 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 percent in

[11] Inter-rater reliability is the degree of agreement between 2 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.

[12] 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.

13 Some results from 2020 and 2021 may differ slightly from previous reporting (within a few decimal places) due to a different approach to weighting in 2022. The weighting approach used in 2022 was conducted using the svy function of Stata, which incorporates complex sample designs more robustly than probability weights in SPSS (the previous approach). Both approaches are means of mitigating sampling bias, and the practical significance of these differences is minimal.

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 2022 is likely to have been considerably lower. Indeed, the student survey in 2022 found that about 58.3 percent of students in 2022 had moved one or more times due to crises in the past academic year. This, along with teacher mobility, is still likely to impact greatly on student performance.

3. The 2022 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. The Manahel programme was implemented in a reduced sample of schools (n = 157) compared to 2021 (n = 435) due to budgetary constraints. The proportion of students sampled from Idleb and Aleppo remained the same in 2020, 2021, and 2022 (approximately 75.0 percent and 25.0 percent, respectively). All the 40 schools randomly sampled had been in the programme for years. One of the limitations of a smaller sample size, particularly when disaggregating data, is low statistical power. Therefore, data disaggregation is performed by gender and grade only. Lastly, comparisons of results between 2022 and prior time points must be made with caution.

FINDINGS

This section reports findings according to the study's five main research questions. Results statistically significant at the $p < 0.05$ level are referred to as 'significantly' lower or higher in the text. For research question 2, only descriptive statistics are reported.

DESCRIPTION OF THE SAMPLE

The sample was equally balanced between grade 3 and grade 4 students, each group representing approximately 50.0 percent of the overall sample (see Table 2). Girls made up 53.2 percent of the sample, while boys accounted for the remaining 46.8 percent. Students ranged in age from 7 to 14 years old. Most (82.7 percent) were on-age for their grade, but 0.5 percent were underage, and 16.8 percent were over-age. Less than 10.0 percent of the students reported missing school more than one day in the past week and about 58.3 percent of students had moved one or more times due to the various crises they had faced in the past academic year.

Mirroring the relative population distribution in the two provinces, most of the sample came from Idleb (75.0 percent) compared to Aleppo (25.0 percent). Within Idleb, students were relatively equally divided between the districts A, B, C, D (13.3 percent, 13.3 percent, 10.0 percent, and 6.7 percent of the overall sample, respectively). A slightly larger proportion came from the A and B districts(33.3 percent and 23.3 percent, respectively). About 60.0 percent of students from the Aleppo sample came from C district, while 40.0 percent came from D district

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 Idleb province, 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, 2021 and in 2022 are classified as 'progressing' and 'proficient' readers, and how have those proportions changed over the period?

As with the 2020 and 2021 Manahel learning assessments, the 2022 assessment classified students' EGRA scores into proficiency bands established by the 2017 Idarah assessment. All data is derived solely from the

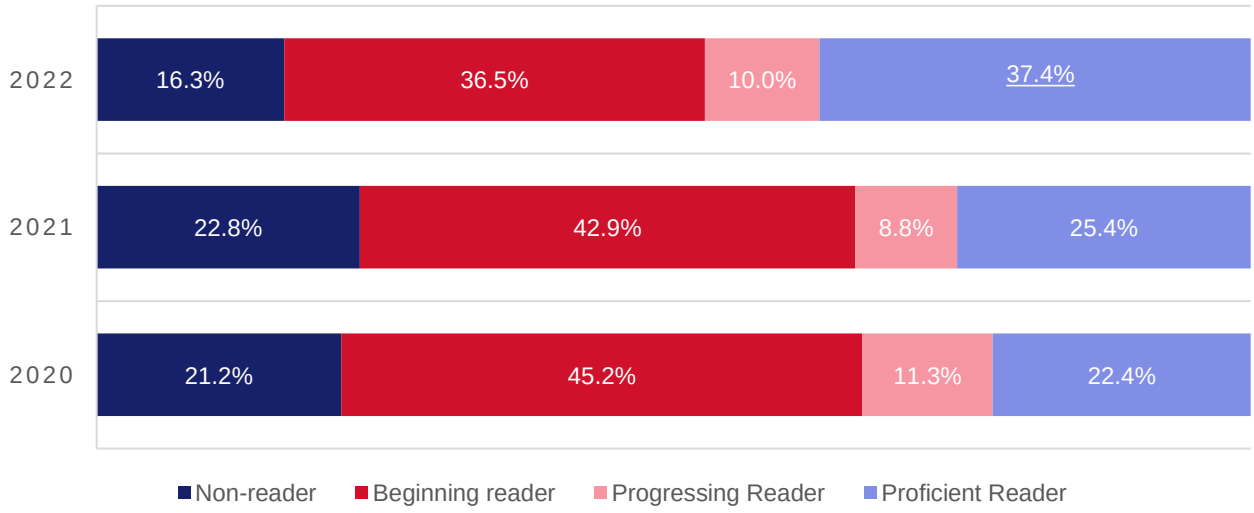
oral reading fluency (ORF) and reading comprehension subtasks. These bands tie the ability to read at a 'proficient' level—the highest band—directly to comprehension, while lower bands also consider fluency. The reading proficiency bands are defined as follows:

- **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 less than 80.0 percent of questions correctly on the reading comprehension subtask.
- **Progressing readers** are students who read 23 CWPM or more on ORF and answered less than 80.0 percent of the reading comprehension subtask correctly.
- **Proficient readers** are students who answered 80% or more of questions correctly on the reading comprehension subtask.

Note that 2022 sample numbers were much lower than previous years, so results comparing 2022 with previous years should be interpreted with caution. Both 2021 and 2022 grade 3 EGRA results show a significant and consistent improvement on those of 2020. The percentage of progressing and proficient readers has risen nearly 14 percentage points from 33.7 percent in 2020 to 47.4 percent in 2022. This means that from just one third of students having the basic reading skills to engage with grade 3 texts Manahel schools now have nearly half about to do so. In addition, in 2022 there is a moderate decrease in the percentage of non-readers from 2020 from 21.2 percent of students to 16.3 percent.

Overall, significant differences were found between the reading proficiency classification of students in grade 3 in 2022 and students in grade 3 in 2021 and 2020. In 2022, 37.4 percent of students met the reading proficiency benchmark of scoring 80.0 percent or higher on reading comprehension – significantly higher compared to 25.4 percent of students in 2021 (see Figure 2). In 2020, 22.4 percent of students met the reading proficiency benchmark of scoring 80.0 percent or higher on reading comprehension compared to 37.4 percent of students in 2022 (see Figure 2).

Figure 2: Proportion of 2022 grade 3, 2021 grade 3, and 2020 grade 3 students by reading proficiency level



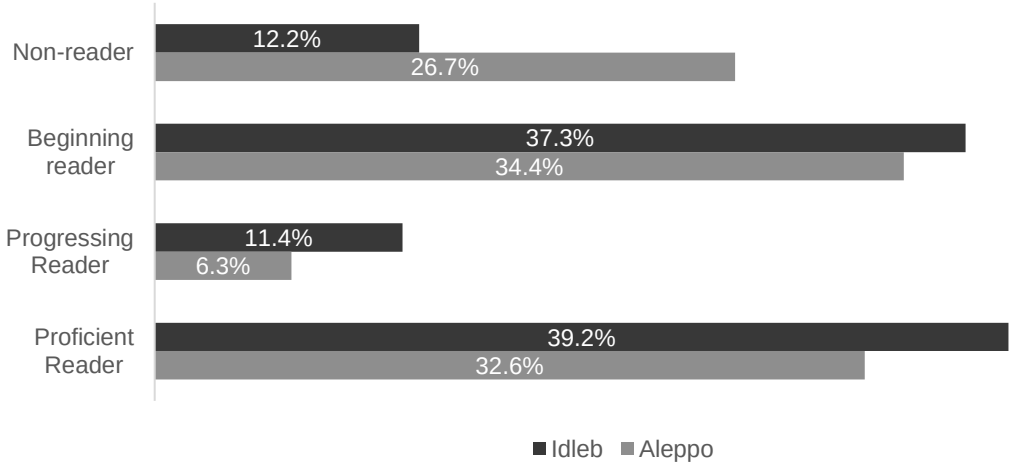
Note: Underlined scores are statistically significant between reading proficiency bands at p<0.05.

Results by province. No significant difference was found between the reading at proficiency level in Idleb and Aleppo in 2022 (see Figure 3)[14]. Aleppo had 32.6 percent of students as proficient readers and Idleb has 39.2 percent of students reaching this level. This is in contrast with the results in 2020 and 2021 where significantly more students were at the proficient level in Idleb as compared to Aleppo (24.6 percent of proficient readers in Idleb in 2020, compared to 11.0 percent in Aleppo; 28.0 percent of proficient readers in

[14] This is likely due to the small number of learners in each category in Aleppo, where 30 learners were non-readers, 33 were beginning readers, 7 were progressing readers, and 30 were proficient readers.

Idleb in 2021, compared to 11.3 percent in Aleppo). Note that 2022 sample numbers were lower than previous years, so results at this level of disaggregation should be interpreted with caution.

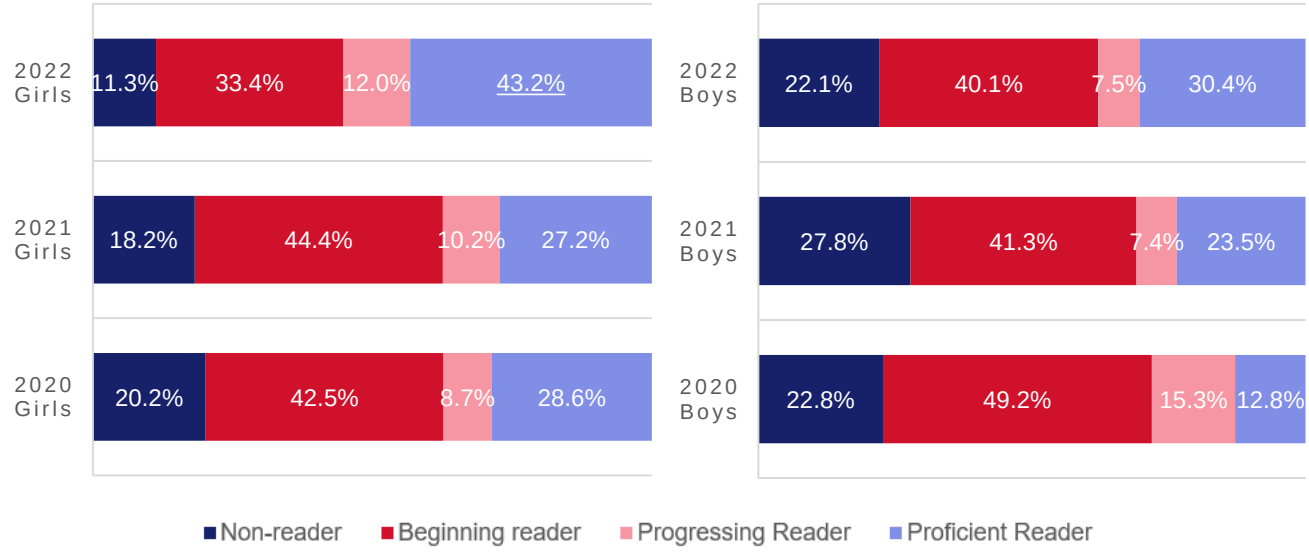
Figure 3: Comparison of grade 3 student reading proficiency levels in Idleb and Aleppo schools in 2022



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

Results by gender. A significantly larger proportion of girls and boys were proficient readers in 2022 as compared to 2020 and 2021. In 2022, 43.2 percent of girls were proficient readers, compared to 28.6 percent in 2020. The proportion of boys in the proficient reader category in 2020 was 12.8 percent and in 2022 30.4 percent. Boys continued to be over-represented in the 16.7 percent of 2022 students who could not read a word of the grade level text, with almost a quarter of boys being non-readers, not significantly better than in 2020. Over the same period the proportion of girls who were non-readers had fallen by nearly half to 11.3 percent.

Figure 4: Comparison of grade 3 student reading proficiency levels disaggregated by gender for 2022, 2021, and 2020.



Note: Underlined scores are statistically significant between reading proficiency bands at p<0.05.

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

Research Question 2: How has last year's grade 3 cohort progressed in reading and maths outcomes in grade 4?[15]

The purpose of RQ2 is mainly to track student progress in reading and mathematics within a cohort across an additional year of schooling. As the Manahel intervention focuses mainly on early grade students, this RQ provides some measure of the sustainability of improvements made to student reading and numeracy skills in the early grades and to what extent they get built on and developed in the following grade.

Overall, EGRA and EGMA results show that students in grade 4 in 2022 outperformed students in grade 3 in 2021 based on administration of the same grade 3 level tests to both grades. This indicates that students improve their learning with an additional year of schooling, as is expected. However, these results do not indicate if students in grade 4 are performing at the expected level. Not surprisingly, students in grade 4 are performing at ceiling[16] in the listening comprehension subtask.

Reading Outcomes

Overall, a much higher proportion of grade 4 students were proficient readers in 2022 than grade 3 students in 2021 (69.6 percent compared to 25.4 percent, respectively). Conversely, in 2021, 22.8 percent of grade 3 students were non-readers, compared to 6.3 percent of grade 4 students in 2022. Similarly, 42.9 percent of grade 3 students were beginning readers in 2021, compared to 16.3 percent of grade 4 students in 2022. This seems to indicate that many students who were beginning to read in 2021 in grade 3 had become fluent readers a year later in grade 4.

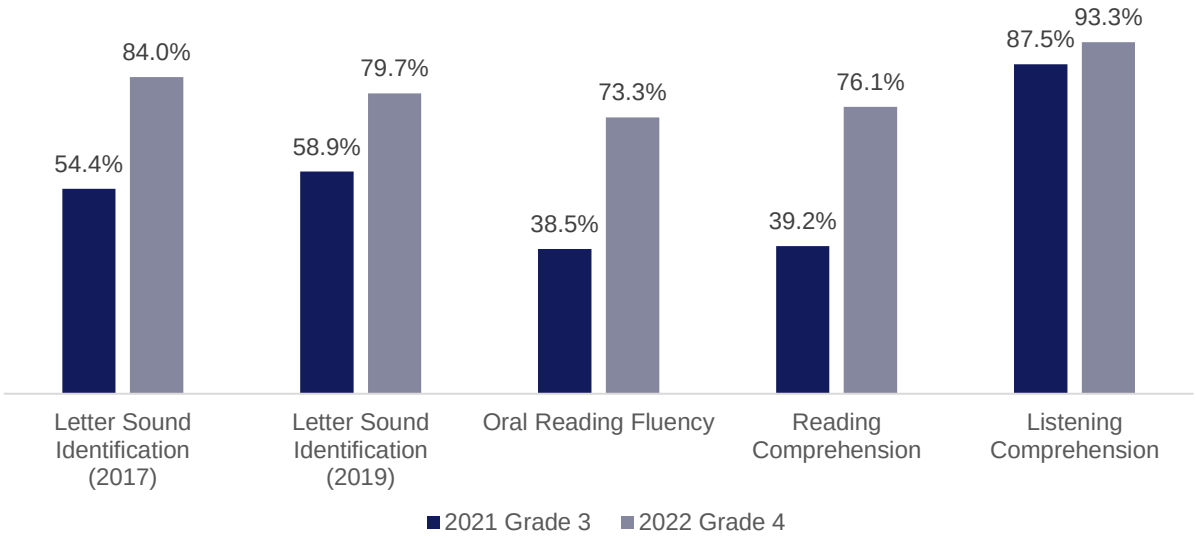
Among EGRA fluency scores, grade 4 students in 2022 had a letter sound fluency score of 53.8 correct letter sounds per minute (CLSPM); a letter sound fluency score—with Arabic modifiers—of 47.8 correct letter sounds per minute (CLSPM); and an ORF score of 40.1 CWPM. In 2021, grade 3 students read 31.3 CLSPM, 33.8 CLSPM, and 18.7 CWPM. These are very significant increases.

Among accuracy scores, grade 4 students in 2022 had higher scores than grade 3 students in 2021 on every EGRA subtask (see Figure 4). In 2022, grade 4 students had an average of 73.3 percent of ORF words correct and 76.1 percent of reading comprehension questions correct while grade 3 students in 2021 averaged 38.5 percent of ORF words correct and 39.2 percent of reading comprehension questions correct. Students showed the greatest gains in the critical metrics of oral reading fluency and reading comprehension.

[15] Descriptive analysis will be performed for this research question due to a reduced sample size in 2022 and the variation of intervention dosage among the selected sample.

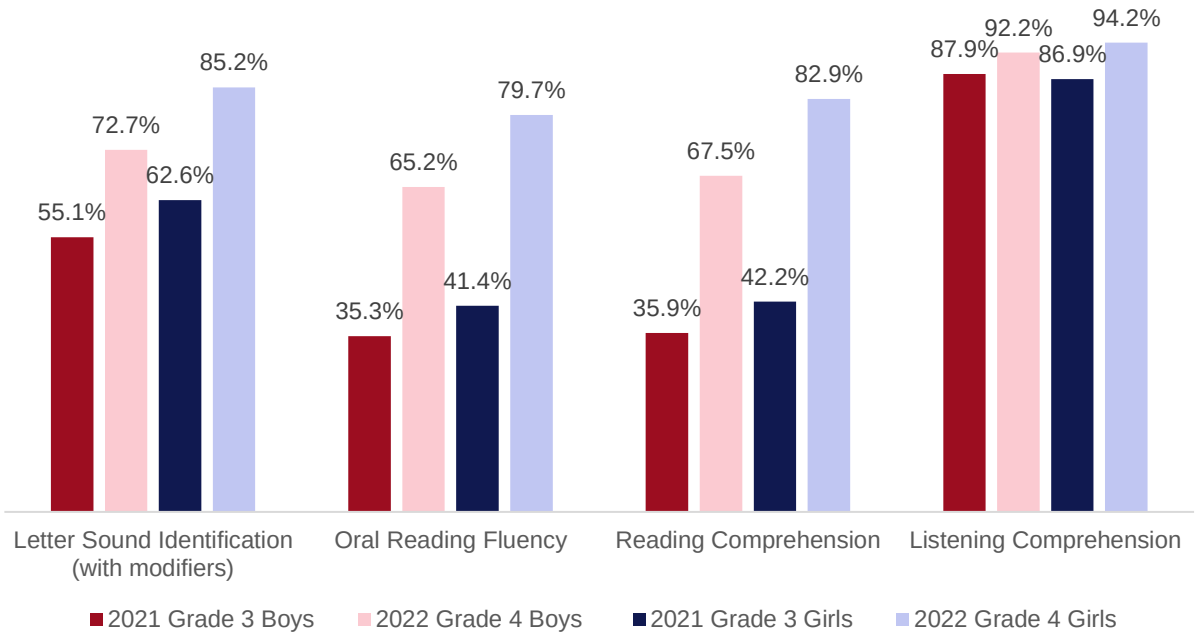
[16] A ceiling effect occurs when a high proportion of students in the sample have maximum scores on the observed variable. The ceiling effect makes it difficult to discriminate the performance among the students achieving maximum scores.

Figure 4: Reading accuracy scores by subtask and year



By gender. A higher proportion of both boys and girls in grade 4 met the reading proficiency benchmark of scoring 80 percent or higher on reading comprehension than grade 3. In grade 4, 58.8 percent of boys met the benchmark—compared to 23.5 percent in grade 3. Among girls in grade 4, 77.9 percent met the reading benchmark, compared to 27.2 percent of grade 3 girls. The same trend was seen for accuracy scores (see Figure 5). Boys and girls in grade 4 had higher accuracy scores on every subtask than boys and girls in grade 3. Both boys and girls showed the greatest gains for the key subtests of oral reading fluency and reading comprehension.

Figure 5: Reading accuracy scores by subtask, grade, and gender

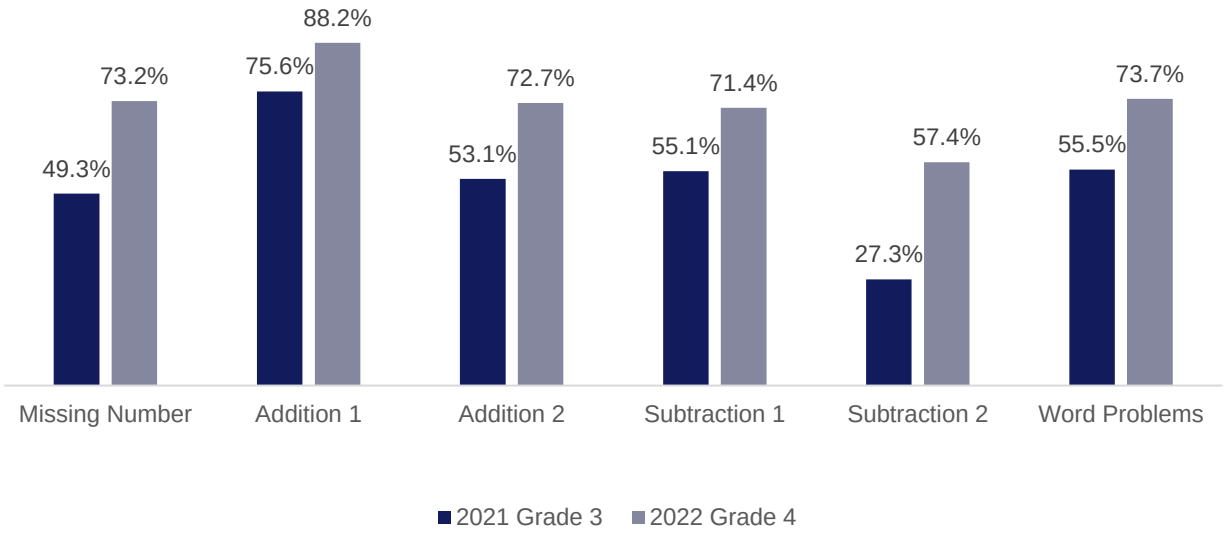


Mathematics Outcomes

Grade 4 students had markedly higher mathematics scores compared to grade 3 students. In EGMA fluency scores, grade 4 students in 2022 had 12.9 correct addition problems per minute (CADDPM); and 8.5 correct subtraction problems per minute (CSUBPM). In 2021, students had 9.4 CADDPM and 5.9 CSUBPM.

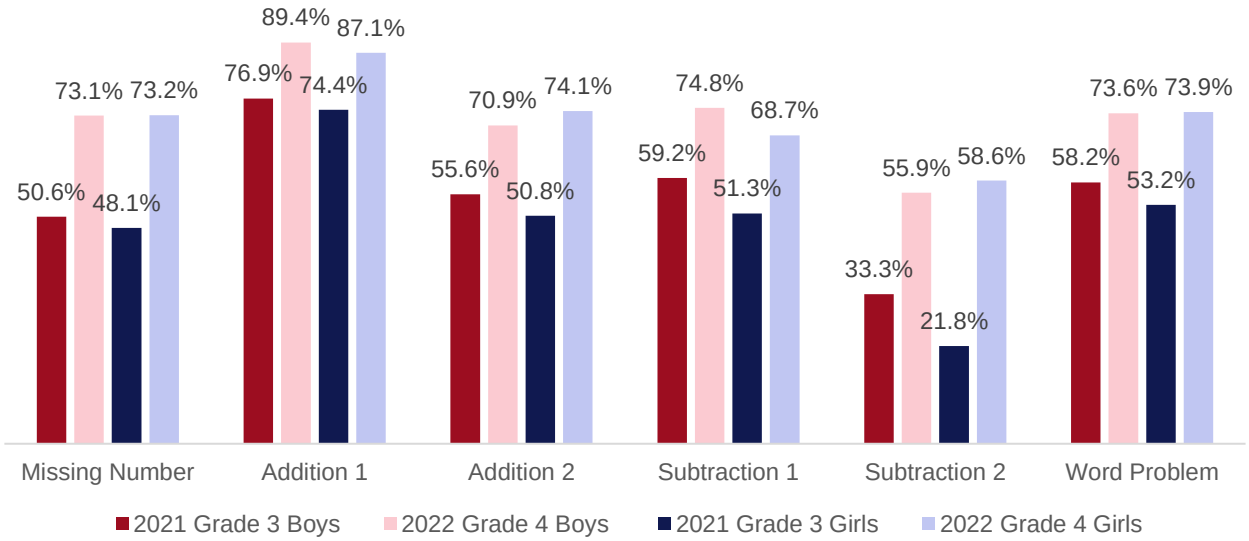
In accuracy scores, students in 2022 again had higher scores than students in 2021 in every EGMA subtask, as shown in Figure 6, with notable gains in the more advanced subtasks of addition, subtraction, and word problems.

Figure 6: Mathematics accuracy scores by subtask and year



By gender. A higher proportion of boys and girls in grade 4 had higher fluency and accuracy scores than boys and girls in grade 3 (see Figure 7). This further supports the trend that students, irrespective of gender, in grade 4 had substantially improved their learning since grade 3.

Figure 7: Mathematics accuracy scores by subtask, year, and sex



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 two different cohorts at the same time point early in the 2022/23 academic year using the same tests but with the different levels of exposure to the

intervention (two years for grade 3; three years for grade 4). For comparability, this question assumes that students have been enrolled in schools with continuous exposure to Manahel interventions since they enrolled in grade 1, and that cohorts are comparable in relation to 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 and churn. 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.

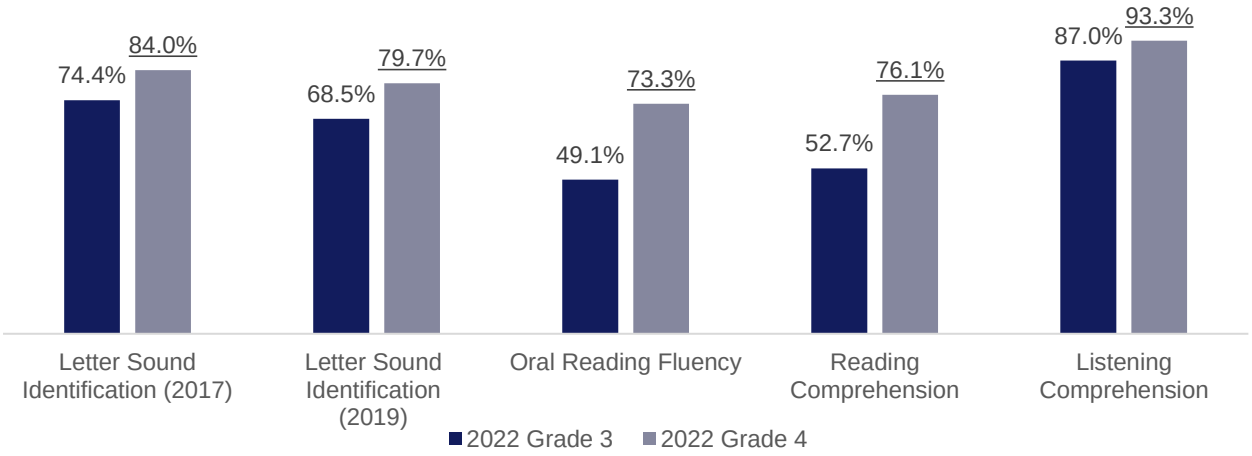
Reading Outcomes

Overall, grade 4 students significantly outperformed grade 3 students in every EGRA subtask (see Figure 8**Error! Reference source not found.**). 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 had an average accuracy score of 73.3 percent on ORF compared to 49.1 percent for grade 3 students and 76.1 percent for reading comprehension compared to 52.7 percent for grade 3 students.

A lower proportion of grade 4 students received zero scores in all subtasks. This was significant for all subtasks, with the exception of 2017 (without modifier) Letter Names and Listening Comprehension – the simplest subtasks with the lowest proportion of zero scores in each grade. For example, 20.8 percent of Grade 3 students received a zero score in reading comprehension, compared to only 7.6 percent of Grade 4 students.

Zero score results for individual subtasks can be found in Annex F, Table F11.

Figure 8: 2022 reading accuracy scores by grade

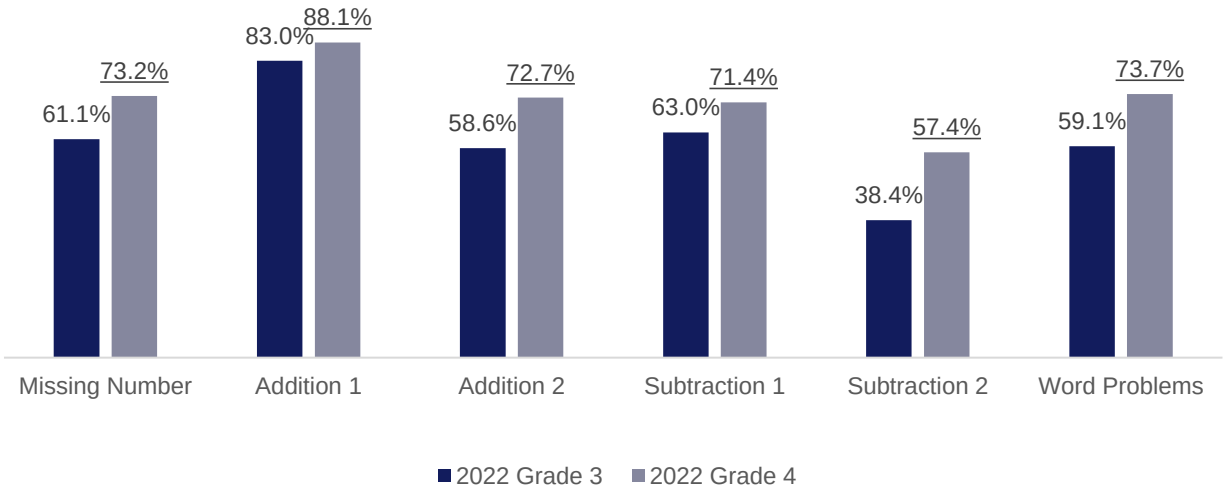


Note: Underlined scores are statistically significant between grades at p<0.05.

Mathematics Outcomes

Overall, grade 4 students significantly outperformed grade 3 students in every EGMA subtask**.** Results are presented in Figure 9below. Some of the largest gains were addition 2, subtraction 2, and word problems – subtasks that measure more complex mathematics skills.

Figure 9: 2022 mathematics accuracy scores by subtask and grade



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

In analysing the jump in performance between grade 3 and 4 it is important to note that under ED rules weak students who fail the end of year tests in grade 3 may be held back and repeat the grade. This will help ensure that some students who are failing in grade 3 do not transition to grade 4. This could indicate that grade 4 students who still cannot read already spent two years in grade 3 or were automatically transitioned to grade 4.

RESEARCH QUESTION 4: NON-READERS

Research Question 4: What proportion of Grade 3 students in 2022 are classified as 'non-readers' and what proportion cannot complete simple (addition) mathematical calculations (zero scores in reading and mathematics tasks)?

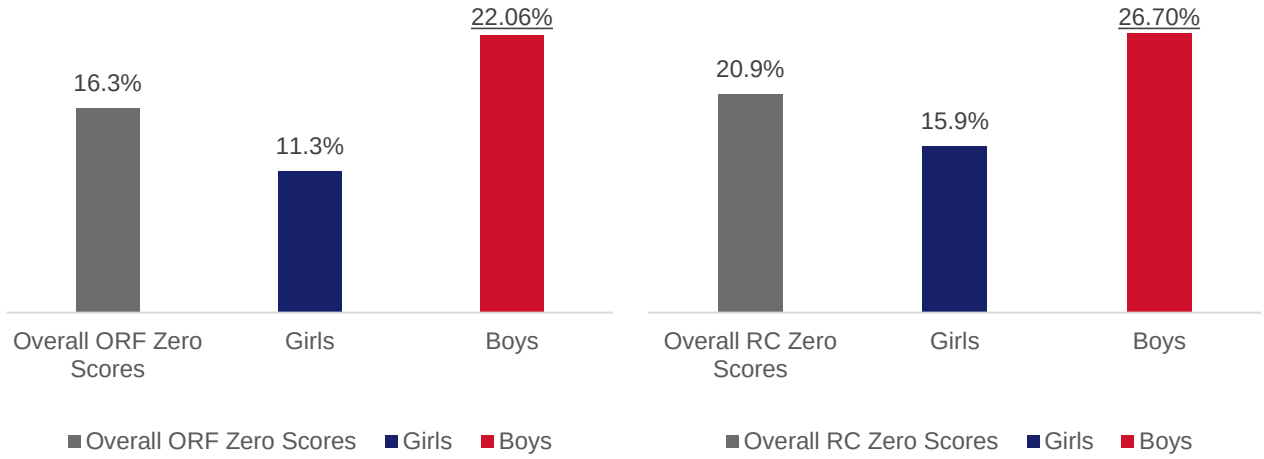
The purpose of this research question is to measure the proportion of learners who are failing to learn to read and understand the simplest of mathematical concepts after at least two full years in school. This will also serve as an important metric for SEP II.

Reading Outcomes

Overall, 16.3 percent of learners in Grade 3 could not read a single word of grade-level text, and 20.9 percent of learners could not answer a single reading comprehension question. A significantly higher proportion of boys received zero scores in both oral reading fluency and reading comprehension in grade 3. This indicates that girls are performing better than boys in both these complex reading subtasks and that a higher proportion of boys have failed to master essential decoding and reading skills.

Zero score results for individual subtasks can be found in Annex F.

Figure 10: Proportion of zero scores for oral reading fluency and reading comprehension



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

Further, analysis by province revealed no significant differences[17] between the proportion of boys and girls receiving zero scores in both oral reading fluency and reading comprehension in Aleppo, where 31.9 percent (n=17) of boys received a zero score in ORF and 38.9 percent (n=22) received a zero score in reading comprehension, compared to 22.5 percent (n=13) and 28.4 percent (n=16) of girls, respectively (see Annex F for zero score results by individual subtask, province, grade, and gender).

In Idleb the gap between boys and girls who are non-readers is particularly stark, with 18.4 percent of boys and 6.6 percent of girls being unable to read a word of a grade level text and 22.1 percent boys and 10.7 percent girls being unable to answer any reading comprehension question correctly.

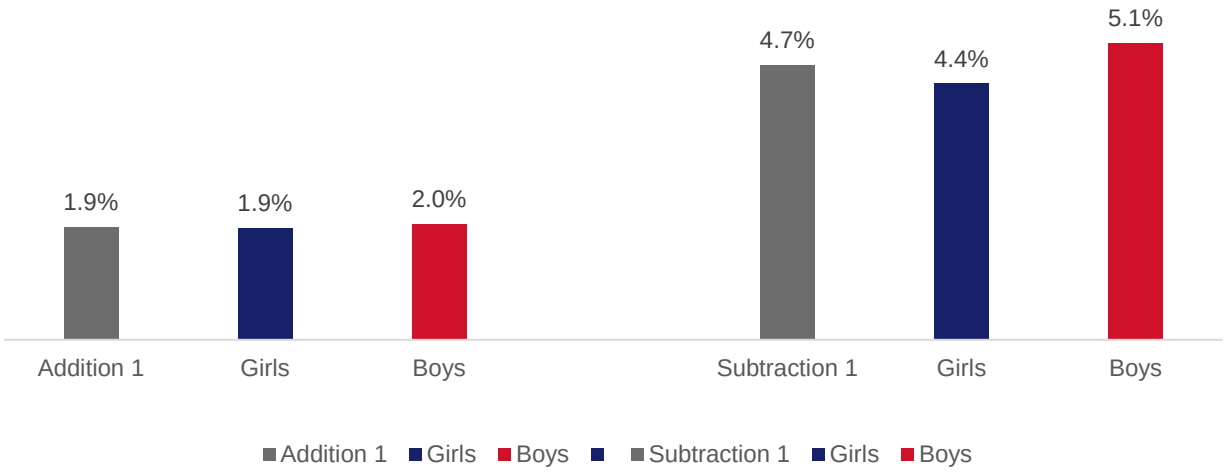
Mathematics Outcomes

The RQ focuses in on the two simplest of the mathematical functions, assuming that if a student cannot do a single digit addition sum or a single digit subtraction sum then they will not be able to do more complex calculations. Overall, single digit subtraction has a higher proportion of students scoring zero (4.7 percent of learners, n=21) as compared to single digit addition (1.9 percent, n=7), indicating that weak students find

[17] The lack of statistical significance is most likely due to the small sample size of students in Aleppo in each of these grades, so results should be interpreted with caution.

subtraction harder than addition. However, no significant differences are observed between boys and girls or any further disaggregation on both these subtasks due to the very low proportion of zero scores (see Annex F for zero score results by grade, gender, and province).

Figure 11: Proportion of zero scores for addition and subtraction



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

RESEARCH QUESTION 5: GENDER GAP

Research Question 5: 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?

This research question aims to understand the differences in boys' and girls' performance for students in grades 3 and 4 in 2022.

Reading Outcomes

In 2022, 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, which is not a reading skill) and in grade 4 (again, except for listening comprehension). Table 5 shows boys' and girls' accuracy scores by grade and the difference between boys' and girls' scores.

For the most basic reading skills measured by letter sound identification and listening comprehension, a bigger gender gap was observed in grade 3. For the advanced reading skills such as oral reading fluency and reading comprehension, the bigger gender gap was observed in grade 4. In grade 3, the difference between girls' and boys' ORF scores was 12.8 percentage points, while it grew to 14.5 percentage points in grade 4. Similarly, the difference between girls' and boys' reading comprehension scores was 12.5 percentage points in grade 3 and 15.4 percentage points in grade 4.

Table 5: 2022 reading accuracy scores by grade and gender

Reading Subtask	Grade 3			Grade 4			Bigger Gender Gap in…
	Boys	Girls	Difference (Girls - Boys)	Boys	Girls	Difference (Girls - Boys)	
Letter Sound Identification 2017*†	67.4%	80.3%	12.9%	78.6%	88.3%	9.6%	Grade 3

Reading Subtask	Grade 3			Grade 4			Bigger Gender Gap in…
	Boys	Girls	Difference (Girls - Boys)	Boys	Girls	Difference (Girls - Boys)	
Letter Sound Identification 2019*†	60.1%	75.6%	15.5%	72.7%	85.2%	12.5%	Grade 3
Oral Reading Fluency*†	42.1%	54.9%	12.8%	65.2%	79.7%	14.5%	Grade 4
Reading Comprehension*†	45.9%	58.4%	12.5%	67.5%	82.9%	15.4%	Grade 4
Listening Comprehension	85.0%	88.6%	3.6%	92.2%	94.2%	2.0%	Grade 3

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

Mathematics Outcomes

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. As in previous EGMA data collection, boys had higher scores than girls in all mathematics subtasks in grade 3. In grade 3, differences were statistically significant for missing number, addition 1, and subtraction 1, as outlined in Table 6. However, in grade 4, only subtraction 1 scores were significantly different between boys and girls. Grade 4 boys and girls had comparable scores in missing number accuracy, addition 1 and 2 accuracy, subtraction 2 accuracy, and word problem accuracy. Boys still outperformed girls on the more basic maths functions.

Table 6: 2022 mathematics accuracy scores by grade and gender

Maths Subtask	Grade 3			Grade 4			Bigger Gender Gap in…
	Boys	Girls	Difference (Boys - Girls)	Boys	Girls	Difference (Boys - Girls)	
Missing Number*	64.5%	58.3%	6.2%	73.1%	73.2%	-0.1%	Grade 3
Addition 1*	86.7%	79.9%	6.8%	89.4%	87.1%	2.3%	Grade 3
Addition 2	61.1%	56.4%	4.7%	70.9%	74.1%	-3.2%	Grade 3
Subtraction 1*†	69.5%	57.6%	11.9%	74.8%	68.7%	6.1%	Grade 3
Subtraction 2	42.1%	35.2%	6.9%	55.9%	58.6%	-2.7%	Grade 3
Word Problems	61.5%	57.1%	4.4%	73.6%	73.9%	-0.3%	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.

DISCUSSION

This section presents key findings from the results of the five 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 2020 to 2021 to 2022. There were significant differences between the reading proficiency classification of students in grade 3 in 2021 and students in grade 3 in 2022. In 2021, 25.4 percent of students met the reading proficiency benchmark of scoring 80.0 percent or higher on reading comprehension compared to 37.4 percent in 2022. This difference (68 percent improvement) was statistically significant. The overall increase in students' reading proficiency levels in 2022, compared to 2020 and 2021, may be due to the effectiveness of interventions following the 2020 COVID-19 school closures. Additionally, survey data revealed a lower proportion of non-readers for students who reported reading with a caregiver, completing online modules, and Skype chatting with teachers during COVID-19 school closures, compared to those who did not report engaging in these activities. Particular attention needs to be paid in future to the failure of between a quarter and a fifth of boys to learn to read by grade 3. This proportion of boys who cannot decode a single word in a grade level text has remained stubbornly consistent over the last three years of the project and needs specific targeted intervention and remediation, from grade 1 onwards, using improved and varied teaching methodologies.

In 2022, as in 2021 and 2020, proficient readers were more likely to come from Idleb than Aleppo. However, in 2022, this gap had closed dramatically with 39.2 percent of students in Idleb being proficient readers, compared to 32.6 percent of Aleppo students. For Idleb this was an improvement of over 11 percentage points, while for Aleppo it was an improvement of over 21 percentage points. This improvement in Aleppo could be due to the focal points being generally better qualified with greater field-level expertise than their Idleb counterparts, and Aleppo ED being particularly committed to the project outcomes. However, there is need for further exploration to understand how such a significant improvement was achieved.

Like 2021 assessment, results showed girls in grades 3 and 4 outperforming boys in all reading subtasks in 2022. In addition, a significantly higher proportion of boys than girls received zero scores in oral reading fluency and reading comprehension in grade 3, further highlighting better performance by girls in these higher order reading subtasks. Results also point to the difference between boys' and girls' literacy outcomes expanding in these two key subtasks as students transition from grade 3 to grade 4. While the 2022 learning assessment did not collect contextual data beyond reading and mathematics outcomes, 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. Recently, the programme also reported an increased interest in girl child education amongst parents and caregivers. This is related to the use of girls' education centres, which use school classrooms after school. The programme targets Grade 5 and 6 girls who are at risk. These sessions focus on literacy and numeracy and psychosocial support along with life and parenting skills and peer support. Their caregivers are invited to participate in these sessions. The programme has over 1500 girls and 1300 caregivers registered in the centres. Parents and girls have provided testimony to how this innovation has encouraged girls to perform better in class and caregivers to treat their female charges and daughters differently – understanding the value of school, learning and ambition. As a result of the programme parents reported that their daughters show more ambition, discipline and enjoyment in life. Such attitudinal changes are likely to positively impact younger girls in the same families and in the broader community.

Results indicate that grade 4 students performed at 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. The improvements were pronounced in ORF and reading comprehension, where accuracy scores improved from 38.5 percent in 2021 to 73.3 percent in

2022 in ORF; and 39.2 percent correct to 76.1 percent correct in reading comprehension. The grade 4 progress in these areas might be attributed to the effectiveness of the Manahel interventions and/or another year of exposure to interventions for grade 4, as well as some filtering of weak learners to repeat grade 3.

The improvement can be attributed to a number of factors, including the role of increasingly competent lead teachers in Manahel schools leading in implementing literacy and numeracy supported by programme related factors including the introduction of additional literacy and numeracy lessons in the early grades and learning circles for teachers; following up on struggling students; peer learning among students under the supervision of their teachers; teachers gaining in confidence in teaching and remediating literacy and numeracy; and the Bank of Ideas Facebook page where teachers upload videos of their teaching and share ideas on the teaching of literacy and numeracy as well as sharing motivational testaments to their impact on students' reading and numeracy skills; and the reading challenge campaigns. In addition, Manahel's surveys indicate that parents and society's view of children's education has become more positive. One programme that Manahel believes has had a fundamental impact is the 'One Million Pages Reading Challenge' it initiated. This has mobilised schools, teachers and students in a competitive space to get excited about reading as they try and out-perform other students and schools. The challenge has captured the imagination and led to teachers creating charts to record student page counts and uploading pictures and videos of their students reading.

As mentioned in the 2021 report, the effort and concentration in encouraging students in Aleppo after the worrying performance in the 2020 assessment seems to have had success in bringing their performance on par with students in Idleb. In 2022, no significant differences were found between two provinces in the reading profiles of learners. Overall, the programme's and ED's efforts to encourage struggling readers in Aleppo seems to have succeeded in improving students' reading proficiency. This includes ensuring that Aleppo schools get the same package and programme dosage as Idleb schools and that only qualified personnel become lead teachers along with stabilising school communities and providing a supportive environment for the intervention.

Mathematics

Maths results between grade 3 and 4 students showed similar trends as reading with steady improvement from grade 3 to 4, especially in the higher-level skills of double-digit addition, double-digit subtraction, and word problems. This relates to grade 4 students being exposed to higher order mathematical problems in class. However, it is impossible to tell if students in grade 4 are performing at the correct grade level given that both grade 3 and grade 4 students are assessed on the same EGMA.

In line with the data from the 2021 assessment, in 2022, boys in grade 3 outperformed girls in all mathematics subtasks, however, in 2022 grade 4 boys and girls performed comparably on all subtasks except subtraction 1. The gap between boys' and girls' scores on every single subtask was greater in grade 3 than grade 4.

While 8 students (2 percent) could not do any simple addition sum and 19 students (4.7 percent) could not do any simple subtraction no significant differences were observed in the zero scores for grade 3 boys and girls. It is also clear that 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 and performing better in some subtasks (though the difference is not statistically significant). This narrowing of the gap may be due to specific programme efforts aimed at encouraging girls in mathematics.

CONCLUSION AND RECOMMENDATIONS

This section uses the analysis from the findings and discussion to draw conclusions, make recommendations, and provide some possible next steps based on this evidence.

The key conclusion is that students in the schools that Manahel supports in NW Syria continue to show substantial and sustained progress in both literacy and numeracy. This is indicated in various metrics.

The grade 3 reading results show a significant and consistent improvement over the last three years. There has been a reduction in the grade 3 students who are non-readers or beginning readers from just under 80 percent of students in 2020 to a little over 50 percent in 2022. This would indicate that while most students after three years in school were unable to read grade level texts in 2020, by 2022 half can. This was achieved during a period of recovery from the impact of school lockdowns during COVID-19 and an improvement in the security situation. As discussed above this improvement can be attributed to the project, to improved security, and to the role of the EDs. However, while celebrating the overall improvement, attention needs to be paid to the consistent failure of over a fifth of grade 3 boys to read a single word.

Furthermore, the gap between the performance of students in grade 3 and grade 4 remained large. Students in grade 4 had an average accuracy score of 73.3 percent on oral reading fluency, compared to 49.1 percent for grade 3 students and 76.1 percent accuracy for reading comprehension compared to 52.7 percent for grade 3 students. This would seem to imply that those students who progress to grade 4 have built solid reading foundations in the earlier grades which they then grow in grade 4. It may also reflect the impact of weak students being held back for an extra year in Grade 3.

Another interesting result is that the reading results from Aleppo have improved significantly. The security situation in Aleppo has improved over the last few years and is now similar to that in Idleb. In addition, the closing of the gap may also relate to the way that the Aleppo ED is engaging with the project and the quality of the focal points the ED has deployed, through replacing all the unqualified lead teachers with qualified ones.

Overall, in 2022, grade 3 girls significantly outperformed boys in all reading skills and continued to do so in grade 4. In 2020 28.7 percent of girls could read with some fluency, by 2022 this was 55.2 percent. Perhaps the most significant result is that while boys have performed consistently better than girls in mathematics since 2020, girls are now closing the gap and, by grade 4, the 2022 results show that gap has almost completely closed, with girls performing slightly better than boys in all the higher order mathematics tests.

While the sample size was relatively small, being 400 students across grade 3 and 4 so the findings need to be treated with care, we can conclude that this is a significant achievement for the programme. This is particularly so, given the protracted conflict in the area and that school closures across the world during the pandemic have resulted in learning loss and tended to lead to slower learning trajectories in the subsequent years. The remote learning interventions that Manahel put in place early in the pandemic in 2020, along with other in-class support activities, teaching and learning materials, increased literacy and numeracy classes, peer learning, learning circles for teachers, literacy clubs and competitions seem to be working effectively. Overall, students in the programme have shown progress in both literacy and numeracy outcomes year on year.

Although reading and maths scores have held up reasonably well and generally improved between 2020 and 2022, 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 students with the foundations that they require for engaging with the upper primary curriculum, and for their future education and work lives.

The recommendations are that follow are structured around the level at which they could be implemented: school, system, and by Manahel. Given the budget constraints, which limited the analysis carried out in this report, and the end of the present project cycle looming, this report makes recommendations for future innovations, research and analysis.

SCHOOL RELATED RECOMMENDATIONS

1. Early grade reading scores using EGRA show a significant improvement from 2020 to 2022, indicating that teachers, supported by the lead teachers, are effectively implementing classroom teaching practices and the additional literacy sessions that the Education Directorates have introduced with Manahel's support. However, a stronger focus on the use of continuous assessment, along with the introduction of teacher-led early grade reading assessment (TEGRA), to identify and track struggling students and support in developing in-school strategies of remediation will be helpful in guiding teachers towards a targeted teaching approach focused on specific skills that students struggle with. The learnings from continuous assessment could help teachers in differentiating instruction according to student needs. This will support students most at risk of not gaining basic and higher-level reading skills. Such interventions and remediation are particularly required to help the fifth of all grade 3 boys who have failed to learn to read even a word after three years in school.
2. Through the end of the programme and in SEP II, early grade teachers should work with boys to build reading fluency and their foundation towards reading proficiency. In grade 3, with students – particularly boys – who cannot read even a word of a grade level text 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. This should be supplemented with teachers adopting methods and incentives which appeal to boys and help keep them in school and learning. 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.
3. Teachers in grades 1 and 2 should make sure that all students have understood the basic mathematical functions (addition/subtraction) while teachers in grade 3 should focus on more complex mathematics skills to ensure that students master mathematics operations and real-world thinking and are therefore better prepared for the more complex mathematics taught in grade 4.

PROJECT RELATED RECOMMENDATIONS

4. Redacted

SYSTEM RELATED RECOMMDENDATIONS

5. All of the above school-based recommendations will be more successful if supported from within the system. In particular, the EDs, Education Assemblies and school leaders should assist schools in implementing TEGRA so that they can interpret and analyse their continuous and termly assessment results in the early grades to inform their teaching and to build remediation measures around areas of weakness. This will in turn help solve the gender-related performance gaps with boys improving their reading performance and girls their performance in numeracy.

RECOMMENDATIONS FOR FUTURE RESEARCH AND LEARNING

The research and learning agenda should focus on a 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)

Further research should be undertaken to investigate the relationship between student survey responses regarding attendance, displacement, and access to literacy sessions and non-readers and/or students who cannot solve simple mathematical calculations. This will require the tracking of individual students over years. Building on this data is the need to know if the same students are struggling with both reading and numeracy and the under-researched interplay between the two skills.

A key part of the learning agenda for SEP II should be action research around the introduction of TEGRA, which was discussed above. The results from TEGRA and the gap in performance between students when

tested by their teachers and with a specialised team of male enumerators from outside the school should be explored with teachers and field officers being active participants in the research.

In 2022, grade 4 students showed ceiling effects on the listening comprehension task. Future research needs to focus on the reliability and validity of the EGRA assessments. In SEP II, a scientific process is required to develop grade-specific annual and semester-based reading and numeracy benchmarks and proficiency levels. 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. While this would take some time, SEP II could consider working with the Education Directorates to align the proficiency levels with the global proficiency framework.

SEP II should observe grade 2 and 3 teachers in mathematics classes to analyse their interaction with both boys and girls to see if the actions and biases of the teachers are related to female under-performance. This should focus on who is getting asked questions, who is speaking in small group work, who is coming up to the board, whose work is being celebrated etc. If this action research finds that boys are being advantaged in these classes SEP II should feed the findings into learning circles for mathematics teachers where the impact of teacher attitudes and actions can be exposed, and strategies discussed to assist mathematics teachers change their attitudes and actions.

FCDO could consider commissioning research into the processes that underpin the progress of students, from pre-school to grade 1, from grade 1 to grade 2, from grade 2 to 3, and grade 3 to 4. The steep and significant improvements in performance by grade 4 students over grade 3 students in this 2022 Learner Assessment may be in part a result of selection of students to progress to grade 4 based on their test results, with weak students being held back in grade 3 to repeat year. This needs to be explored and clarified.

IMPORTANCE OF TEACHER PAY

SEP II could administer the EGRA/EGMA in a random sample of schools that the project is supporting but also in non-SEP schools supported by other NGOs as well as in schools in which teachers are not being paid by the project and a randomized sample of schools where teachers are being paid. This would provide some measure of understanding of how different school support interventions and providing teacher pay impacts learner performance.

POSITIVE DEVIANCE STUDIES

Manahel contends that SEP II undertake detailed case studies of individual project schools which have seen a robust improvement of learner results 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. This research would be shared with the Education Directorates and the instructors in order for them to understand the elements which are related to high performance on difficult circumstances. While transferring such lessons from one school to another is problematic, if the EDs and instructors are aware of these lessons they can help and nudge under-performing schools towards adopting such lessons.

STUDENT GENDER AND VULNERABILITY RESEARCH

Manahel has implemented a GESI review during the extension period. This review provides a status report and reflects on GESI focused activities that can be applied during future programming. It is also proposed that SEP II undertakes small-scale research to understand if girls' well-being is comparable to boys in the later years and widen the Manahel 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.

SEP II could also prioritise introducing learning circles to support teachers to create gender-responsive pedagogy and a growth mindset, and measure how these interventions are perceived. Given that boys are struggling with reading skills to a greater extent than girls learning circles, supported by research, are needed to explore incentives and special activities that appeal to boys and get them motivated to learn, compete and persevere in school.

Finally, SEP II could undertake a study to explore attendance by girls in the early grades 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.