

LET'S READ FLUENTLY

PILOT EVALUATION REPORT

21 APRIL 2023



مؤسسة الملكة رانيا
QUEEN RANIA FOUNDATION

Contents

Executive summary.....	4
Introduction.....	8
Methods	27
Impact Evaluation - Findings	38
Implementation and Process Evaluation – Findings	55
Conclusions.....	85
References.....	91
Appendix A: Additional Descriptive Subgroup Analysis.....	93
Appendix B: Zero Score Analysis for Secondary Outcome Measures	95
Appendix C: Data collection tools	101
Appendix D: Criteria for moving to trial phase	154

ABOUT THE EVALUATOR

The evaluation team includes the following staff from NatCen, as well as Integrated, and Oxford MeasurEd, School-to-School International:

Sashka Dimova, Hannah Woodbridge, Enes Duysak, Eliza Garwood, Julia Ruddick-Trentmann, Natasha Phillips, Tien-Li Kuo, Andi Fugard, Julia Griggs, Daniel Phillips, Nedjma Koval, Leen Al Refai, Zeid Qiblawi, Rachel Outhred, Lydia Marshal, Matthew Murray, Carol da Silvia.

The lead evaluator was Sashka Dimova.

ACKNOWLEDGMENTS

We would like to thank all the teachers, parents and pupils at the schools participating in the study for their time and sincere answers about their experiences. We also appreciate the efforts of the coaches responsible for delivering the 'Let's Read Fluently!' intervention. We are also grateful to the core team at Queen Rania Foundation (Maysoon Masoud, Emilee Rauschenberger, Rami Al Assad) for supporting the evaluation, providing guidance and for supplying recruitment information. We would also like to thank the team at the Education Endowment Foundation led by Christine Kelly for guidance and advice. Many other colleagues were involved in setting up this evaluation and we are grateful to Dr Helen Abadzi and Julie Helson for their contributions throughout this evaluation.

This research project has been funded by the Education Endowment Foundation, in partnership with the BHP Foundation, as part of the "Building a global evidence ecosystem for teaching" project.

Executive summary

Let's Read Fluently!

'Let's Read Fluently!' (LRF!) is an intervention centred around a practice-focused pedagogy method and reading practice book with the main aim of supporting children in developing foundational literacy skills in Arabic. In this pilot evaluation of LRF! two different approaches of the intervention were delivered: a Whole Class (W/C) model, delivered to all pupils in a classroom, and a Literacy Catch-Up (C/U) model, which targeted the lowest-achieving 20% of pupils in a class. Both approaches were delivered by teachers for one semester via three 30-minute classes each week.

The pilot evaluation was intended to inform, and test the intervention's feasibility and evidence of promise, and assess readiness for trial. The pilot was also intended to provide preliminary evidence on the impact of LRF!, the mechanisms of change and lessons to inform future scale-up.

The pilot started in September 2021 with recruitment of schools and randomisation. Delivery of LRF! for the C/U model in Grades 2 and 3 took place in semester 1, while delivery of the W/C and C/U Grade 1 took place in semester 2 of the 2021/22 academic year. The pilot study used a randomised design, with schools randomised into two treatment arms and one control arm. The evaluation reached a total of 587 pupils (W/C: 180; C/U: 114; Control: 294 pupils) in 24 schools (W/C: 8; C/U: 8; Control: 8) that met the eligibility criteria.

This pilot was funded by Queen Rania Foundation (QRF) and was supported by the Education Endowment Foundation (EEF) in partnership with the BHP Foundation, as part of the "Building a global evidence ecosystem for teaching" project. Queen Rania Teacher Academy (QRTA) led on the implementation of the intervention, including training. The evaluation was undertaken by a consortium of partners made of Integrated, Oxford MeasurEd (OM), School-to-School (STS) and led by the National Centre for Social Research (NatCen).

Key conclusions

- The evidence from the impact evaluation suggests that delivery of the W/C approach could improve pupils' literacy attainment. In the case of the C/U approach, there is no evidence of improvement. However, the pilot RCT was small and was not designed to measure impact robustly.
- Perceived outcomes, as reported by teachers, coaches and parents indicate improvement in pupils' literacy, engagement and confidence with reading. Evidence for the C/U approach was mixed. Evidence also suggested some signs of a potentially adverse psychological impact on Grade 1 pupils in the C/U model.
- LRF! was feasible to deliver, and most schools delivered LRF! as intended with small adaptations. This mainly included allowing more time for delivery through lengthening LRF! sessions, adjusting the pace for pupils with lower literacy ability, or providing individual coaching to C/U pupils who struggled with the content.

- The pilot identified a number of potential improvements needed to the interventions, including modifications to the following components: the training for C/U resource room teachers, the format for LRF! delivery, the content of the practice book, and the selection process for C/U pupils.
- Evidence in relation to the evaluation procedures indicates that a clustered RCT design with allocation to the school level will be suitable for a future scaled evaluation. There were no challenges in relation to recruitment, randomisation and retention of schools. However, the reach data indicates that it was easier to recruit schools in the near south region in comparison to schools in Amman, the middle excluding Amman region and the near north region. Recruitment for any future evaluation should consider these regional differences.
- The inconsistencies in evidence of promise and the suggestions for improvements in relation to the feasibility of C/U suggest that further development is required before C/U is reassessed for readiness for trial. Particular attention must be paid to the training for resource room teachers, and to adaptations needed for resource room delivery.

How was the pilot conducted?

To provide insights that can inform future trial design, the evaluation was designed as an experimental study, with schools randomly allocated to one of the pilot arms (W/C, C/U or control). A comprehensive implementation and process evaluation (IPE) using focus groups discussions (FGDs) with teachers, parents and pupils, interviews with school stakeholders, classroom observations, a teacher survey and a pupil survey was carried out to provide insights to the research questions.

The primary aim of the impact evaluation (IE) component was to assess the suitability of a future experimental design and the appropriateness of the selected outcome measures. Given that a small number of schools were participating in the pilot, the IE was not designed to detect significant differences between the treatment and control schools, but it was used to provide an initial assessment of the potential impact and to assess support for the theory of change. The primary outcome was Arabic literacy attainment based on a set of pre-literacy items and the Jordanian Early Grade Reading Assessment (EGRA, hereafter EGRA+pre-lit), whereas the secondary outcomes included letter sound identification, speed and accuracy of word decoding and reading comprehension derived from the EGRA+pre-lit sub domains.

The W/C model was delivered to Grade 1 children in Semester 2 of the 2021-22 academic year, while the C/U model was implemented with children in Grades 2 and 3 in Semester 1, and with Grade 1 in Semester 2 of the 2021-22 academic year. The evaluation of the pilot started in July 2021.

What are the findings?

This pilot evaluation aimed to investigate evidence of promise, feasibility of the intervention, feasibility of the efficacy trial and readiness for trial. Detailed findings for each dimension are presented in Appendix D.

Evidence of promise:

In the case of the W/C approach, findings from the impact and the IPE showed evidence of promise. The IE estimated that children in W/C schools made greater improvements in literacy in comparison to children in the control schools. Teachers reported that the W/C approach had a positive impact on pupils' reading comprehension, and pupils' engagement and confidence in reading. This was corroborated by coaches and parents, who perceived positive changes in children's literacy. Overall, findings from the IE and IPE data suggest that the W/C model can improve literacy attainment.

Findings for the C/U approach were more mixed. The IE found no differences in pupils' literacy attainment between schools assigned to implement C/U and control schools. This was the case for pupils in all grades. Findings from the surveys and FGDs indicate that some teachers observed improvements in pupil performance and confidence in reading, particularly amongst Grade 3 pupils in C/U schools. However, resource room teachers for Grade 1 pupils in C/U schools were far less positive in their survey responses. Similar to teachers, coaches also observed more positive impact for pupils in Grades 2 and 3. Coaches also noted that LRF! could have some negative psychological impacts for Grade 1 C/U pupils, as the material was too difficult for them. Findings from parents and pupil reports also suggest that LRF! could have been too difficult for children in C/U Grade 1.

The findings from the impact evaluation, however, should be treated with caution and should not be interpreted as showing that there has or has not been an impact with high confidence. This is because, due to the small sample size, secure conclusions cannot be drawn from the impact estimates. Any difference in outcomes cannot be interpreted as being directly attributable to LRF!, but instead as showing indicative evidence of promise.

Feasibility of intervention:

The key inputs and outputs as delivered in the trial are acceptable to schools. Findings from a variety of sources suggest that LRF! was delivered as intended with high attendance at training and most sessions were delivered in both the C/U and W/C approach. The practice book was seen by everyone as a very important resource used during every LRF! session. Almost all children had a copy of the practice book and many of them read the practice book at home. Selection of pupils for the C/U approach was done using a diagnostic tool. Views on the appropriateness of the diagnostic tool were mixed.

Participants made practical suggestions about improvements that could be made to the intervention in the following broad categories:

Training and coaching – Coaches described instances when C/U resource room teachers did not grasp the training, which suggested that they may require additional support. Coaches also recommended adaptations to the training materials for C/U resource room teachers.

Content of the LRF! sessions - Coaches and teachers in both intervention arms felt that some parts of the sessions (e.g. the 'You do' process) were too challenging for some pupils, or the content was difficult to cover in the allocated time. As a result, parts of delivery across the C/U and W/C models were adapted to improve engagement among pupils with lower language ability or by extending the time allocated for each session.

Content of the practice book - Even though the practice book was seen as a useful aid, and all sessions were based on the book, the perceptions on its content were less positive. Parents, coaches and teachers from both intervention arms perceived the material to be too difficult for some pupils in Grade 1 or those struggling with literacy. They thought that different versions of the book for pupils with different abilities would be a solution. They also expressed some concerns about how the content was presented and made suggestions for adaptations.

Diagnostic tool – The majority of teachers and coaches did not think the diagnostic tool was able to identify the lowest performing pupils in Grade 1 or to account for learning difficulties, suggesting that the process for selecting pupils in C/U Grade 1 should be reconsidered. Coaches and teachers also voiced concerns about the test environment of the tool proving unsettling for some pupils in both intervention arms.

Feasibility of the efficacy trial

Some of the key evaluation procedures such as recruitment and randomisation were executed well during the pilot study, suggesting that a clustered RCT design with allocation at the school level will be feasible. Retention rates were high for both the W/C and C/U approaches, and the primary and secondary outcome measures were seen as appropriate to identify progress in literacy.

Assessing readiness for trial

While this evaluation provides evidence to suggest that implementation of the W/C and C/U approaches was successful, it also identified some recommended intervention adaptations or conditions that are needed to make LRF! succeed in an efficacy trial.

In sum, there is indicative evidence that W/C is ready for trial conditional on modifications in respect to the time allocated for LRF! delivery, and in respect to the content of the practice book. While the delivery and evaluation partners felt that they have the capacity and knowledge to deliver and evaluate a scaled-up version of LRF!, they also felt that allowing sufficient time for adaptations will be key to the success of the evaluation.

The pilot found that the C/U approach would require substantial changes before it can be reassessed for readiness for trial.

Additional findings

Despite the reported success in improving literacy attainment, W/C teachers indicated that LRF! had been more beneficial for children with lower reading abilities. Our additional exploratory analysis based on the EGRA assessment also found evidence that LRF! could be more beneficial for lower performing children.

INTRODUCTION

Context: The importance of developing literacy among Arab pupils

A strong foundation in literacy is a crucial element predicting educational success. Evidence shows that early literacy difficulties can persist, limiting children's ability to achieve their potential (Brombacher et al., 2012).

Pupils in the countries that use the Arabic language and script for instruction are performing at a low level in international and internal examinations (Eckert et al, 2020). For example, results using the Early Grade Reading Assessment (EGRA) conducted in Jordan since 2012 have shown that primary school-aged children are failing to reach reading comprehension benchmarks (RTI International, 2018). Existing evidence suggests that it is very unlikely that pupils will make up for learning loss during the next stages of their education, leaving these children at a significant disadvantage throughout their schooling and life (World Bank, 2019). Therefore, providing the right support in the early years of schooling is essential for reducing this 'performance gap'.

Learners of Arabic face unique challenges: the script is comparatively complex, pupils can face visual perception challenges, and pupils use various Arabic variants at home which differ from the formal Modern Standard Arabic (MSA) taught in schools (Abadzi, 2017; Eckert et al, 2020). Pupils entering school are consequently tasked with absorbing the standard Arabic language vocabulary alongside developing their literacy skills (e.g. reading and grammar knowledge) to make sense of a text. Considering the linguistic challenges that readers in Arabic face, it is important to identify approaches that will help pupils with literacy attainment. One of these approaches is Let's Read Fluently! (LRF!).

Context: The Let's Read Fluently! intervention

The LRF! approach involves a practice-focused pedagogy and pupil practice book developed by cognitive psychologist Dr Helen Abadzi and the Al Qasimi Foundation in the UAE. The approach draws on insights from studies in linguistics and cognitive science that account for the Arabic script's visual complexities and the relationship between memory function and reading. It has been developed to help pupils build 'low level' neurological functions - rapidly distinguishing letter shapes, chunking and decoding sounds and words. Similar approaches to early literacy teaching in Cambodia, the Gambia and Egypt have shown evidence of promise. There are also early results from a small-scale pilot conducted in the UAE, which suggest an LRF! approach may have a positive impact on Arabic reading fluency in early grade pupils (Eckert et al, 2020).

There are two models of LRF! implementation, a Whole Class Teaching and Learning approach (W/C) and a Literacy Catch-Up (C/U) approach, both lasting one semester. The W/C model targets pupils in Grade 1 and is delivered to classes via three 30-minute classes per week. In class, the teacher adopts a 'I do', 'we do', 'you do' pedagogical approach using the practice book. The C/U model targets the lowest-achieving 20% of pupils in Grades 1 to 3 and is delivered through small group tuition for 3 sessions per week.

WHY

It is estimated that early readers in Arabic need a level of automaticity¹ in oral reading fluency² of 45-60 words per minute (RTI, 2012). This fluency allows working memory to be freed-up for comprehension. Data from the use of the Early Grade Reading Assessment (EGRA) tool in 2018 suggests that only around 19% of Grade 2 and Grade 3 pupils meet, or exceed, the lowest levels of this benchmark. Alongside that, a significant number of pupils in Jordan (16.6%) scored zero in oral reading fluency in 2018 (RTI, 2018).

Early grade interventions can be beneficial in terms of helping pupils as they progress in grade level. The 2018 Program for International Student Assessment (PISA) found that, for 15-year-old Jordanian pupils', attainment levels were behind the OECD average by an equivalent of more than one grade in reading. Only one in five pupils performed at or around the average OECD reading score and two in every five performed below the minimum proficiency level in reading (QRF, 2020).

There are also concerns with global levels of literacy, and in 2019 the World Bank announced its 'Literacy Makes Sense' approach to reduce what it describes as 'learning poverty'. Within the context of Jordan, the report estimated that 52% of Jordanian 10-year-olds are unable to read and understand a short age-appropriate piece of text.

When diacritics³ are used, Arabic is a transparent language – that is, there is a reliable relationship between letters and sounds. Given this, these low EGRA scores likely reflect a gap in phonics skills (EEF, 2021). The importance of phonics is reflected in the EEF's Teaching and Learning Toolkit and other literature (Seidenberg, 2017; Castles et al, 2018). It is important to note that LRF! and its evaluation are taking place amid the ongoing COVID-19 crisis, which has led to sustained school closures. This context reinforces the need for interventions to support literacy acquisition and strong evidence to understand what works (UNICEF, 2020).

WHO

Teachers are at the core of the LRF! intervention as they both receive training and then deliver the intervention to pupils. W/C classroom teachers deliver LRF! to all students in Grade 1. The C/U model is delivered by resource room teachers⁴ to the lowest-achieving 20% of students in Grades 1 to 3 via small group tuition. Classroom teachers in schools selected for the C/U model attended training but were not directly involved in delivering LRF! to pupils.

Teachers received a one-day training course from the Queen Rania Teacher Academy (QRTA) and up to three follow-up coaching sessions to support them in using the LRF!

¹ Automaticity is defined as being able to complete a task with no conscious effort, in much the same way as you are able to read this footnote or calculate 2×2.

² Oral reading fluency is the ability to read connected text quickly, accurately and with expression. In doing so, there is no noticeable cognitive effort associated with decoding the words on the page.

³ Diacritics are marks placed above or below (or sometimes next to) a letter in a word to indicate the short vowels.

⁴ Resource room teachers are responsible for the optimal use of the resource room, and teaching and providing support to pupils that are struggling in the classroom

practice book. They delivered the LRF! intervention to pupils in Grades 1 to 3. Thus, both pupils and teachers can be considered intervention recipients.

All pupils who were in the second semester of Grade 1 in schools and classes selected for the **W/C implementation model** were eligible for the intervention. The rationale was that introducing the practice book at this stage allows for alignment with the sequence of letter-sound introductions that are set out in the Grade 1 textbooks used in standard literacy teaching (i.e. usual practice).

The lowest achieving 20% of pupils in a class in the second semester of Grade 1, or in the first semester of Grades 2 and 3 (Grades 2/3) and in a class selected for the **C/U implementation model**, were eligible for the intervention. Semester two was the earliest it was deemed feasible to identify struggling readers in Grade 1. Similar to the W/C model, the alignment with the curriculum was a reason for starting LRF! implementation for Grade 1 C/U pupils in semester two. QRF felt that the intervention could be more easily administered to pupils in Grades 2/3 in semester one.

QRF recruited primary schools and QRTA provided training and support to teachers delivering LRF!. School principals and supervisors attended three-hour orientation sessions which informed them about what teachers need to do as part of the intervention and how to equip them with the skills to support implementation.

WHAT

Teacher training and coaching

QRTA trained classroom and resource room teachers on how to use the LRF! method and practice book.

Classroom and resource room teachers attended the one-day face-to-face training. Teacher training and coaching focused on just one of the two implementation models. Teachers were asked to deliver either the W/C or C/U intervention (not both) based on the randomised assignment. The training aimed to provide participants with an understanding of the following:

- the rationale for LRF!;
- teachers' role in delivering LRF!;
- the learning experience teachers are being asked to facilitate;
- how to appropriately communicate the project to parents/carers (including conducting a face-to-face awareness raising meeting for parents);
- how to support the involvement of parents/carers – for example, encouraging pupils' use of the LRF! practice book at home with parents/carers, and supporting parents in this, (i.e. with WhatsApp messages).

The training included opportunities to practice the new teaching and learning techniques, and to explore potential barriers and how they can be overcome. School Supervisors⁵ have no formal role in implementing the intervention but were invited to attend the training

⁵ Based on their specialization, school supervisors assume technical support role to early grades teachers, resource room teachers and subject matter teachers. They also have an evaluation role within the school.

session alongside teachers in their District. Principals of schools in the intervention group attended a separate orientation session.

Following the training sessions, teachers received up to three coaching visits (accompanied by classroom observations from the coaches). The coaching visits were delivered by QRTA staff. To facilitate these, QRTA used a coaching model designed to enhance implementation effectiveness on the part of teachers.

WHAT

Whole Class Model (W/C)

The W/C teaching and learning approach is targeted at pupils in the second semester of Grade 1. The rationale for delivery in the second semester is that introducing the practice book at this stage allows an alignment with the sequence of letter-sound introductions specified in the ‘business-as-usual’ Grade 1 textbooks.

The intervention was delivered in three 30-minute sessions each week, for 12 weeks. Following approval from the Jordanian Ministry of Education (MoE), it was agreed that the classroom time required for this would be taken from one of the three ‘free activity periods’⁶ in the Jordanian school curriculum, and two of the seven Arabic language classes. First, using large versions of the textbook, the classroom teacher introduced the letter-sound, or letter combinations, and modelled how to ‘read’ it. (“I do”). This was followed by an opportunity for the whole class to practice ‘reading’ using either the choral or echo method (“we do”). These two steps were to be completed in the first 10 to 15 minutes of the session. Following this, learners were asked to independently work through the pupil practice book, taking each item in turn and with their finger on the text sounding out the letter, or word (“you do”). At this stage of independent pupil practice, the teacher’s role was to encourage engagement with the task and to provide feedback (namely, reinforcement and corrections). This stage of independent practice with teacher feedback should have been around 15-20 minutes, which is half to two-thirds of the session. This is a key feature of the LRF! model, as research in cognitive science indicates that individuals need to independently and repeatedly practice decoding to develop the automaticity needed for fluent reading (Abadzi & Martelli, 2014).

All pupils received a copy of the practice book. They were encouraged to take it home for extra practice with their parents/carers, with teachers supporting this form of parental engagement in two ways: (1) by raising awareness through an introductory meeting with parents, and (2) through communicating via existing channels (e.g. WhatsApp messages to parents) about the support needed with practice at home following the lesson.

In the two sessions that were conducted in the Arabic classes, LRF! sessions replaced ‘business-as-usual’ teaching. Delivery of LRF! was aligned with the existing curriculum content as much as possible so that pupils had a coherent learning experience.

⁶ There is a set curriculum for the Free activity period, including set activities that teachers implement with students. This primarily focuses on enhancing personal skills and values.

Literacy Catch-Up Model (C/U)

The C/U model targeted the lowest achieving 20% of pupils in a class, either in the second semester of Grade 1 or in the first semester of Grades 2/3. The rationale for the former is that it was felt that the first semester of Grade 1 is too early to identify struggling readers. For the latter, the rationale was simply to help facilitate project planning. The C/U model was intended to be delivered to groups of five-to-six pupils with similar literacy learning needs. In effect, this is a form of extra small group tuition, using the practice book as the learning material.

Selection of pupils for the intervention was carried out by teachers, using the coarse grained diagnostic tool developed by the Early Grade Reading and Mathematics Program (RAMP)⁷ - already in use in Jordanian classrooms. In the teacher training, teachers were advised to exclude pupils who have severe learning difficulties,⁸ as the LRF! intervention was not designed with their needs in mind.⁹ That is, QRTA advised that eligible pupils are those for whom it is plausible that a change in teaching and learning approach will be helpful.

In this model the intervention was delivered by resource room teachers. The scheduling of extra support to struggling learners (referred to as 'Resource Room teaching' in Jordan) was agreed between the classroom teacher and the resource room teacher. Schools have autonomy over this scheduling, and this extra support was provided during the normal school day. The MoE gave permission to use one 'free activity period' for literacy catch-up. Teachers were advised to schedule the remaining two sessions in a way that minimised any disruption to normal learning.

HOW

Pupils are taught to process written text more quickly by firstly repeating individual letters and words to the point of automation. This is intended to enable them to decode reading faster, in order to read more fluently and free up working memory to recall important information and think critically. Time engaged in practice and receiving timely feedback (namely, reinforcement and corrections) are seen as important predictors of reading ability.

The LRF! practice book is designed to encourage perceptual learning for decoding, as well as reading practice to attain fluency. It includes a number of design features intended to tackle barriers to literacy and current understanding about what works for early readers:

- Small font sizes negatively affect letter identification, so the book uses large font sizes and spacing.
- The Arabic script is dense and complex, and so creates a higher cognitive load for new readers than other languages. The book and LRF! model more generally, place importance on repetition and teacher feedback.
- New letter shapes are introduced slowly, one by one.
- It follows a phonics-based approach in which children gradually decode words using their phonics knowledge rather than using other clues or seeking help

⁷ For more information on RAMP see https://pdf.usaid.gov/pdf_docs/PA00THHW.pdf

⁸ The description used in QRTA communications with teachers is that LRF! is suitable for 'pupils who are academically behind but don't suffer from mental or physical illnesses.

⁹ The description used in QRTA communications with teachers is that LRF! is suitable for 'pupils who are academically behind but don't suffer from mental or physical illnesses.

- Pattern analogies can assist learning, so common sounds are stressed (e.g., da di du, which links the 'd' sound with each of the short vowels).
- Pupils need to see meaning in text, so real words and sentences are introduced as soon as possible.
- The use of pictures in the text are minimised to ensure pupils learn letter sounds, rather than guessing.

The pupil practice book stresses repetition of patterns, alongside lots of practice in recognising them. See the examples below:



ب

بَ بٍ بِ	بَبٍ بَبِ
بَبِ بَبِ	بُبٍ بُبِ
بَبِ بَبِ	بُبِ بُبِ
بَبِ بَبِ	بُبِ بُبِ
بَبِ بَبِ	بُبِ بُبِ

الْوَزْدُ الْبَرُّ الْبَدْرُ

الْيَدُ الْبِنْتُ الْبِنَاتُ

الْبَالُونُ الْبَرْدُ

الْوَزْدُ الْبَرُّ الْلَيْلُ

Repetition of sounds

The practice book includes text with subtle differences to encourage pupils to recognise common words, even when presented slightly differently. Invented words are also included for each new letter that is introduced. Invented words give pupils the opportunity to practice phonics and to improve pupil's ability to recognise the most common sounds for letters.

أَنَا لَدَي تُوْت لَذِيذُ.
 أَنَا لَدَي لَبَن لَذِيذُ.
 أَنَا لَدَي لُبَان لَذِيذُ.
 أَنَا وَ أَنْتِ لَدَيْنَا تُوْت

Same word –
 slightly different
 font size and
 spacing

WHEN

The intended delivery of LRF! is across 14 weeks, 12 weeks of core material and 2 weeks of revision sessions. Only the 12 weeks of core material were delivered for the pilot study, as the EGRA tests reduced the number of weeks available in each semester. This delivery model applied to both the C/U and W/C approaches.

For semester one, delivery started on 3 October 2021 and was due to run until 23 December 2021. However, due to COVID-19, there was a national closure of schools from early December 2021 to February 2022. This meant that the final weeks of the intervention were delivered in February and March 2022. In semester two, the intervention was delivered from 6 March 2022 until 2 June 2022, which includes a week-long break for Eid. This was later than the original intended delivery from 13 February to 19 May 2022 due to the national closure of schools in winter 2021/22.

TAILORING

Teachers had some flexibility over how they facilitated the sessions; however, the content of each session was set. Teachers were asked to allow for 15-20 minutes of individual independent practice after each lesson.

Teachers were encouraged to draw on their professional judgement about tailoring instruction according to pupils' needs. Teachers were expected to ensure they adequately progress through the content of the practice book, while at the same time ensuring pupils are able to adequately master each 'lesson' as they do so.

CONTROL CONDITION

Pupils in control schools received teaching as usual. Intervention schools were not offered incentives to participate. However, control schools received stand boards as an incentive.

Prior research

In the summer of 2021 a pre-pilot of both the W/C and C/U models was conducted over an 8-week period in the summer.¹⁰ The pre-pilot was carried out to inform the design of the

¹⁰ The original pre-pilot intervention launched in March 2021 but paused after less than two weeks due to COVID related schools' closures. Subsequently, the pre-pilot relaunched in two community centers in summer 2021. Pupils who participated in the pre-pilot had either finished Grades 1 or 2.

pilot by considering how the intervention was implemented, and whether any components of the intervention should be adjusted. This included the practice book, the structure of the teacher and pupil interaction, the teacher training and the approaches to delivering C/U and W/C. To do this, the pre-pilot evaluation gathered contextual information, teacher feedback, classroom observation and fidelity of implementation (FOI) data.

The pre-pilot results suggested a set of adaptations for the pilot. Table 1 below summarises the changes that were made. No adaptations were made to coaching because the delivery of LRF! is perceived to be expensive, so despite recommendations to increase the number of coaching sessions, no increase was made in order to keep delivery cost low. No adaptations were also made to diagnostic tool and teacher rewarding achievement.

Table 1: LRF! – adaptations from the pre-pilot

Adaptations made to:	
Practice Book	Classroom instruction
Recorded the lessons for the purpose of supporting teachers with an accurate model of reading in preparation for their lessons and supporting parents and/or pupils when they practice at home.	Encouraged teachers to draw on their professional judgement and knowledge of their pupils with regard to the flow of the routine of (I do, we do, you do) i.e. do 1-2 pages in one cycle and the rest in another cycle, with careful attention that the independent reading time is not less than 15 minutes in total. The purpose of this change is increasing pupils' attention span and engagement.
Reviewed invented words to ensure they follow word phonetic rules in Arabic.	Emphasised starting where pupils are standing, explicitly introducing the new letter, diacritics if they feel the need.
Reviewed language accuracy.	Emphasised the importance of achieving mastery before proceeding to the next lesson.
Added a self-tracking tool in the footnote of each page. This is a question asking the pupil: How many times did you read this page?	Emphasised techniques that motivate pupils to increase their independent reading stamina. i.e. increasingly challenging them to add to the minutes they are on task, asking them "how much did you exercise your brain today?".
Added another tool to track practice at home. For each lesson, pupils can colour a figure of a brain carrying weights to reflect their answer to the question ("How much did you exercise your brain today?")	

Logic model

A logic model was developed for each strand of the intervention (C/U and W/C) in advance of the evaluation. Both LRF! logic models were updated by the evaluation team in collaboration with QRF (see Figure 1 and Figure 2 below). The updating process included a logic model workshop on 27 July 2021 attended by key members of the QRF and evaluation teams.

Figure 1: LRF! logic model – Whole Class

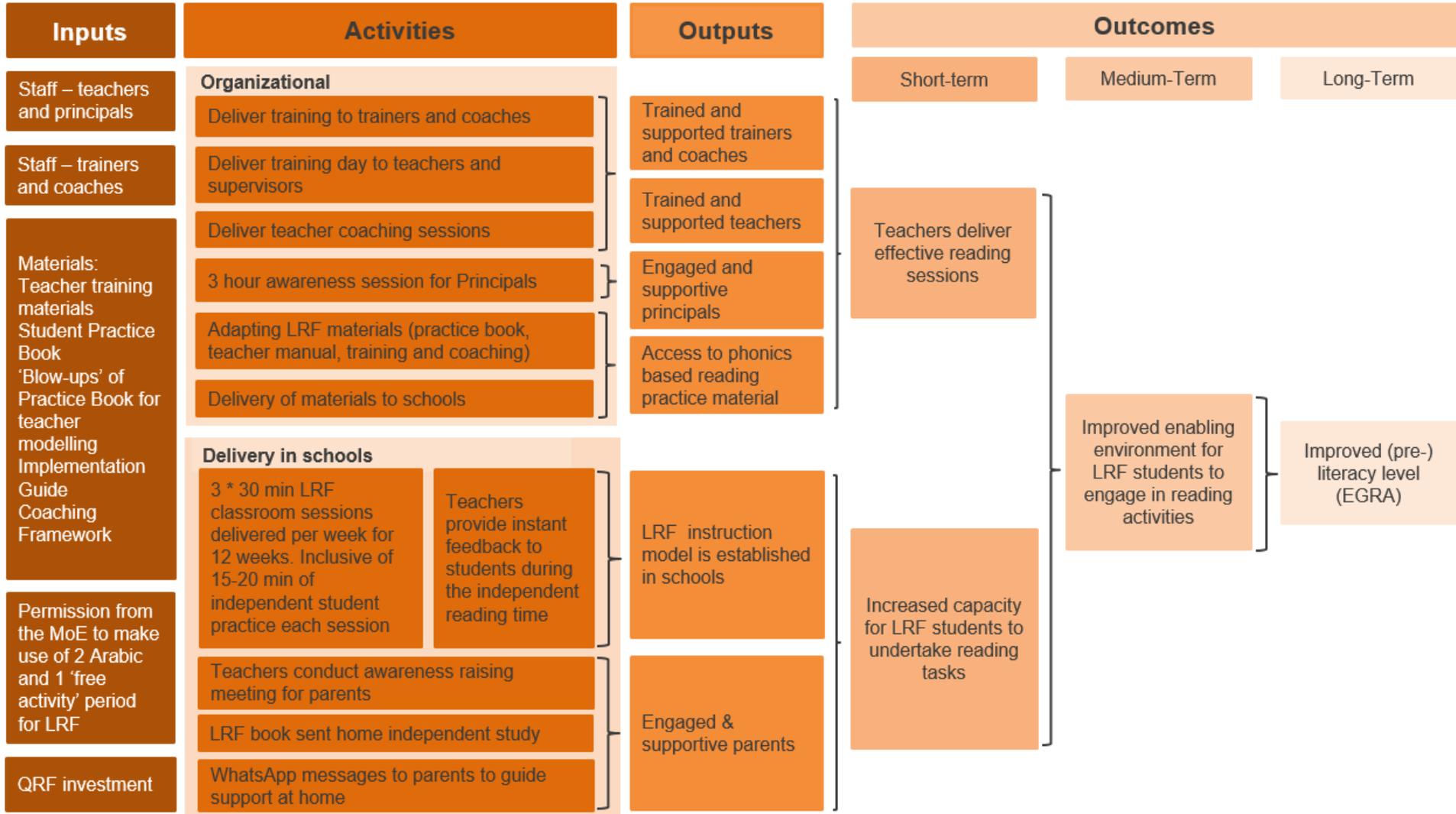
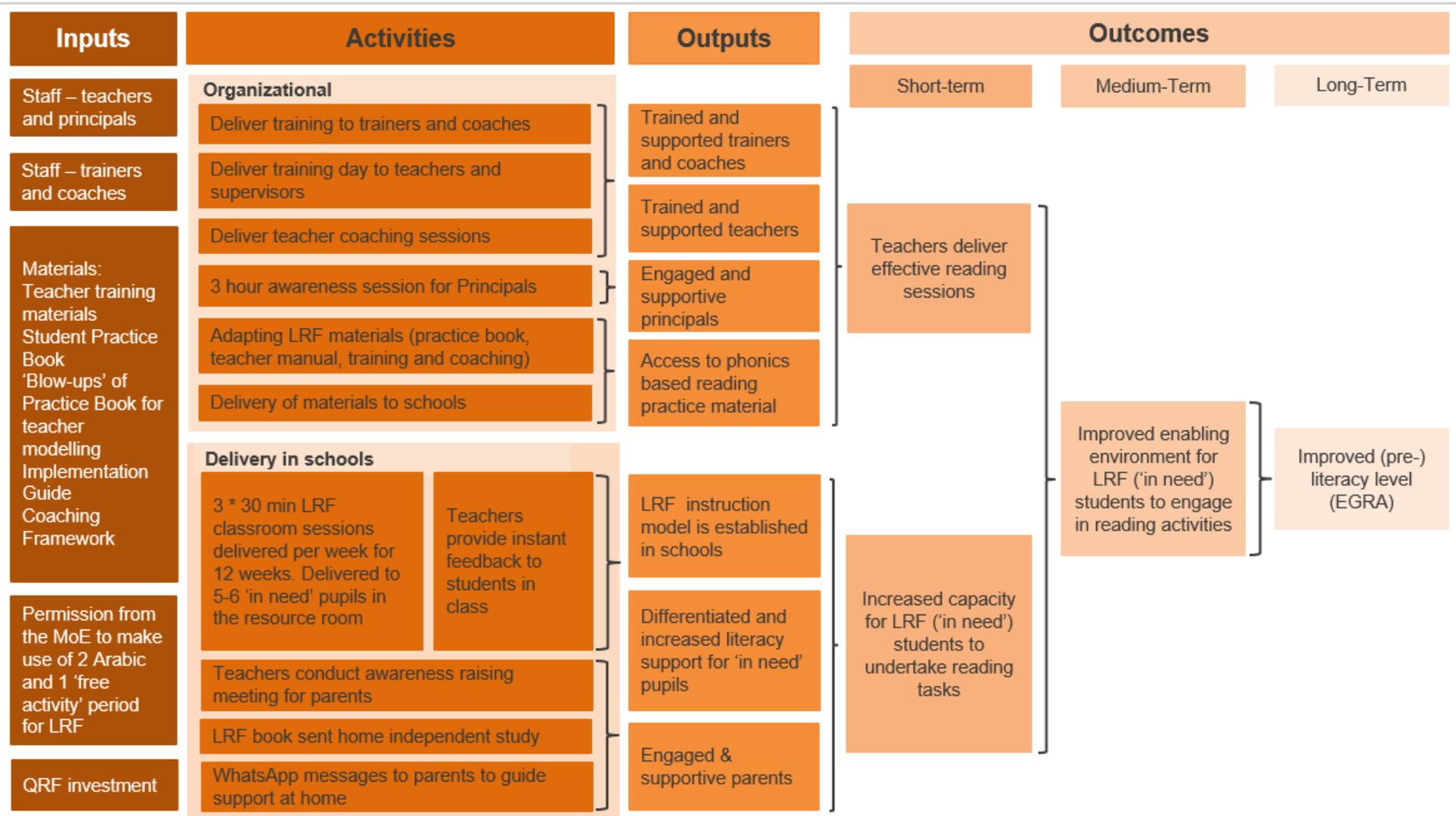


Figure 2: LRF! logic model – Catch Up



RESEARCH QUESTIONS

The pilot aims to answer the following research questions (RQs) within the following domains:

Evidence of promise

Evidence of promise explores early indication of outcomes. These include measured outcomes as part of the impact evaluation and perceived outcomes in the implementation and process evaluation (IPE).

1. In what ways, and to what extent, does LRF! affect school, teacher, pupil, and parental practice as compared to business-as-usual teaching and learning?
2. How do principals, teachers, parents and pupils perceive the intervention and any changes that it has delivered?
3. Is there evidence to support the logic models?
4. Is there any evidence of unintended consequences (negative or positive) as a result of the implementation of LRF!?

Feasibility of intervention

The feasibility of the intervention questions assess the practical delivery of the intervention.

5. Was LRF! delivered as intended in terms of dosage, nature and quality? What modifications were made, with what implications?
6. What is the learning about teacher's use of the Coarse-Grained Diagnostic (RAMP) tool? How successful is it at identifying the most appropriate pupils for the C/U model?
7. What were the facilitators and barriers to engagement in the teacher training, teacher coaching and supervisor and principal orientation sessions?
8. To what extent do teachers develop sufficient skills and confidence through the training and coaching?
9. What do we know about how teachers need to be supported (coached) during delivery?
10. Are there any key contextual factors that appear to facilitate or impede successful implementation of LRF!?

Assessing feasibility of the efficacy trial(s)

The feasibility of the efficacy trial(s) questions analyse the pilot study to see the viability of an efficacy trial.

11. What does the pilot tell us about the feasibility of the process components of an efficacy trial, e.g., school recruitment, retention, or data collection in both intervention and control groups?
12. What does the pilot tell us about the feasibility of the resources of an efficacy trial, e.g. measurement instruments or specific equipment used?

13. What does the pilot tell us about the feasibility of the management components of an efficacy trial, e.g. problems with data collection or variability of collected data?

Assessing readiness for trial

The assessing readiness for trial questions use the other three domains to assess whether the intervention and pilot study are able to proceed to efficacy trial.

14. What changes, if any, are needed to the logic models?
15. What changes to the intervention, implementation models, support or materials need to be made?
16. What can we learn from the pilot about minimal detectable effect size estimates, intra-cluster correlations, pre-and-post correlations and sample sizes?
17. Is there any evidence of contamination between the control and treatment groups? For example, from the supervisors who attend the training alongside teachers in their district spreading, or promoting, aspects of LRF! to other schools under their supervision.

To answer the outlined RQs the evaluation triangulated across different data sources. Table 2 maps the data collection activities by RQ.



Table 2: Research activities against evaluation domains and research questions

Research Question	Research activity							
	EGRA tests	Focus Group Discussions (FGDs) with teachers	FGDs with parents	FGDs with pupils	FGDs with coaches	Classroom observations	Teacher survey	Pupil survey
EVIDENCE OF PROMISE								
1. In what ways, and to what extent, does LRF! affect school, teacher, pupil, and parental practice as compared to business-as-usual teaching and learning?		◆	◆	◆	◆		◆	◆
2. How do principals, teachers, parents and pupils perceive the intervention and any changes that it has delivered?		◆	◆	◆	◆		◆	
3. Is there evidence to support the logic models?	◆	◆	◆	◆		◆	◆	◆
4. Is there any evidence of unintended consequences (negative or positive) as a result of the implementation of LRF!?		◆	◆	◆	◆		◆	◆
FEASIBILITY OF INTERVENTION								
5. Was LRF! delivered as intended in terms of dosage, nature and quality? What modifications were made, with what implications?		◆			◆	◆	◆	
6. What is the learning about teacher's use of the Coarse-Grained Diagnostic (RAMP) tool? How successful is it, in use, at identifying the most appropriate pupils for the Literacy Catch-Up implementation model (C/U model only)?		◆			◆			
7. What were the facilitators and barriers to engagement in the teacher training, teacher coaching and supervisor and principal orientation sessions?		◆			◆		◆	

Research Question	Research activity							
	EGRA tests	Focus Group Discussions (FGDs) with teachers	FGDs with parents	FGDs with pupils	FGDs with coaches	Classroom observations	Teacher survey	Pupil survey
8. To what extent do teachers develop sufficient skills and confidence through the training and coaching?		◆			◆	◆	◆	
9. What do we know about how teachers need to be supported (coached) during delivery?		◆			◆		◆	
10. Are there any key contextual factors that appear to facilitate or impede successful implementation of LRF!?		◆	◆		◆	◆	◆	◆
FEASIBILITY OF TRIAL								
11. What does the pilot tell us about the feasibility of the process components of an efficacy trial, e.g., school recruitment, retention, or data collection in both intervention and control groups?	◆	◆	◆	◆	◆	◆	◆	◆
12. What does the pilot tell us about the feasibility of the resources of an efficacy trial, e.g. measurement instruments or specific equipment used?	◆	◆	◆	◆	◆	◆	◆	◆
13. What does the pilot tell us about the feasibility of the management components of an efficacy trial, e.g. problems with data collection or variability of collected data?	◆	◆	◆	◆	◆	◆	◆	◆
READINESS FOR TRIAL								
14. What changes, if any, are needed to the logic models?		◆	◆	◆	◆		◆	◆
15. What changes to the intervention, implementation models, support or materials need to be made?		◆	◆	◆	◆	◆	◆	◆
16. What can we learn from the pilot about minimal detectable effect size estimates, intra-cluster	◆							

Research Question	Research activity							
	EGRA tests	Focus Group Discussions (FGDs) with teachers	FGDs with parents	FGDs with pupils	FGDs with coaches	Classroom observations	Teacher survey	Pupil survey
correlations, pre-and-post correlations and sample sizes?								
17. Is there any evidence of contamination between the control and treatment groups?					♦	♦		

PROJECT TEAMS

The evaluation team included staff from NatCen who led the partnership, as well as Integrated, Oxford MeasurEd (OM) and School-to-School International (STS). **NatCen** was the lead partner and accountable to QRF. NatCen led on project coordination and management, evaluation design, analysis and reporting. **Integrated** was a subcontracted partner. They led EGRA testing, IPE data collection and contributed to the evaluation design, analysis and reporting. **OM** was a subcontracted partner. They developed and refined the primary outcome measure and contributed to the evaluation design and analysis. OM was responsible for developing a technical report about the design of the primary outcome measure. **STS** was a subcontracted partner. They brought expertise in child literacy and acted as a ‘critical friend’ throughout the evaluation, contributing to the evaluation design, analysis and reporting stages.

The project delivery team included staff from QRF, QRTA and Advisem. **QRF** had overall responsibility for delivery and provided inputs on the design of the evaluation. **QRTA** was responsible for implementation of the intervention, including training and coaching sessions. The team at **Advisem** worked closely with QRF and provided advice on the design of the evaluation.

Staff members of the evaluation and project delivery teams are presented in Table 3 and Table 4.

Table 3: Evaluation project team

Name	Project role	Role and team
Sashka Dimova	Principal Investigator and strategic lead	Research Director, Evaluation, NatCen
Hannah Woodbridge	Day-to-day project manager	Senior Researcher, Children and Families, NatCen
Julia Ruddick-Trentmann	Project management support	Researcher, Children and Families, NatCen
Andi Fugard	Impact Evaluation Lead	Deputy Director, Evaluation, NatCen
Enes Duysak	Impact Evaluation Support	Senior Researcher, Evaluation, NatCen
Tien-Li Kuo	Impact Evaluation Support	Researcher, Evaluation, NatCen
Eliza Garwood	IPE support	Senior Researcher, Children and Families, NatCen
Natasha Phillips	IPE support	Researcher, Children and Families, NatCen
Nedjma Koval	Data collection lead	Integrated
Leen Al Refai	IPE Lead, project management	Integrated
Zeid Qiblawi	EGRA, IPE Data Collection Support	Integrated
Rachel Outhred	Outcome measure development lead	OM
Lydia Marshall	Support for outcome measure development, evaluation design, analysis and reporting	OM
Carol da Silva	Provide feedback to theory of change	STS

Daniel Phillips	Quality assurance	Director, Evaluation, NatCen
-----------------	-------------------	------------------------------

Table 4 Delivery project team

Name	Project role	Role and team
Lubna Dirini	Manager, Lead trainer	QRTA
Amani Alker	Coach	QRTA
Mohammad Salameh	Coach	QRTA
Maysoon Masoud	Project lead	QRF, Research & Program Development Manager
Rami Al Assad	Administrative & financial management	QRF, Project Management Specialist
Emilee Rauschenberger	Quality Assurance	QRF, Research Manager
Haneen Alabed	Material development lead and trainer of trainers	QRF, Research & Program Development associate Manager
Hanif Pabani	Quality Assurance	Advisem, Impact Evaluation consultant
Julie Helson	Quality Assurance	Advisem, Impact Evaluation consultant
Louis-Pierre Michaud	Quality Assurance	Advisem, Impact Evaluation consultant



DATA PROTECTION, ETHICS, AND TRIAL REGISTRATION

Data Protection

We recognise the need for data security and operate to extremely high standards of confidentiality and anonymity.

NatCen is fully accredited to ISO 27001 and subject to annual external audits of procedures to maintain accreditation. We also hold Cyber Essentials Plus Certification. We were previously registered under the Data Protection Act and are now fully GDPR compliant.

EGRA assessments were undertaken by Integrated, with pupils assigned a unique identifier. Test results were submitted to EGRA's Prodigy tool and a pseudonymised dataset transferred to NatCen and OM. EGRA data was stored with back-end provider Prodigy and sent directly to authorised Integrated personnel. Integrated stored the IPE data on a dedicated drive that could only be accessed by authorised personnel.

All information stored, processed and/or transmitted by Integrated is protected in a manner consistent with contractual and legal restrictions proportionate to the level of sensitivity, value and risk of that information to Integrated, its partners and/or clients. Sensitive information is secured against disclosure, modification, and access by unauthorised individuals while both holding and transferring it. Personnel with authorised access are obliged to maintain data confidentiality through measures such as legally binding provisions in employment contracts, as well as a signed code of conduct for all employees.

Data shared with NatCen is stored on NatCen's secure network, with access to the project folder restricted to authorised personnel only. The data is backed up, and NatCen carried out regular testing to ensure this process was effective.

To ensure integrity and confidentiality, all data and files held by NatCen are classified to one of three different levels, with each level having its own specific requirements for how the data are stored, handled, and transmitted. Any data containing personal details is deemed to be 'Respondent Confidential'. For such data, protection against the disclosure of respondent identities – whether by direct association with a name or address or by indirectly associating information disclosed – is built into all stages of the process.

OM had access to EGRA data to conduct analyses. NatCen transferred EGRA data to OM using the NatCen's Secure File Transfer Services. The data was stored on OM's secure servers and only named team members have access to this data.

The partnership agreed a data retention period with QRF. Once this period has expired all partners will securely erase project data (with explicit permission from QRF).

Ethical Approval

This project was submitted to NatCen's Research Ethics Committee (REC), made up of senior NatCen staff and external experts where appropriate, for scrutiny in advance of data collection. NatCen's ethics procedure meets the requirements of the UK Economic and Social Research Council (ESRC) and the UK Government Social Research (GSR) Professional Guidance. The evaluation was undertaken according to NatCen procedures designed to ensure our research is conducted in line with five principles outlined by the Government Social Research (GSR) guidance:



- Sound application and conduct of social research methods, and appropriate dissemination and utilisation of the findings.
- Participation based on valid informed consent.
- Enabling participation.
- Avoidance of personal and social harm.
- Ensuring that participants are not identifiable in the outputs.

Trial registration

The trial was registered on the Open Science Foundation on 26/11/2021 (<https://osf.io/vj7a4>).

METHODS

Pilot design

The pilot evaluation was designed as a three-arm cluster randomised controlled trial (RCT). In total 24 schools took part in the pilot evaluation. Schools were stratified by region and urban/rural classification prior to randomisation to ensure balance across pilot arms across strata after randomisation. Randomisation occurred as planned, and all eligible schools had an equal chance of being assigned to the control group, (W/C) model or (C/U) model.

The primary outcome was Arabic literacy attainment among Grade 1, Grade 2, and Grade 3 pupils in Jordan. The secondary outcomes were the specific sub-domains of Arabic literacy attainment for the same grades. The same primary and secondary outcomes were measured at baseline and endline using the same tools. Grade 1 pupils in schools assigned to the W/C intervention and Grades 1, 2 and 3 pupils in schools assigned to the C/U model were eligible to participate in the intervention. Fifty percent of pupils in Grade 1 in W/C schools were tested in Arabic literacy attainment, whereas 50% of Grade 1, 2 and 3 pupils in C/U schools who in the lowest achieving 20% of their class were tested in Arabic literacy attainment.

Table 5 presents the pilot design, including the unit of randomisation, stratification, variables and measures of outcomes.

Table 5: Study design – pilot

Trial design, including number of arms		Pilot Evaluation
Unit of randomisation		School level
Stratification variables		Region and urban/rural classification
Primary outcome	Variable	Arabic literacy attainment
	Measure (instrument, scale, source)	<p>Source: EGRA Grade 2 assessment with the addition of a set of pre-literacy items</p> <p>Instrument: EGRA + pre-literacy tool</p> <p>Scale: scale scores with a mean of 500 and a standard deviation of 100.</p>
Secondary outcome(s)	Variable(s)	Specific sub-domains of Arabic literacy attainment



	Measure(s) (instrument, scale, source)	Source: EGRA Grade 2 assessment with the addition of a set of preliteracy items Instrument: EGRA + pre-literacy tool Sub-domains: Oral reading fluency, letter sound identification, syllable identification, and reading comprehension
Baseline for primary outcome	Variable	Arabic literacy attainment
	Measure (instrument, scale, source)	Source: EGRA Grade 2 assessment with the addition of a set of preliteracy items Instrument: EGRA + pre-literacy tool Scale: scale scores with a mean of 500 and a standard deviation of 100.
Baseline for secondary outcome	Variable(s)	Specific sub-domains of Arabic literacy attainment
	Measure (instrument, scale, source)	Source: EGRA Grade 2 assessment with the addition of a set of preliteracy items Instrument: EGRA + pre-literacy tool Sub-domains: Oral reading fluency, letter sound identification, syllable identification, and reading comprehension

Randomisation

Every school recruited for the pilot was randomly allocated to one of the pilot arms (W/C, C/U or control). Schools were randomly allocated to groups at the beginning of the 2021/22 academic year.

Schools were stratified by region and urban/rural classification prior to randomisation to ensure balance across pilot arms across strata after randomisation. Jordan has three geographical regions (middle, south, and north). To equally represent schools in the middle region, and due to logistic reasons, we divided these regions into four geographical regions (Amman, middle excluding Amman, near south and near north¹¹). Given that we had four geographical regions and a school could be either in a rural or urban area, we had eight strata.¹²

Randomisation was carried out by an analyst at NatCen in September 2021. Randomisation was undertaken in Stata 16 and both the 'do' and 'log' files were saved as a record of the randomisation process.

Recruitment

All primary schools in Jordan were eligible for the pilot as long as they satisfied the following conditions:

- having pupils in Grades 1, 2 and 3;
- being a single shift school;

¹¹ See LRF! pilot protocol for details on how we produced geographical regions (<https://www.qrf.org/sites/default/files/2022-05/Let%27s%20Read%20Fluently%20Pilot%20Protocol.pdf>).

¹² We had the following strata: Amman urban, Amman rural, Middle excluding Amman urban, Middle excluding Amman rural, North urban, North rural, South urban and South rural.



- being a school from middle, near north and near south;
- not being part of any other literacy interventions, other than the Reading and Writing Project which has been implemented in Grades 1-3 in all the Ministry of Education (MoE) school in Jordan since 2011;
- having a resource room teacher;
- not delivering blended teaching¹³;
- not being in a Syrian refugee camp.

QRF was responsible for recruiting schools satisfying the eligibility criteria listed above in September 2021. They also liaised with the MoE, who provided written permission for the intervention and evaluation to take place, as well as for the use of the EGRA assessment. QRF explained to prospective schools what participation in the evaluation would involve during the recruitment process.

In total, QRF recruited 24 primary schools: 8 primary schools in the W/C intervention, 8 in the C/U intervention, and 8 in the (shared) control group (completely randomised).

IMPACT EVALUATION

Outcome measures

As shown in Table 5, the primary outcome was Arabic literacy attainment whereas the secondary outcomes included letter sound identification, syllable identification, oral reading fluency and reading comprehension as sub-domains. The pilot study used an assessment of Arabic literacy that was well-targeted to the proficiency levels of learners in the LRF! intervention. This was determined through a robust process (outlined in a technical report produced by OM, see Outhred et al., 2022) including a desk review of previous administrations of EGRA and a pre-pilot. The assessment includes items that span pre-literacy through to reading fluency proficiency levels. This was necessary to adequately measure the full range of proficiency levels across the sample at baseline and capture higher proficiency levels at endline. We refer to the tool as EGRA+pre-lit, as it combines some EGRA subtasks previously administered in Jordan and a set of new pre-literacy items developed specifically for this study. Baseline primary and secondary measures were identical to the endline primary and secondary measures. We collected baseline data for Grade 2 and Grade 3 both C/U and their control group in September 2021, while endline data collection for these groups took place in March 2022. Grade 1 W/C and their control group had baseline data collected in February 2022 and endline data collection took place in June 2022. Baseline data collection for Grade 1 C/U and their control took place in March 2022, while their endline data was collected in June 2022.

The primary outcome measure was a single learning metric for reading, which was obtained by administering the EGRA+pre-lit. We used Item Response Theory (IRT) to produce an interval-scaled measure that took the difficulty level of individual items into account¹⁴ so that an underlying construct (i.e. one that cannot be measured directly) becomes one that can be measured on an interval scale where a higher score indicates higher levels of proficiency and lower scores indicate lower levels of proficiency. The latent unidimensional traits that we sought to measure using the pupil tests were abilities/performance in Arabic literacy. We used the Rasch

¹³ Due to the COVID related safety measures, in many large public schools pupils attended schools on alternative day shifts. Schools eligible for the study were only those in which students attended school on daily basis.

¹⁴ More details on IRT can be found in the [LRF! pilot protocol](#) and in a separate technical report produced by OM.



model, a special case of an IRT model, to estimate the probability of answering the item correctly as a logistic function of the difference between the person’s ability and the item’s difficulty.¹⁵ The formula and details on how the Rasch model worked for dichotomous responses can be found in a technical report produced by OM (Outhred et al., 2022). Briefly, all administered items (see Appendix C) were included in the IRT model, which produced a scale score per participant. Therefore, the overall Arabic literacy attainment score calculated in the IRT model was rescaled to have a mean of 500 and a standard deviation of 100.

The secondary outcomes were the specific sub-domains of Arabic literacy attainment (i.e. oral reading fluency, letter sound identification, syllable identification, and reading comprehension). We also produced separate metrics for these sub-domains from the EGRA.

The first sub-domain in this trial was oral reading fluency (ORF). ORF measures pupil’s ability to orally read text. As part of EGRA testing, pupils were given a short story and asked to read it within one minute. The story consisted of 42 words. The ORF was measured as the number of correct words read per minute. The second sub-domain of EGRA was letter sound identification (LSI). It assessed pupil’s ability to associate sounds with letters. In this task, pupils were given 100 letters and asked to read the letter sounds within a minute. A per-minute score for LSI was created to assess pupil’s ability on LSI. Pupils also completed a sub-domain of EGRA used to assess pupil’s ability to identify syllables. In this sub-domain, pupils were given 100 Arabic syllables to read within one minute. The number of correct syllables was used to create a per-minute score, which reflects pupil’s ability to identify syllables. The last sub-task of EGRA implemented in this pilot trial was for reading comprehension (RC). Once a pupil completed reading the short story for the ORF measure, pupils were asked five questions related to the short story to assess their reading comprehension. They were given a maximum of 15 seconds for each question. The RC score was the number of correct answers, out of a maximum possible score of 5.

Impact evaluation analysis

Primary outcome pilot analysis

In line with the EEF analysis guidance,¹⁶ the primary outcome analysis followed an intention-to-treat (ITT) approach.¹⁷ The analyses used a three-level multilevel model to account for the clustering of pupils (level 1) in classes (level 2) and schools (level 3). The treatment assignment was at the school level (level 3). This model included school and class-level random effects and accounted for the baseline EGRA+pre-lit.

A separate model was estimated for each intervention type (the W/C approach and the C/U approach). The basic form of the model is:

$$Outcome_{ijk} = \beta_0 + \beta_1 Baseline_{ijk} + \beta_2 Intervention_k + \beta_3 Strata_k + u_{jk} + e_{ijk}$$

¹⁵ Test scores measure two things. Firstly, they measure the proficiency of the test taker. Secondly, they measure the difficulty of the test. The logistic function of the Rasch model places the difficulty of the item and the proficiency of the test taker on the same scale. This involves ranking the difficulty of the items and the ranking the test takers. In our item person map, where a test taker and an item are placed at the same point in a scale, that test taker has a 50 percent probability of answering that item (or an item of equal difficulty) correct.

¹⁶ Our analysis approach is based on EEF (2018) Statistical analysis guidance for EEF evaluations.

¹⁷ The Intention-to-Treat approach compares pupils assigned to the treatment and pupils assigned to the control group, irrespective of whether pupils assigned to the treatment group actually receive the intervention.



Where pupils (i) were clustered in classes (j) within schools (k). The intervention effect was estimated by β_2 , β_3 represented strata fixed effects for the schools (i.e., their geographical location), β_0 was the intercept, β_1 was the slope for baseline scores, u_{jk} a classroom-level random intercept, and e_{ijk} the residual term. In line with the EEF analysis guidance, other additional covariates were not considered. The analysis was carried out using Stata 17 and both the syntax used and outputs of analysis were saved as a record of process.

The difference between the intervention and control groups at endline was expressed as a standardised effect size using Hedges' g with 95% confidence intervals. Following EEF guidelines, the numerator came from the unstandardised effect estimate given by β_2 in the multilevel model specified above, which was adjusted for baseline score and strata. The denominator was the unconditional pooled standard deviation in the primary outcome at endline. The formula is provided below:

$$g = \frac{\beta_2}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}}$$

Where n_1 and n_2 were the number of pupils in both groups and s_1^2 and s_2^2 were the within-group variances in outcomes at endline.

The primary analysis also included detailed descriptive analysis: histograms, means, quartiles, and SDs, for all measures, groups, and time points.

We also reported school and class-level Intracluster Correlation Coefficients (ICCs¹⁸) alongside 95% confidence intervals in analyses. We used the following model with a random intercept by school, v_k , a classroom-level random intercept, u_{jk} , the residual term, e_{ijk} , and no fixed effects for strata so it was arithmetically possible that the variance of $v_k > 0$:

$$Outcome_{ijk} = \beta_0 + u_{jk} + v_k + e_{ijk}$$

Secondary outcome pilot analyses

The secondary outcome analysis involved first providing summary statistics and an unadjusted mean difference between the intervention and control group for secondary outcome measures.

For all defined secondary outcomes (i.e. oral reading fluency, letter sound identification, syllable identification, and reading comprehension), we followed the ITT approach using a basic model for each intervention type (the W/C approach and the C/U approach) similar to that of the primary outcome analysis. This statistical model took the form of multilevel model, where pupils were clustered in classes within schools, and accounted for stratification factors used at randomisation.

$$Outcome_{ijk} = \beta_0 + \beta_1 Baseline_{ijk} + \beta_2 Intervention_k + \beta_3 Strata_k + u_{jk} + e_{ijk}$$

¹⁸ The ICC measured similarity between units in the same cluster; in this case, pupils within the same classroom. Units within the same cluster may exhibit similarities due to being exposed to similar environmental characteristics. This must be accounted for when conducting sample size calculations, since similarity between units reduced the amount of unique information each new observation contributed to the sample.



Where pupils (i) were clustered in classes (j) within schools (k). The intervention effect was estimated by β_2 , β_3 represented strata fixed effects for the schools (i.e., their geographical location), u_{jk} a classroom-level random intercept, and e_{ijk} the residual term.

For all secondary outcome measures we also used standardised scores (z-scores), reporting confidence intervals at 95% level, and the effect size using Hedges' formula previously described.

IMPLEMENTATION AND PROCESS EVALUATION

The IPE used several different data collection tools. This enabled data to be collected across the IPE domains of interest, which were informed by EEF's framework for IPE (EEF, 2019).

Unless otherwise stated, all data collection was administered by Integrated. Integrated have a team of trained enumerators who are experienced in collecting data within Jordanian schools.

Training attendance data

The Queen Rania Teacher Academy (QRTA) gathered attendance data for each LRF! training session. This data recorded the names of those invited to attend training, the names of those who attended, and their roles within schools. For example, whether they were a teacher or supervisor, and what grade(s) they taught. This information was recorded in Excel and analysed to produce numerical counts of the number of training attendees and absences.

Classroom observations

Classroom observations were conducted in all sections implementing either the W/C or C/U intervention. These observations assessed pupil engagement with LRF!, pupil use of the practice book, and the administration of the 'I do', 'We do', 'You do' approach. All sections were observed twice, once a few weeks after initially implementing the intervention and once again shortly before the intervention finished. Some schools in the second semester (Grade 1 C/U and W/C) received additional follow up observations due to teacher and pupil absence. The purpose of the two rounds of observation was to assess any changes across the semester. The data was collected using a Fidelity of Implementation tool and a Pupil Engagement tool (see Appendix C).

Teacher Survey

The teacher survey covered experiences of LRF! training and coaching, engagement with parents, and perceptions of LRF!. For the W/C model, all classroom teachers completed a survey at the same time as their second observation. This was because it was believed that teachers had enough experience to reflect on LRF! and it reduced participant burden by combining the two elements of fieldwork.

Data was collected using the Alchemer survey software and results were shared in Excel.

Pupil Survey

The pupil survey was comprised of multiple questions assessing engagement and interest in reading, access and usage of the internet, and for intervention pupils at endline, reflections on LRF!. All pupils participating in the evaluation, including those in the control arm, completed the pupil survey. For semester two, the survey was conducted alongside the EGRA+pre-lit at both baseline and endline. This enabled assessment of change across the intervention period. For semester one, the pupil survey was not delivered at baseline, as the material was still being developed, and was only delivered at endline. Due to this delay in developing the baseline survey,



reported C/U baseline data for both intervention and control contain only Grade 1 pupils. The C/U endline data for both intervention and control contains the Grade 1 pupils and those pupils in Grades 2 and 3. Subsequently, the endline sample is larger compared to the baseline and any change in results between the two surveys may be attributable to this variation in the sample.

The pupil survey data was collected using the Alchemer survey software and results were shared in Excel (see Appendix C).

Focus groups discussions with teachers, pupils, parents, and coaches

FGDs took place with several stakeholders across all three arms of the intervention. Table 6 below summarises these FGDs.

Table 6: FGD summary

Stakeholder	Control	Whole class	Catch Up	Total
Teachers	2x Grade 2/3 classroom and resource room teachers (June 2022)	1x classroom teachers (June 2022)	2x Grades 2/3 resource room teachers (December 2021)	10 FGDs
	1x Grade 1 classroom and resource room teachers (June 2022)		1x Grade 1 resource room teachers (June 2022)	
			3x Grades 1, 2, and 3 classroom teachers (June 2022)	
Pupils	N/A	1x pupil (June 2022)	1x Grades 2/3 pupils (December 2021)	3 FGDs
			1x Grade 1 pupils (June 2022)	
Parents	1x Grades 1, 2 and 3 parents (June 2022)	4x parents (June 2022)	2x Grades 2/3 parents (December 2021)	8 FGDs
			1x Grade 1 parents (June 2022)	
Coaches	N/A	1x Coaches (June 2022)	1x Grades 2/3 coaches (January 2022)	3 FGDs
			1x Grade 1 coaches (June 2022)	
Total	4 FGDs	8 FGDs	12 FGDs	24 FGDs

Schools were sampled to ensure that FGDs took place across the geographical regions and with a mix of urban and rural settings. For C/U resource room teacher sessions, all teachers were invited to attend due to the small population size (8 resource room teachers total). School principals helped to recruit parents by sending an open invitation to all parents of pupils participating in LRF!. At all FGDs, verbal consent was given by participants at the start of the session, including consent to record the FGD. Participants were also reminded of their right to withdraw from the research.

FGDs generally took place within schools, although some teacher FGDs were conducted virtually or at Integrated's head office in Amman. This was to increase attendance, as some teachers were not willing or able to travel across the country. Additionally, only 19 FGDs were originally included in the research design, however due to poor attendance amongst parents and teachers, five additional FGDs were required. Difficulty in travelling and a lack of reimbursement were seen as the main challenges to attendance.



The FGDs covered a range of content. In the case of the control arm, the FGDs aimed to get a better understanding of business-as-usual literacy learning. For both W/C and C/U, the FGDs aimed to capture experiences and perceptions of LRF!. Table 7 below summarises the length and content of the sessions.

Table 7: FGD content

Stakeholder	Aim of FGD	FGD length
W/C classroom teacher	<ul style="list-style-type: none"> Teachers' understanding and perceptions of the intervention Teachers' experiences of implementation to date, including barriers and enablers to delivery Teachers' insights into perceived impacts to date 	45 – 60 minutes
C/U resource room teacher		
C/U classroom teacher	<ul style="list-style-type: none"> Teachers' understanding and perceptions of the intervention Teachers' experiences of implementing the diagnostic tool and engaging with pupils and parents Teachers' insights into perceived impacts to date 	25 – 30 minutes
Control classroom and resource room teacher	<ul style="list-style-type: none"> Teachers' typical approach to teaching children to read Teachers' experiences of barriers and enablers to teaching reading 	45 minutes
Pupils (same topic guide for both W/C and C/U)	<ul style="list-style-type: none"> Pupils' perceptions of LRF! lessons Pupils' usage of the practice book and their perceptions of it Whether they feel they are making progress with reading or not, and why they think this 	40 minutes
W/C parents	<ul style="list-style-type: none"> Parents' understanding and perceptions of the intervention Parents' usage and perceptions of the practice book Parents' insights into perceived impacts to date 	45 – 60 minutes
C/U parents		
Control parents	<ul style="list-style-type: none"> Parents' typical approach to supporting children to learn to read Parents' perceptions of enablers and barriers children face when learning to read Parents' insights for improvement in reading instruction/learning 	50 minutes
Coaches (same topic guide for both W/C and C/U)	<ul style="list-style-type: none"> Coaches' perceptions of the intervention Coaches' experiences of implementation to date, including barriers and enablers to delivery Coaches' insights into perceived impacts to date 	60 minutes

For all FGDs that took place in semester one, notes were taken summarising the discussion. These were translated into English and shared with NatCen for analysis. In semester two, all FGDs were audio recorded and transcribed. These transcripts and audio files, along with notes in English, were then shared with NatCen for translation and analysis. NatCen used an approved external



provider to translate the Arabic transcripts into English. As described in the study [protocol](#), the original research design planned to use the NatCen's framework approach¹⁹ to conduct thematic analysis. However, due to translation taking 4 weeks to complete and the limited timeframe for analysis, NatCen opted for a quicker, more efficient approach. This approach involved highlighting key themes within each FGD to organise data from each FGD, and then it grouped the emerging themes into findings.

MODIFICATIONS TO THE EVALUATION DESIGN

Additional analysis

Subgroup analysis

The protocol did not set out a subgroup analysis. However, as part of the further investigation of our results, we completed a descriptive subgroup analysis. Three categories of pupils' overall abilities in Arabic literacy were created based on the baseline EGRA+pre-lit, with boundaries set at the terciles of the distribution to create comparatively sized groups reflecting a low, medium, and high achieving pupils at baseline. This was completed for the W/C model only as the C/U model targets the lowest achieving 20% of pupils. The results are presented in Appendix A: Additional Descriptive Subgroup Analysis.

Zero score analysis

We also undertook zero score analysis for the secondary outcome measures. As part of the zero score analysis, we estimated the percentage of pupils scoring zero for every secondary outcome measure. These percentages were calculated for the intervention and control groups at baseline and endline for both approaches. The results of the analysis are reported in Appendix B: Zero Score Analysis for Secondary Outcome Measures.

Readiness for trial criteria and workshops

One of the main objectives of this pilot evaluation was to assess whether LRF! is ready for an efficacy trial. In addition to what was specified in the study protocol, and as part of this core objective a list of criteria were developed specific to the pilot evaluation to determine whether the crucial components for an efficacy study are in place. The specific criteria were agreed through discussions between QRF, EEF and NatCen, and they fall into the following categories: evidence of promise, feasibility of intervention, and feasibility of the efficacy trial. Broadly speaking the evidence of promise criteria was used to check whether there is preliminary evidence of impact on intended outcomes and to determine if unintended consequences as a result of the implementation of LRF! were zero or minimal. The evidence of feasibility dimension explored whether LRF! was implemented as intended in most schools (e.g. the schools engage with LRF!, the practice book is used, the training is effective), and considered the quality and appropriateness of the intervention materials, training and the diagnostic tool. Last, the criteria around the feasibility of the efficacy trial aimed to determine whether LRF! is 'ready for trial' by checking if enough is in place to allow the intervention to take place at scale.

The implementation and evaluation partners felt that developing a rating system and quantifying if the criteria have been met would be inappropriate as most of them are concepts that cannot be readily quantified (e.g. quality of training could not be assessed against certain thresholds).

¹⁹ The Framework approach is a type of thematic analysis which evidences the relationship between themes and anonymised cases in published findings.



Therefore, we did not define a weighting or coding system to assess success or failure on certain criteria. Instead, we agreed to present a comprehensive summary of the evidence in two virtual workshops. These workshops took place in May and July 2022. The main aim of the workshops was to present the evidence on each criterion and to create a space in which all implementation and delivery partners could reflect openly on the pilot study, and discuss what went well and what could be improved. The research questions and criteria, the source of data, and the key conclusions from the assessment are presented in Appendix D. Considering that there are implementation differences across the C/U and the W/C model, and across Grade 1 and Grade 2 and 3 pupils we presented the evidence for the W/C Grade 1, C/U Grade 1 and C/U Grade 2 and 3 separately.

Modifications to the IPE

Some of the modifications to the intended IPE design include the following:

- For the C/U model, only resource room teachers completed the survey. In semester one, for Grades 2/3, resource room teachers completed the survey via telephone in January. This was due to the sudden closure of schools in early December, and to limit the number of fieldwork activities taking place in the final week of teaching, the survey was postponed until the new year. The same resource room teachers would have completed the survey for Grade 1 C/U in semester two, at the same time as the second observations;
- The baseline pupils survey in C/U schools Grade 2 and 3 was not administered and collected due to the school closure in December 2021;
- Interviews with school supervisors were not undertaken due to logistical difficulties.

TIMELINE

Table 8: Timeline of the evaluation activities

Date	Activity
July 2021	Inception meetings
July 2021	MOE and MOSD approval
27 July 2021	Logic model workshop
August 2021	Review and update intervention materials based on pre-pilot findings
August 2021	Confirm pre-test logistics and recruit sample
11 August 2021	Ethical approval for pre-test
22 – 26 August 2021	Pre-test fieldwork
31 August 2021	Pilot sampling criteria agreed
2 – 13 September 2021	Pilot sampling and recruitment
9 – 22 September 2021	Dissemination of pre-test findings and finalised EGRA+PreLit tool
22 September 2021	Teacher training for RAMP diagnostic tool
24 September 2021	Ethical approval for EGRA testing
25 September and 2 October 2021	C/U teacher training for LRF!
26 – 29 September 2021	Grade 2/3 C/U and control baseline data collection
30 September 2021	Logic model signed off

3 October 2021 – March 2022	Grade 2/3 C/U intervention delivery
October and November 2021	Development and sign off of IPE materials
31 October 2021	Ethical approval for IPE activities
10 November 2021	Trial registered
14 – 18 November 2021	Grade 2/3 C/U initial classroom observations
5 – 9 December 2021	Grade 2/3 C/U IPE data collection: follow up classroom observations, and FGDs with parents, pupils, and teachers
27 January 2022	Grade 2/3 C/U coaches FGD
2 February 2022	W/C teacher training for LRF!
17 February 2022	Grade 2/3 C/U teacher survey
20 – 24 February 2022	W/C baseline data collection
27 February – 3 March 2022	Grade 1 C/U and Control baseline data collection
6 March – 2 June 2022	Grade 1 C/U and W/C intervention delivery
13 – 17 March 2022	Grade 1 C/U and W/C initial classroom observations
20 – 24 March 2022	Initial Grade 2/3 C/U and Control endline data collection
24 – 28 April 2022	Follow up Grade 2/3 C/U and Control endline data collection
8 – 12 May 2022	Grade 1 C/U and W/C IPE data collection: follow up classroom observations, teacher surveys, and FGDs with parents, pupils, and teachers.
5 – 9 June 2022	Control parents and teachers FGD
5 – 9 June 2022	Grade 1 C/U, W/C, and Control endline data collection
21 June 2022	Grade 1 C/U and W/C coaches FGD



IMPACT EVALUATION - FINDINGS

Participant flow including losses and exclusions

The participant flow diagram is presented in Figure 3 below. The diagram shows that 112 primary schools were approached for the pilot evaluation and among these, 24 primary schools agreed to participate in the pilot evaluation. As part of the recruitment process, QRF developed a recruitment log that summarised reasons for schools' non-participation. Table 9 below shows that the main reason schools did not take part in the study was that QRF was unable to reach them because they had different phone numbers. This was the case for 63 out of 88 primary schools that did not participate. Other reasons for non-participation included the following: schools preferred not to take part (14 schools), schools delivered blended teaching (10 schools), and schools did not have official resource room teacher (1 school).

Table 9 also shows the numbers of schools that agreed to take part and the number of schools that the delivery team tried to reach by geographical region. It is clear from the table that while it was easier to recruit schools in near south, it was more challenging in Amman, middle excluding Amman and near north regions. For instance, QRF had to contact 8 primary schools in near South to recruit 6 schools from the region, while it had to contact 55 primary schools in Amman to recruit 6 schools. The future recruitment process should consider these regional differences.

The recruited 24 primary schools (764 pupils) remained in the trial. Following randomisation, 8 schools (269 pupils) were assigned to the W/C intervention group, 8 schools (150 pupils) to C/U intervention group, and the remaining 8 schools (219 pupils for W/C and 149 pupils for C/U) were assigned to the control group.

Once schools selected pupils to take part in the research, all children were given the opportunity to opt out if they wished. This was firstly done through verbal agreement and then more formally through pupils' clicking a button at the start of the online EGRA test. No pupil opted out from the pilot evaluation. Baseline data was collected for 764 pupils from 24 primary schools.

Before collecting endline data, 2 primary schools (59 pupils) from the C/U intervention group dropped out. One of the schools dropped out because the trained teacher was absent from work for personal reasons. For the second drop-out, the reason was that the resource room teacher preferred to implement the traditional teaching approaches. QRF recruited two replacement schools from the random list of schools provided to them by NatCen during the recruitment process. Given that the replacement schools could implement the intervention in the second term only, baseline data was collected for Grade 1 pupils only.

At the time of analysis, data had been collected from 711 pupils in 24 primary schools.

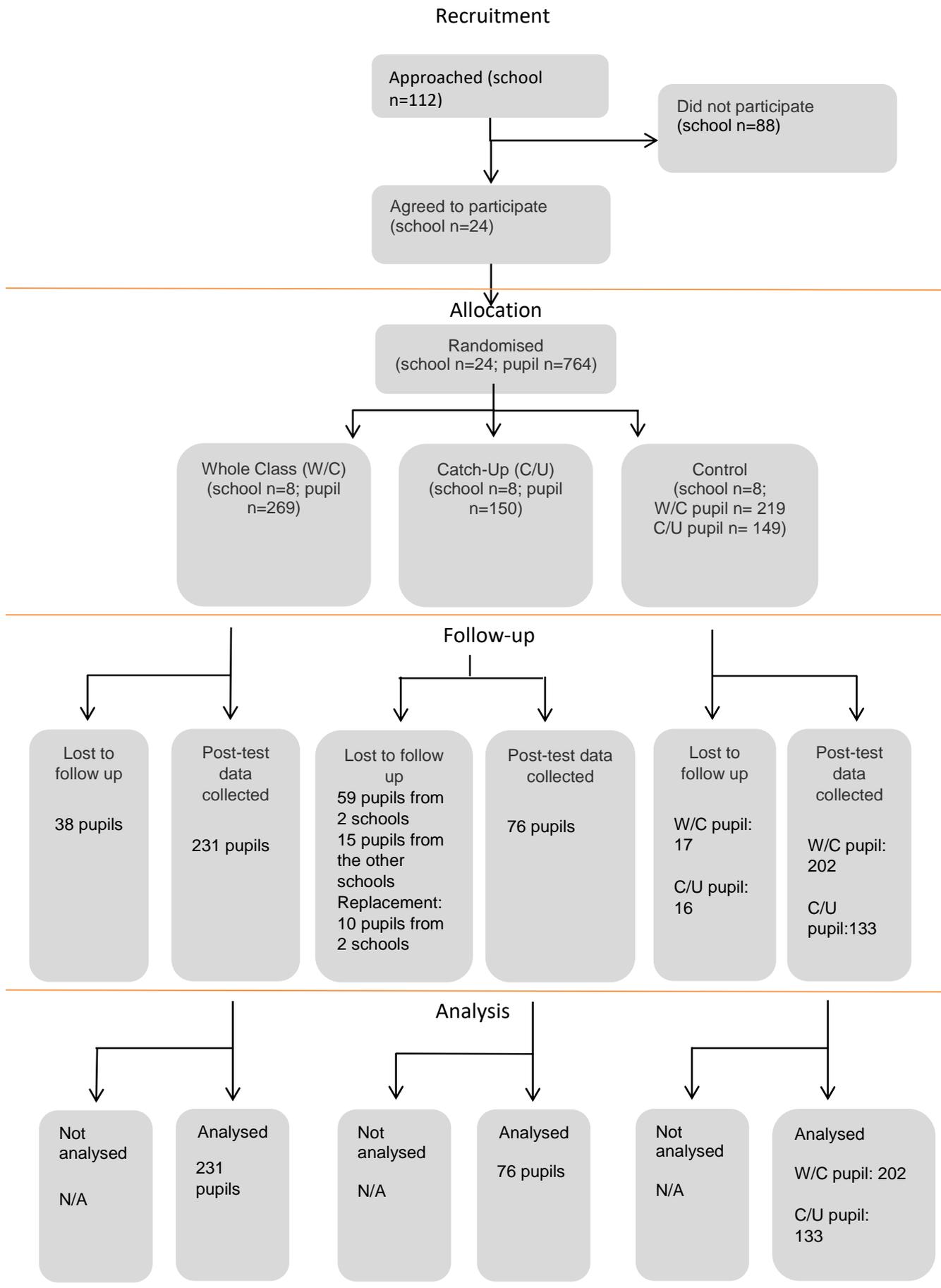


Table 9: Common reasons why schools across the different regions did not take part in the evaluation

Reasons for not participating	Number of schools by geographical regions				
	Amman	Middle excluding Amman	Near North	Near South	Total
Prefer not to participate	8	4	2	0	14
Unable to reach school/wrong number	41	14	6	2	63
Delivering blended teaching	0	0	10	0	10
No official resource room teacher	0	0	1	0	1
Total number of schools not participating	49	18	19	2	88
Number of schools agreed to participate	6	6	6	6	24
Total number of schools	55	24	25	8	112



Figure 1: Participant flow diagram



ATTRITION

No W/C intervention school or control school dropped out from the trial. Table 10 below shows the pupil level attrition from the trial for the W/C model. For this model, baseline data was available for 269 pupils in 8 W/C intervention schools and 219 pupils in 8 control schools. At the analysis stage, data was available for 231 pupils in 8 W/C intervention schools and 202 pupils in 8 control schools.

The overall ratio between pupils analysed and pupils randomised in the W/C model is 433 to 488. The retention rate is 88.7% and the attrition rate is 11.3%. In the control group for the W/C model, the retention rate is higher than the overall retention rate for the W/C model. It is 95.5% (an attrition rate of 4.5%) with 202 pupils in the analysis from 219 pupils randomised. In the intervention group, the retention rate is 85.9% (an attrition rate of 14.1%) with 231 pupils analysed from 269 pupils randomised.

Table 10: Pupil level attrition from the trial (primary outcome) – W/C

		Intervention	Control	Total
Number of pupils	Randomised	269	219	488
	Analysed	231	202	433
Pupil attrition (from randomisation to analysis)	Number	38	10	55
	Percentage	14.1%	4.5%	11.3%

Two C/U intervention schools dropped out from the trial due to reasons mentioned above. These schools are not included in the final analysis and have been listed under the “lost to follow-up” category in the participant flow diagram. QRF recruited two replacement schools, which implemented the intervention in the second term only. Consequently, only Grade-1 pupils could participate in the trial from these replacement schools. No C/U control schools dropped out from the trial.

Table 11 below shows the pupil level attrition from the trial for the C/U model. At randomisation stage baseline data was available for 150 pupils in 8 C/U intervention schools and 149 pupils in 8 control schools. After excluding the dropped schools and including the replacement schools baseline data was available for only 101 pupils from the C/U intervention schools. At the analysis stage, data was available for 76 pupils in 8 C/U intervention schools and 133 pupils in 8 control schools.

The overall retention rate for the C/U model between randomisation and endline data collection is 69.9% (an attrition rate of 30.1%). The retention rate in the control group was significantly higher than that in the intervention group. The retention rate for the control group is 89.3% (an attrition rate of 10.7%) with 133 pupils in the analysis from 149 pupils randomised while the retention rate for the C/U intervention group is 50.7% (an attrition rate of 49.3%) with 76 pupils analysed from 150 pupils randomised.



Table 11: Pupil level attrition from the trial (primary outcome) – C/U Model

		Intervention	Control	Total
Number of pupils	Randomised	150	149	299
	Analysed	76	133	209
Pupil attrition (from randomisation to analysis)	Number	74	16	90
	Percentage	49.3%	10.7%	30.1%

The data collection team mobilised a “clean-up” crew to test pupils that missed EGRA endline testing. The crew completed up to three additional visits in some schools. They also kept a log recording the reasons why pupils did not complete the endline assessment. The main reasons pupils were not tested were the following:

- illness (e.g. chickenpox) (40% of pupils not tested)
- changing a school or place of residence (20% of pupils not tested)
- persistent attendance issues (11% of pupils not tested)
- teacher absence (11% of pupils not tested)
- unknown reasons (11% of pupils not tested)
- family circumstances (4% of pupils not tested)
- helping family with work (shepherds) (2% of pupils not tested), and
- school phobia (1% of pupils not tested).

Primary outcome analysis

This evaluation’s primary outcome was a single learning metric for reading known as EGRA+pre-lit, which is described in the outcome measures section. This metric reflects pupils’ overall abilities/performance in Arabic literacy and was rescaled to have a mean of 500 and a standard deviation of 100. A higher score on the primary outcome indicates a better overall ability in Arabic literacy.

Estimates of effect sizes from the primary outcome analysis will be highly imprecise due to the small sample size of the pilot trial. We therefore caution against drawing conclusions around the impacts of the interventions.

W/C Model

In total, there were 433 pupils with both baseline and endline data that were included in the primary outcome analysis for the W/C model. For pupils in the analysis, the baseline EGRA+pre-lit score had a mean of 492.9 and standard deviation of 86.9 due to following concurrent equating.²⁰ The range for the analytical sample was from 320.4 to 809.7. The endline EGRA+pre-lit score had a mean of 549.1 and standard deviation of 109.4. The range for the analytical sample was from 320.4 to 822.6. The distribution of the baseline and endline EGRA+pre-lit score is illustrated in Figure 4 and in Figure 5, below. The spread of the outcomes resembles a normal distribution, indicating that the assumptions underlying the parametric tests used in the analysis are not violated. The endline EGRA+pre-lit score was correlated to the baseline EGRA+pre-lit score, with r

²⁰ Concurrent equating means that the data were scaled using baseline and endline data pooled together (see technical report produced by OM). If the baseline and endline data is combined, the EGRA+pre-lit score has a mean of 500 and a standard deviation of 100.



= 0.72, suggesting a moderately positive correlation. The derived correlation coefficient can be used to inform the required sample size for future evaluations.

Figure 4: Baseline EGRA+pre-lit Score (W/C Model)

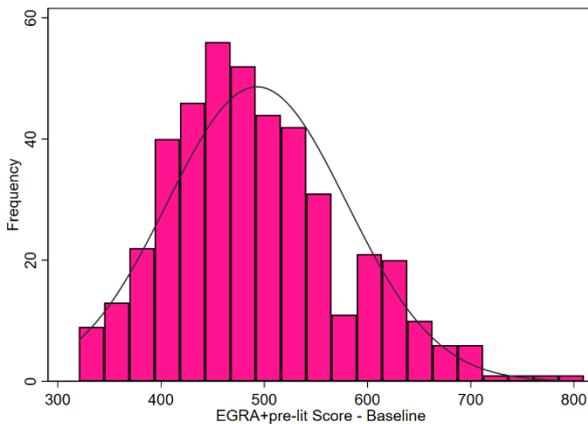
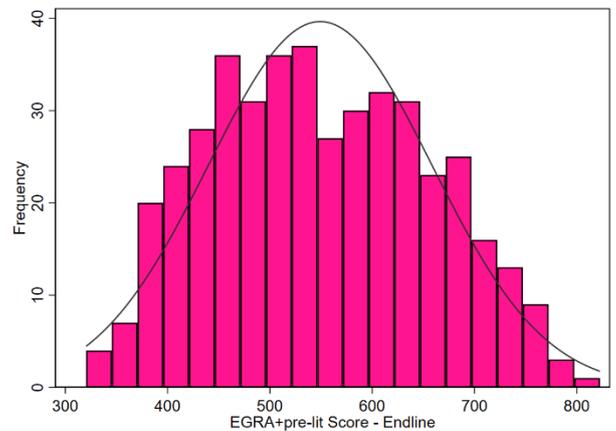


Figure 5: Endline EGRA+pre-lit Score (W/C Model)



The unadjusted EGRA+pre-lit mean score for the intervention group was larger than that for the business-as-usual control group. Table 12, below, illustrates that the unadjusted difference in means between the intervention and control group was 65.8. In the multilevel model that accounted for clustering an adjusted difference in EGRA+pre-lit means between the intervention and control group was 46.1. The p-value associated with the difference in means was $p < 0.0001$, while the effect size associated with this adjusted difference in means is 0.20 (95% CI: 0.01– 0.39) (Table 13).

This indicates that the W/C approach had a positive impact on Arabic literacy. As indicated by the confidence intervals, this is an imprecise estimate with the lower limit being very close to zero. Furthermore, as pointed out earlier, the pilot evaluation was not large enough to establish evidence of causal impact.

Lastly, as outlined in the methods section, we estimated the post-intervention ICCs based on the primary outcome (EGRA+pre-lit). The estimated ICC could be used to help the development of new cluster studies in the area. We have not made any adjustments when calculating the ICCs.²¹ Based on EGRA+pre-lit data, the ICC within schools is 0.18 with 95% confidence interval of 0.065 and 0.4. The ICC in classes within schools is 0.29 with 95% confidence interval of 0.167 and 0.454.

²¹ The ICC measured similarity between units in the same cluster; in this case, pupils within the same classroom. Units within the same cluster may exhibit similarities due to being exposed to similar environmental characteristics.



Table 12: Primary Outcome Analysis W/C and C/U

Outcome	Unadjusted diff. in means	Adjusted diff. in means	Intervention group		Control group		Pooled variance
			n (missing)	Var. of outcome	n (missing)	Var. of outcome	
Arabic Literacy Attainment – W/C Model	65.81	46.16	231 (38)	37072.43	202 (17)	65182.48	50181.76
Arabic Literacy Attainment – C/U Model	-12.51	-5.18	76 (74)	7346.94	133 (16)	38067.53	26936.88

Note. diff.= difference; var.= variance

Table 13: Effect size estimation - primary outcome analysis W/C and C/U

Outcome	Unadjusted means				Effect size		
	Intervention group		Control group		Total n	Hedges g (95% CI)	p-value
n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)				
Arabic Literacy Attainment – W/C Model	231 (38)	579.01 (554.18 – 603.84)	202 (17)	522.25 (487.04 – 557.46)	433	0.20 (0.01– 0.39)	0.00
Arabic Literacy Attainment – C/U Model	76 (74)	477.81 (458.54 – 497.08)	133 (16)	492.42 (459.25 – 525.57)	209	-0.03 (-0.31 – 0.25)	0.64



C/U Model

In total, 209 pupils with both baseline and endline data were included in the primary outcome analysis for the C/U model. The same caveats discussed for the W/C model estimates apply to the C/U subsample, with the additional issues of smaller sample sizes and unequal attrition between the C/U and control group (see Attrition).

The baseline EGRA+pre-lit score had a mean of 438.8 and standard deviation of 70.7. The range for the analytical sample was from 320.4 to 706.7. The distribution of the baseline EGRA+pre-lit score is illustrated in. The endline EGRA+pre-lit score had a mean of 482.5 and standard deviation of 91.9. The range for the analytical sample was from 320.4 to 732.5. The distribution of the baseline and endline measure is provided in Figure 6 and Figure 7, respectively. The distribution is skewed towards the lower end of the range of scores, particularly at baseline. This was expected given that the test was completed by struggling pupils.

For the C/U model, the endline EGRA+pre-lit score was positively correlated to the baseline EGRA+pre-lit score, with $r = 0.63$. The estimated correlation coefficient is suggesting moderately positive correlation between baseline and endline scores.

Figure 6: Baseline EGRA+pre-lit Score (C/U Model)

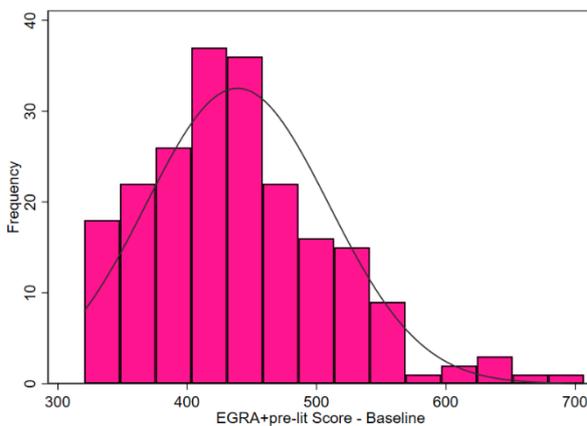
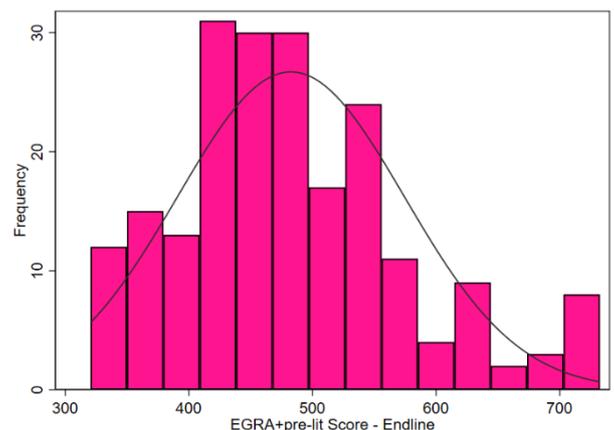


Figure 7: Endline EGRA+pre-lit Score (C/U Model)



On the contrary to the W/C model, the unadjusted EGRA+pre-lit mean score for the intervention group in the C/U model was smaller than that for the business-as-usual control group in the C/U model at endline. Table 12, above, illustrates that the unadjusted difference in means between the intervention and control group was -12.5. After controlling for pre-intervention covariates, an adjusted difference in means between the intervention and control group was -5.1 and was not statistically significant with $p=0.636$. The effect size associated with this adjusted difference in means was -0.03 (95% CI: -0.31 – 0.25) (see Table 13). Therefore, the results provide no indication of any positive or negative impact of the C/U approach on literacy based on the EGRA+pre-lit assessment.

The post-intervention ICCs were estimated from EGRA+pre-lit data from both intervention and control schools. We have not made any adjustments when calculating the ICCs. Based on EGRA+pre-lit data, the ICC within schools is 0.116 with 95% confidence interval of 0.037 and 0.31. The ICC in classes within schools is 0.127 with 95% confidence interval of 0.044 and 0.315.



SECONDARY OUTCOME ANALYSIS

The protocol set out four secondary outcome measures: oral reading fluency (ORF), letter sound identification (LSI), syllable identification (SI) and reading comprehension (RC). The results of the secondary outcome analyses are presented separately for the W/C and C/U models for each secondary outcome measure.

As is the case for the primary outcome analysis, the focus of this analysis is not on attributing causality for changes in outcomes to the W/C or C/U model as the analysis is not powered to detect meaningful differences. Estimates of effect sizes will be highly imprecise due to the small sample size of the trial. We, therefore, caution against drawing conclusions around the impacts of the interventions on the secondary outcome measures.

W/C Model

In total 433 pupils with both baseline and endline data were included in the secondary outcome analysis for the W/C model.

Oral Reading Fluency (ORF)

The first secondary outcome measure in this trial was ORF, which is measured as the number of correct words per minute.²² The baseline ORF score had a mean of 2.27 and standard deviation of 5.60. The distribution of the baseline ORF score is illustrated in Figure 8. The endline ORF score had a mean of 6.33 and standard deviation of 9.75. The distribution of this measure is provided in Figure 9, below. This measure was not perfectly normally distributed, with the majority of pupils obtaining a total score of 0 on this sub-task. There was, therefore, a large floor effect for this measure.²³

The unadjusted mean difference between the intervention and control group was 3.95 (Table 14). The adjusted mean difference between the intervention and control group was 2.78 and was statistically significant with $p=0.004$ (Table 15). The effect size associated with this difference was 0.16 (95% CI: -0.02 – 0.35) (Table 15).

The results indicate that oral reading fluency was similar in both W/C and control schools.

²² This measure takes value between 0 and 42.

²³ See Appendix B for more information on pupils scoring zero on EGRA subtasks.



Figure 8: Baseline ORF Score (W/C Model)

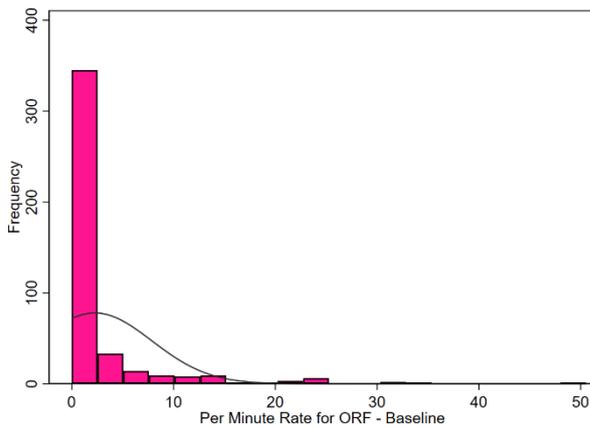
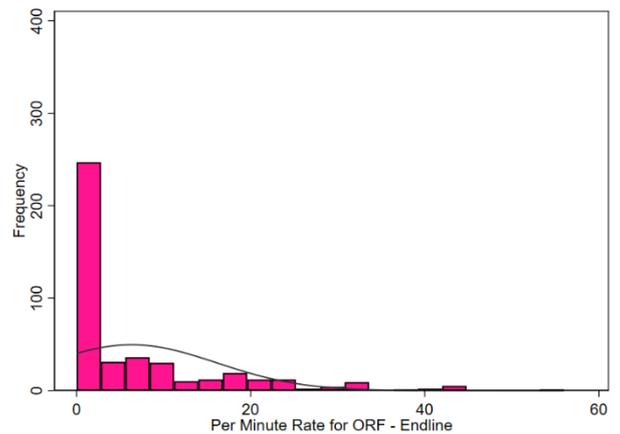


Figure 9: Endline ORF Score (W/C Model)



Letter Sound Identification (LSI)

The second secondary outcome measure was LSI, which is measured as the number of correct letters per minute.²⁴ The baseline LSI score had a mean of 15.90 and standard deviation of 16.30. The distribution of the baseline LSI score is illustrated in Figure 10. The endline LSI score had a mean of 25.40 and standard deviation of 21.62. The distribution of this measure is provided in Figure 11, below. This measure was not perfectly normally distributed, with the majority of pupils obtaining a total score of 0 on this sub-task. There was, therefore, a large floor effect for this measure.²⁵

The unadjusted mean difference between the intervention and control group was 13.7 (Table 14). The adjusted mean difference between the intervention and control group was 10.6 and was statistically significant at $p < 0.001$ (Table 15). The effect size associated with this difference was 0.3 (95% CI: 0.11 – 0.49) (Table 15). The results indicate that the W/C model had a positive effect on letter sound identification.

Figure 10: Baseline LSI Score (W/C Model)

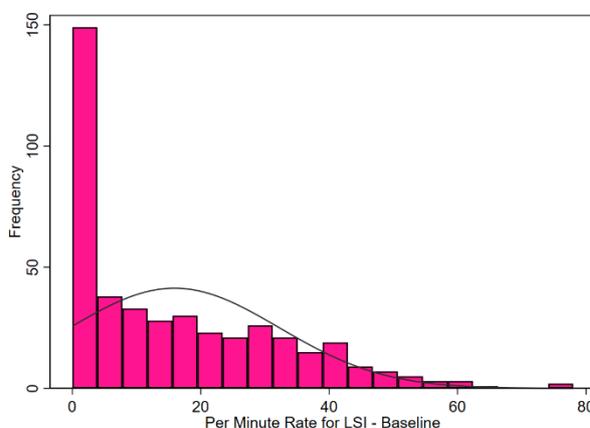
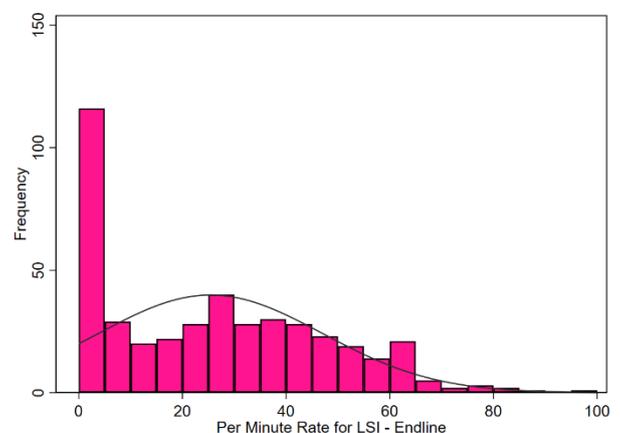


Figure 11: Endline LSI Score (W/C Model)



²⁴ This measure takes value between 0 and 100.

²⁵ See Appendix B for more information on pupils scoring zero on EGRA subtasks.



Syllable Identification (SI)

The third secondary outcome measure was syllable identification (SI), which is measured as the number of correct syllables per minute.²⁶ The baseline SI score had a mean of 6.94 and standard deviation of 9.46. The distribution of the baseline SI score is illustrated in Figure 12. The endline SI score had a mean of 13.98 and standard deviation of 15.24. The distribution of this measure is provided in Figure 13, below. This measure was not perfectly normally distributed, with the majority of pupils obtaining a total score of 0 on this sub-task. There was, therefore, a large floor effect for this measure.²⁷

The unadjusted mean difference between the intervention and control group was 8.23 (Table 14). The adjusted mean difference between the intervention and control group was calculated as 6.42 and was statistically significant at $p < 0.001$ (Table 15). The effect size associated with this difference was 0.23 (95% CI: 0.04 – 0.42), suggesting that the W/C approach had a positive effect on syllable identification (Table 15). The very wide confidence interval suggests that this is a very imprecise estimate.

Figure 12: Baseline SI Score (W/C Model)

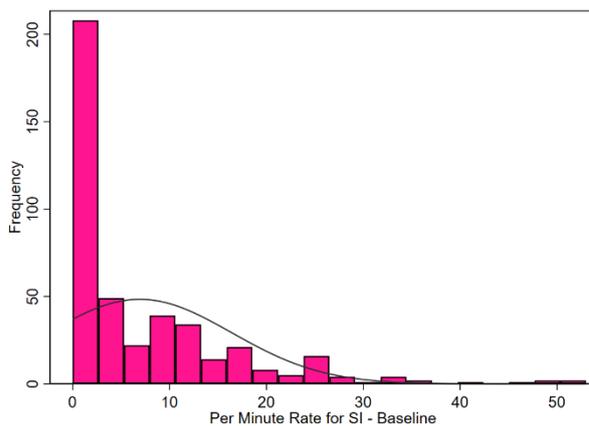
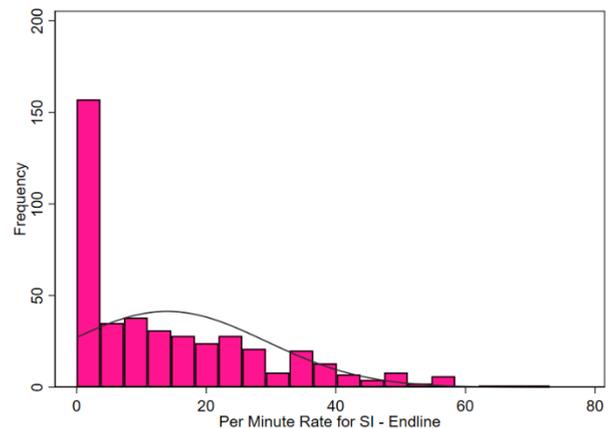


Figure 13: Endline SI Score (W/C Model)



Reading Comprehension (RC)

The last secondary outcome measure was reading comprehension (RC), which is measured as the number of reading comprehension questions a pupil answered correctly.²⁸ The baseline RC score had a mean of 0.16 and standard deviation of 0.50. The distribution of the baseline RC score is illustrated in Figure 14. The endline RC score had a mean of 0.48 and standard deviation of 1.00. The distribution of this measure is provided in Figure 15, below. This measure was not perfectly normally distributed, with the majority of pupils obtaining a total score of 0 on this sub-task. There was, therefore, a large floor effect for this measure.²⁹

The unadjusted mean difference between the intervention and control group was 0.29 (Table 14). The adjusted mean difference between the intervention and control group was calculated as 0.24

²⁶ This measure takes value between 0 and 100.

²⁷ See Appendix B for more information on pupils scoring zero on EGRA subtasks.

²⁸ This measure takes value between 0 and 5.

²⁹ See Appendix B for more information on pupils scoring zero on EGRA subtasks.



and was statistically significant with $p=0.011$ (Table 15). The effect size associated with this difference was 0.16 (95% CI: -0.029 – 0.34) (Table 15). This analysis suggests that the W/C approach had a positive effect on reading comprehension.

Figure 14: Baseline RC Score (W/C Model)

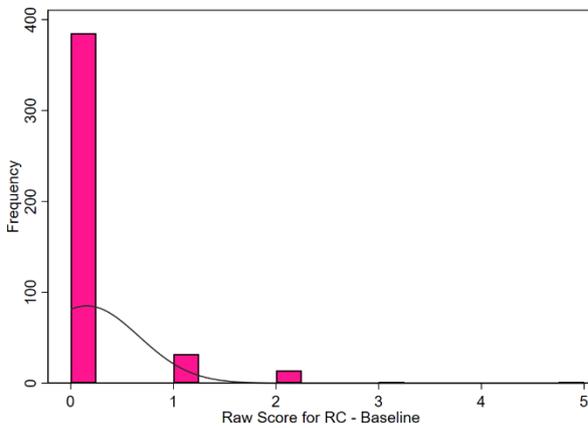


Figure 15: Endline RC Score (W/C Model)

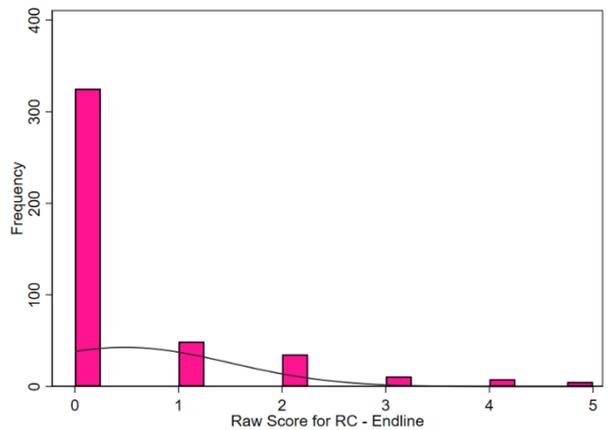


Table 14: Secondary outcome analyses for W/C model

Outcome	Unadjusted diff. in means	Adjusted diff. in means	Intervention group		Control group		Pooled var.
			n (missing)	Var. of outcome	n (missing)	Var. of outcome	
Oral Reading Fluency	3.95	2.78	231 (38)	268.21	202 (17)	296.98	281.63
Letter Sound Identification	13.7	10.6	230 (39)	1017.16	202 (17)	1479.89	1232.96
Syllable Identification	8.23	6.42	231 (38)	859.83	202 (17)	634.58	754.78
Reading Comprehension	0.29	0.24	231 (38)	2.00	202 (17)	2.4	2.19

Note. diff.= difference; var.= variance

Table 15: Effect size estimation - secondary outcome analyses for W/C model

Outcome	Unadjusted means				Effect size		
	Intervention group		Control group		Total n	Hedges g (95% CI)	p-value
n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)				
Oral Reading Fluency	231 (38)	8.25 (6.14 – 10.36)	202 (17)	4.41 (2.04 – 6.79)	433	0.16 (-0.02 – 0.35)	0.00
Letter Sound Identification	230 (39)	31.82 (27.71 – 35.94)	202 (17)	18.66 (13.35 – 23.96)	432	0.3 (0.11 – 0.49)	0.00
Syllable Identification	231 (38)	17.94 (14.16 – 21.72)	202 (17)	9.95 (6.47 – 13.42)	433	0.23 (0.04 – 0.42)	0.00
Reading Comprehension	231 (38)	0.62 (0.44 – 0.80)	202 (17)	0.34 (0.12 – 0.55)	433	0.16 (0.029 – 0.34)	0.01

C/U Model

A total of 209 pupils with both baseline and endline data were included in the secondary outcome analysis for the C/U model.



Oral Reading Fluency (ORF)

The ORF score is measured as the number of correct words per minute.³⁰ The baseline ORF score had a mean of 0.42 and standard deviation of 1.97. The distribution of the baseline ORF score is illustrated in Figure 16. The endline ORF score had a mean of 1.76 and standard deviation of 5.20. The distribution of this measure is provided in Figure 17, below. This measure was not perfectly normally distributed, with the majority of pupils obtaining a total score of 0 on this sub-task. There was, therefore, a large floor effect for this measure.³¹

The unadjusted mean difference between the intervention and control group was -0.20 (Table 16). The adjusted mean difference between the intervention and control group was 0.24 and was not statistically significant with $p=0.659$ (Table 17). The effect size associated with this difference was 0.03 (95% CI:-0.24 – 0.31) (Table 17). The results indicate that there were no differences in oral reading fluency between pupils in C/U and control schools with the effect size estimates close to zero.

Figure 16: Baseline ORF Score (C/U Model)

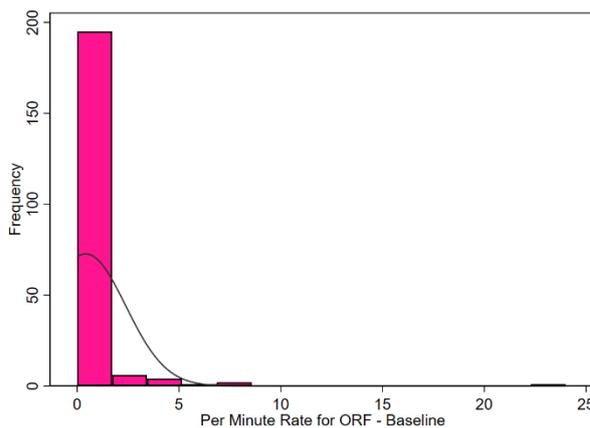
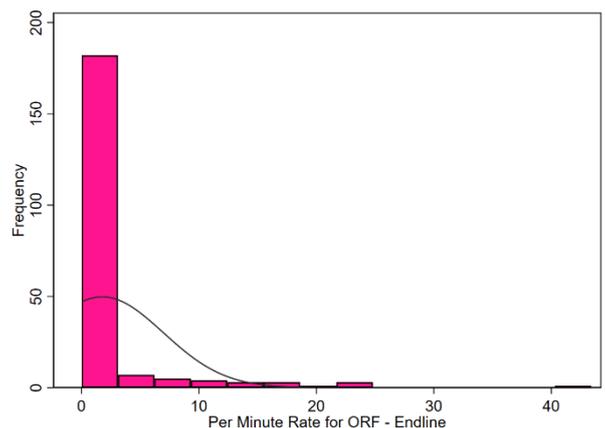


Figure 17: Endline ORF Score (C/U Model)



Letter Sound Identification (LSI)

The LSI score is measured as the number of correct letters per minute.³² The baseline LSI score had a mean of 7.03 and standard deviation of 11.96. The distribution of the baseline LSI score is illustrated in Figure 18. The endline LSI score had a mean of 12.32 and standard deviation of 15.73. The distribution of this measure is provided in Figure 19, below. This measure was not perfectly normally distributed, with the majority of pupils obtaining a total score of 0 on this sub-task. There was, therefore, a large floor effect for this measure.³³

The unadjusted mean difference between the intervention and control group was 0.69 (Table 16). The adjusted mean difference between the intervention and control group was 1.06 and was not statistically significant at $p=0.587$ (Table 17). The effect size associated with this difference was

³⁰ This measure takes value between 0 and 42.

³¹ See Appendix B for more information on pupils scoring zero on EGRA subtasks.

³² This measure takes value between 0 and 100.

³³ See Appendix B for more information on pupils scoring zero on EGRA subtasks.



close to zero, at 0.04 (95% CI: -0.23 – 0.32), suggesting that there was no evidence of impact of the C/U approach on letter sound identification (Table 17).

Figure 18: Baseline LSI Score (C/U Model)

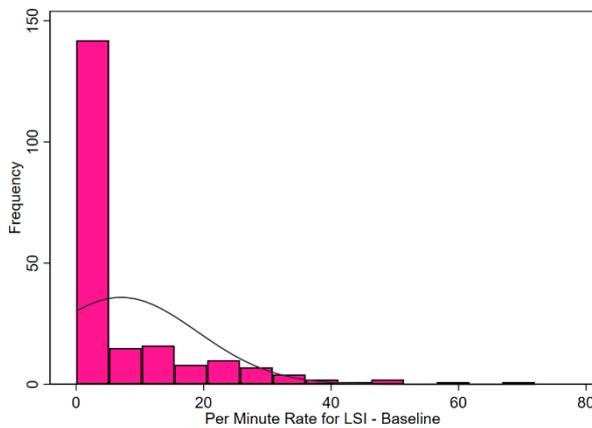
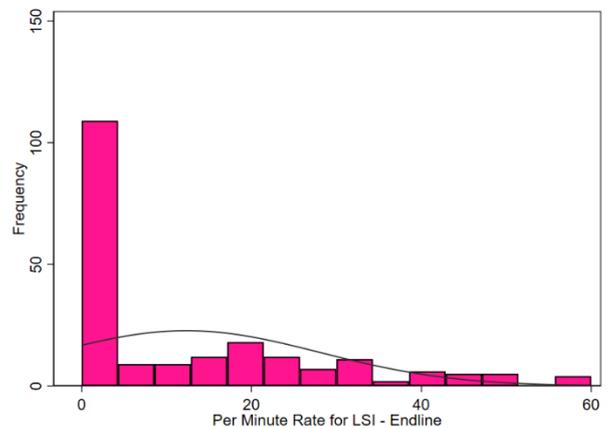


Figure 19: Endline LSI Score (C/U Model)



Syllable Identification (SI)

The SI score is measured as the number of correct syllables per minute.³⁴ The baseline SI score had a mean of 1.61 and standard deviation of 4.21. The distribution of the baseline SI score is illustrated in Figure 20. The endline SI score had a mean of 4.42 and standard deviation of 8.02. The distribution of this measure is provided in Figure 21, below. This measure was not perfectly normally distributed, with the majority of pupils obtaining a total score of 0 on this sub-task. There was, therefore, a large floor effect for this measure.³⁵

The unadjusted mean difference between the intervention and control group was -0.47 (Table 16). The adjusted mean difference between the intervention and control group was calculated as 0.16 and was not statistically significant at $p=0.850$ (Table 17). The effect size associated with this difference was almost zero, at 0.01 (95% CI: -0.26 – 0.29), which indicates no impact of the C/U approach on syllable identification (Table 17).

³⁴ This measure takes value between 0 and 100.

³⁵ See Appendix B for more information on pupils scoring zero on EGRA subtasks.



Figure 20: Baseline SI Score (C/U Model)

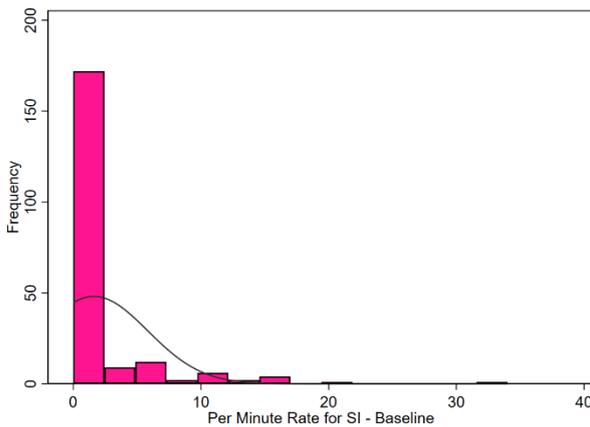
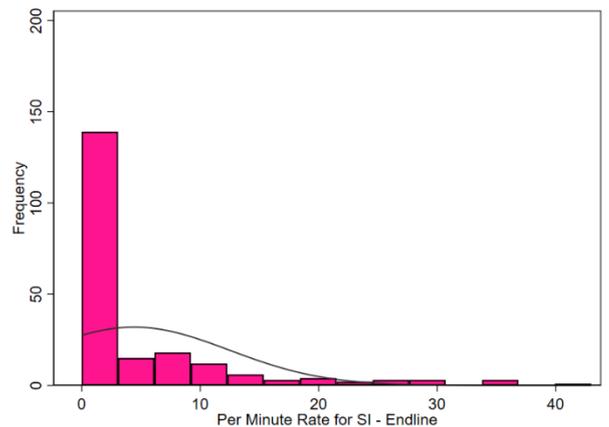


Figure 21: Endline SI Score (C/U Model)



Reading Comprehension (RC)

The RC score is measured as the number of reading comprehension questions a pupil answered correctly.³⁶ The baseline RC score had a mean of 0.02 and standard deviation of 0.18. The distribution of the baseline RC score is illustrated in Figure 22. The endline RC score had a mean of 0.26 and standard deviation of 0.77. The distribution of this measure is provided in Figure 23, below. This measure was not perfectly normally distributed, with the majority of pupils obtaining a total score of 0 on this sub-task. There was, therefore, a large floor effect for this measure.³⁷

The unadjusted mean difference between the intervention and control group was -0.08 (Table 16). The adjusted mean difference between the intervention and control group was calculated as -0.11 and was not statistically significant with $p=0.260$ (Table 17). The effect size associated with this difference was -0.11 (95% CI: -0.40 – 0.16) (Table 17). The results indicate that there was no impact of the C/U approach on reading comprehension.

Figure 22: Baseline RC Score (C/U Model)

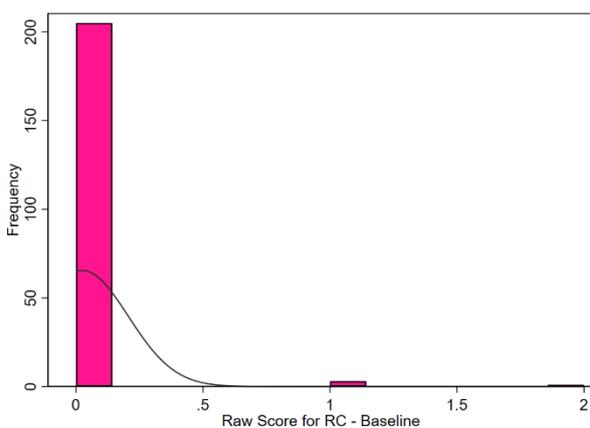
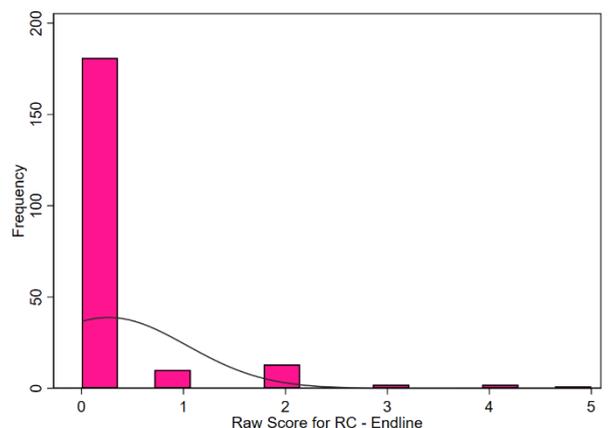


Figure 23: Endline RC Score (C/U Model)



³⁶ This measure takes value between 0 and 5.

³⁷ See Appendix B for more information on pupils scoring zero on EGRA subtasks.



Table 16: Secondary analyses for Catch-Up model

Outcome	Unadjusted diff. in means	Adjusted diff. in means	Intervention group		Control group		Pooled var.
			n (missing)	Var. of outcome	n (missing)	Var. of outcome	
Oral Reading Fluency	-0.20	0.24	76 (74)	21.33	133 (16)	60.18	46.10
Letter Sound Identification	0.69	1.06	76 (74)	276.68	133 (16)	690.43	540.52
Syllable Identification	-0.47	0.16	76 (74)	75.51	133 (16)	185.09	145.39
Reading Comprehension	-0.08	-0.11	76 (74)	0.58	133 (16)	1.16	0.95

Note. diff.= difference; var.= variance

Table 17: Effect size estimation - secondary outcome analyses for Catch-Up model

Outcome	Unadjusted means				Effect size		
	Intervention group		Control group		Total n	Hedges g (95% CI)	p-value
n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)				
Oral Reading Fluency	76 (74)	1.72 (0.68 – 2.76)	133 (16)	1.78 (0.46 – 3.10)	209	0.03 (-0.24 – 0.31)	0.66
Letter Sound Identification	76 (74)	13.23 (9.49 – 16.97)	133 (16)	12.84 (8.37 – 17.30)	209	0.04 (-0.23 – 0.32)	0.59
Syllable Identification	76 (74)	4.47 (2.52 – 6.42)	133 (16)	4.77 (2.45 – 7.08)	209	0.01 (-0.26 – 0.29)	0.85
Reading Comprehension	76 (74)	0.21 (0.04 – 0.38)	133 (16)	0.28 (0.10 – 0.47)	209	-0.11 (-0.40 – 0.16)	0.26



IMPLEMENTATION AND PROCESS EVALUATION – FINDINGS

Fidelity

This section explores implementation fidelity, or the extent to which LRF! followed the intended delivery plan. It includes how much of the intervention was delivered (Dosage), how well the intervention was delivered (Quality), and the nature and extent of changes made to the intervention (Adaptation).

RQ5: Was LRF! delivered as intended in terms of dosage, nature and quality? What modifications were made, with what implications?

Dosage

Training

Based on attendance data from the LRF! training sessions, almost all teachers attended the training. For the C/U model, all seven resource room teachers attended the training (100%).³⁸ For the W/C model, 18 out of 20 classroom teachers attended the training (90%). Reasons for non-attendance were not recorded. As part of the teacher survey, teachers were asked if they had attended the training, and these findings broadly mirror the attendance data. For C/U all resource room teachers said they had attended training. For W/C, 19 out of 20 teachers said they had³⁹. For the one teacher who said they did not attend training, they explained they received training from other teachers within in their school.

Coaching

At the time of the survey,⁴⁰ both C/U and W/C teachers indicated they had received at least one coaching session. The average number of coaching sessions for C/U was 1.75 and for W/C 2.6. This difference is likely due to variation in the survey administration between semester one and two. Despite the variation, for their most recent coaching session, all teachers were observed for the full lesson and clear action points were agreed and provided by the coaches.

LRF! Content

LRF! delivery observed during both W/C and C/U lessons generally followed the intended approach. Across both arms of the intervention, most teachers:

- Gave a ‘wrap up’ of the previous class.
- Pointed to the letters and words as they read.
- Conducted the ‘I do’ section by modelling the reading for pupils.
- Had the pupils echo their (the teacher’s) reading, during the ‘We do’ section.
- Assigned pages for pupils to read during the ‘You do’ section.
- Had the pupils read individually as part of the ‘You do’ section.
- Got pupils to follow the text with their finger as they read.

³⁸ The seven eligible C/U resource room teachers are from six schools. This figure excludes the two schools who dropped out of the LRF! program and the two replacement schools who were separately trained.

³⁹ This is one person more than the QRTA records indicated. However, as the teacher survey was anonymous, we have not excluded the participant.

⁴⁰ W/C teachers answered the survey towards the end of semester two. C/U resource room teachers answered the survey twice, first for Grades 2/3 in January and then again for Grade 1 at the second observation towards the end of the semester two. The C/U average is based on a combination of these scores, but the COVID-19 school closures impacted the number of coaching sessions in the first semester (for Grades 2/3).



- Rotated around the classroom to observe pupils.
- Gave appropriate praise to pupils.

Across the observations, there was little difference in the delivery of LRF! in W/C and C/U schools.

LRF! Practice book

During observations, all schools (both W/C and C/U) had a big print copy of the practice book in the classroom and almost all pupils had a copy of the practice book.

The teacher survey found that W/C pupils were taking the practice book home daily, with 19 out of 20 W/C teachers assigning one-two pages for homework. The other teacher assigned three-four pages for homework. The proportion of pupils completing homework varied between schools, although every W/C teacher indicated that at least some of the pupils completed the assigned pages. The most common response was that over half of pupils completed the assigned pages (45%); no W/C teacher responded that all their pupils completed the homework.

For the C/U model, the teacher survey found that almost every resource room teacher was getting pupils to take the practice book home daily, in line with intended delivery. Although in semester one, for Grades 2/3, one teacher reported that pupils took the practice book home a few times each week. Furthermore, the number of pages set for homework also varied between the two semesters. For semester two, with Grade 1, all C/U resource room teachers assigned one to two pages. For semester one, Grades 2/3, half assigned one-two pages, a quarter assigned three-four pages, and a quarter did not assign any pages. Teachers were not told how much homework to set. Instead, teachers were advised to use their professional judgement based on students' abilities.

Despite these differences, most C/U pupils did not complete the homework. Two out of three Grades 2/3 C/U resource room teachers (67%) and three out of four Grade 1 C/U resource room teachers (75%) indicated that less than half of pupils completed the assigned pages (see Delivery).

During the pupil survey, most pupils across both intervention arms recalled having a copy of the practice book (95% for W/C and 88% for C/U). The majority of these pupils stated they took the practice book home (97% for W/C and 85% for C/U), which is consistent with the findings from the teacher survey. Furthermore, most pupils stated that they read the practice book at home (95% for W/C and 78% for C/U). When asked who they read the practice book with, for both W/C and C/U the most common responses were parents (62% and 64% respectively), followed by individual reading (38% and 25%) and with siblings (17% and 18%). Notably, C/U pupils were far less likely to read the practice book by themselves.

Timing

The teacher survey findings indicate that all teachers (W/C and C/U) delivered at least three sessions per week on average.

However, in FGDs, a recurring view was that the number of LRF! sessions should be spread out across two semesters. C/U resource room and W/C teachers believed this would make it easier to cover all the LRF! content alongside curriculum work (see Quality). Coaches, parents, and pupils commented that due to the difficulty of LRF! content, it should be taught over a whole school year.

Parents were supportive of their children having additional reading practice.



However, coaches acknowledged that in some C/U schools sessions were taking place during maths and Arabic lessons. They explained how this upset parents who felt that their children were missing curriculum content.



QUALITY

RQ2: How do principals, teachers, parents and pupils perceive the intervention and any changes that it has delivered?

RQ4: Is there any evidence of unintended consequences (negative or positive) as a result of the implementation of LRF!?

RQ5: Was LRF! delivered as intended in terms of dosage, nature and quality? What modifications were made, with what implications?

RQ6: What is the learning about teacher's use of the Coarse-Grained Diagnostic (RAMP) tool? How successful is, it in use, at identifying the most appropriate pupils for the Literacy C/U implementation model (C/U model only)?

RQ8. To what extent do teachers develop sufficient skills and confidence through the training and coaching?

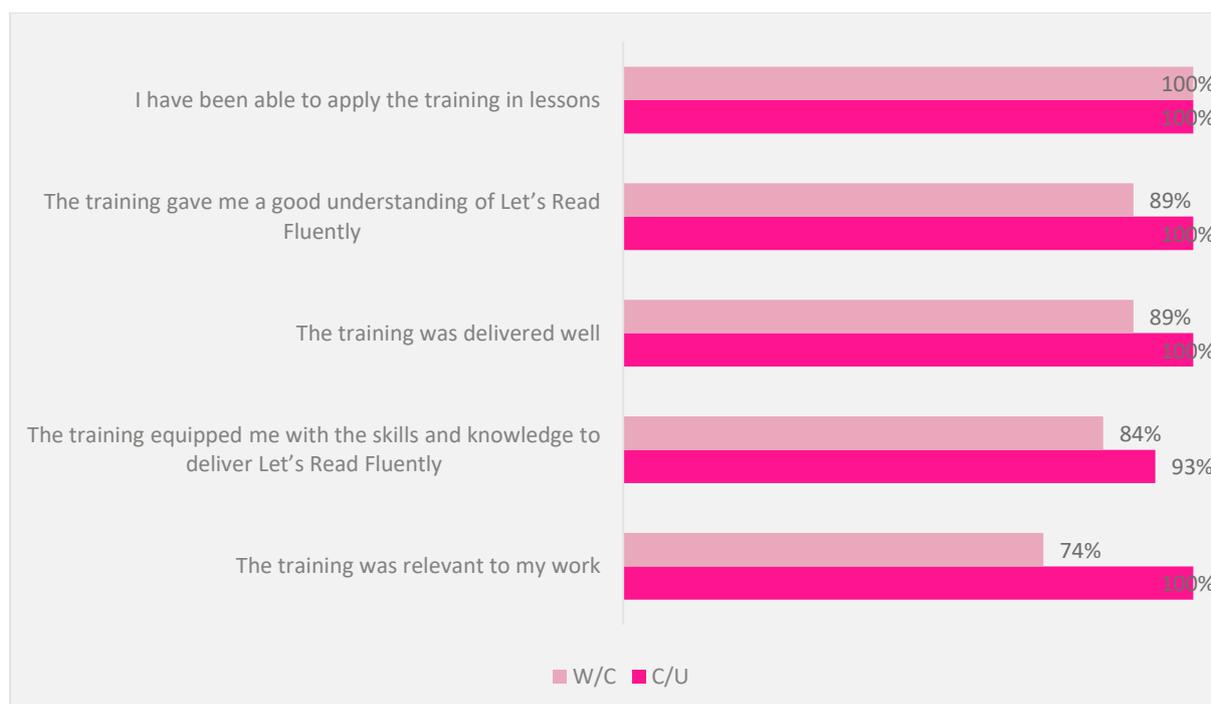
RQ9: What do we know about how teachers need to be supported (coached) during delivery?

RQ10: Are there any key contextual factors that appear to facilitate or impede successful implementation of LRF!?

Training

For both W/C and C/U, the teacher survey found that teachers were broadly positive about the training (see Figure 24). Across both intervention arms, the majority of teachers agreed that the training had given them a good understanding of LRF! and equipped them with the skills and knowledge to deliver the intervention. Moreover, all respondents agreed they had applied the training in lessons.

Figure 24: Perceptions of LRF! Training - Agreement



Source: Teacher survey, W/C n=19 and C/U n=15



In FGDs, feedback on the training was similarly positive from both C/U resource room teachers and W/C classroom teachers. They reported that the training was clear, the teaching method was uncomplicated, and their questions were answered quickly. C/U resource room teachers also noted that they liked that the training was not just theory-based, however they would also like training on the specific difficulties of delivering LRF!, such as invented words.

Despite this positivity about the training, the survey found teachers disagreed over whether the training was an appropriate length. Two in five W/C teachers (42%) and a quarter of C/U resource room teachers (27%) reported the training was too short, while one in ten W/C teachers (10%) and one in eight C/U resource room teachers (13%) stated it was too long. Also, in FGDs, teachers from both intervention arms explained that the length of the training session was insufficient to cover all the material. They reflected that, because of the limited training, they felt underprepared to deliver LRF!. For example, they did not always know how to handle questions about invented words.

During FGDs, coaches also reflected on the quality of the training. They perceived the training to have gone less well for C/U resource room teachers. They attributed this to two key reasons:

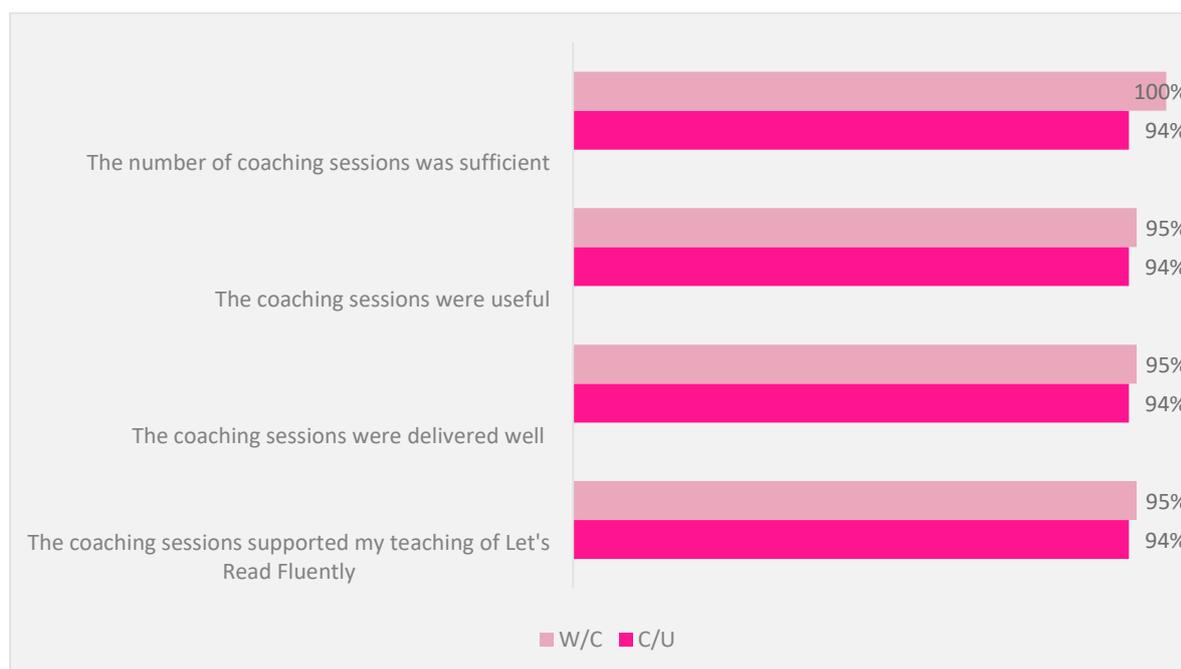
- **Additional effort to administer LRF!.** C/U resource room teachers need to apply additional effort when administering LRF!, as they are working solely with struggling pupils, and so they find implementing LRF! more burdensome. Coaches perceived that this resulted in a reluctance and defeatist attitude from resource room teachers during the training.
- **Greater experience.** W/C classroom teachers were more experienced in delivering different teaching methodologies and had a higher number of years' experience, and so were better prepared for the training than C/U resource room teachers.

Coaching

The majority of teachers surveyed across both intervention arms were positive about the coaching sessions. At least nine in ten teachers (94% C/U and 95% W/C) agreed the coaching sessions were useful, delivered well, and supported their teaching. Almost all teachers (94% C/U and 100% W/C) agreed that the number of coaching sessions was sufficient (see Figure 25).



Figure 25: Perceptions of LRF! Coaching – Agreement



Source: Teacher survey, W/C n=20 and C/U n=16

FGDs were also positive, with both C/U resource room teachers and W/C classroom teachers reporting the sessions were well-timed and at the right frequency. This differs from what coaches reported in FGDs. Coaches reflected that the timeline to deliver the three coaching sessions was too short and so teachers were unable to report any progress. They acknowledged that this was partly due to COVID-19 reducing the length of the school semesters. Coaches also expressed concerns over teacher enthusiasm and engagement with the LRF! intervention between coaching sessions.

Teachers described the coaches as excellent, patient, and relaxed. Teachers appreciated receiving feedback from them, as it gave them practical support with delivery. For example, coaches suggested teaching lessons over a couple of sessions to ensure pupils had learnt the page before moving on. C/U resource room teachers also reflected that the coaching helped them to translate the LRF! training into practical delivery.

In FGDs, coaches reflected on the relationship with teachers. They described how teachers having the same coach enabled coaches to monitor progress and provide continuity of support. However, coaches witnessed difficulties with some male teachers. They described how male teachers were reluctant to communicate via WhatsApp groups or receive coaching sessions. They believed this was because not all male teachers felt comfortable communicating with female teachers or coaches.

Diagnostic tool

FGDs showed that coaches and C/U classroom teachers had mixed opinions about the effectiveness of the diagnostic tool. They stated the tool was appropriate for pupils' abilities and made accurate assessments of the pupils' capabilities. However, a contrasting view was that the tool did not identify the right pupils for the C/U intervention, as it did not consider those pupils who might struggle with literacy for non-learning related reasons, like undiagnosed eyesight problems. C/U resource room teachers also reported that the diagnostic tool did not accurately



capture performance, as there was variability between C/U pupils. For example, you may have a pupil who can read words and another who can only read five letters.

Furthermore, coaches and C/U classroom teachers noted that the tool identified too many pupils as struggling with literacy. For instance, one C/U classroom teacher reported that 18 of their 19 pupils were selected. They attributed this to the COVID-19 disruption to schooling, with pupils having missed in-person schooling, which resulted in many pupils having a weak foundational knowledge. Coaches and C/U classroom teachers suggested that this gave an inaccurate assessment and led to some pupils being removed from the C/U intervention at a later date.

“It is impossible to begin with G1 [Grade 1] pupils who move to G2, after school interruption [due to COVID-19], without finding out how many letters they learned. I don’t test them on what they didn’t learn, it isn’t fair. What kind of assessment that tests their ability for something they weren’t exposed to?” - Coach

Additionally, C/U classroom teachers reflected that due to the COVID-19 school disruption, some pupils, particularly those in Grade 1, were nervous during the diagnostic assessment as they were unfamiliar with the school environment. These teachers believed this led to certain pupils underperforming, and the tool inadvertently identifying the wrong pupils. This was reported alongside other practical barriers to administering the test, like pupils copying answers or needing to complete the test in an individual setting. Subsequently, C/U classroom teachers adapted the tool’s implementation (see Adaptations) and suggested they would be in better position to select pupils for the intervention because of these adaptations.

Delivery

LRF! Content

In FGDs, teachers had mixed perceptions of the ‘I do, We do, You do’ process of LRF!. C/U resource room teachers like the repetition element, as it helped pupils become familiar with letters. However, C/U classroom teachers were critical of this, suggesting pupils were simply repeating rather than learning to recognise letters. Also, parents stated that pupils found the repetitive element tedious.

W/C parents offered contrasting opinions on the use of hand gestures. They suggested the gestures distracted pupils, while another view, which C/U pupils also reported, was that the hand gestures made it easier to remember words.

Across both intervention arms, teachers struggled most with implementing the ‘You do’ stage of LRF!. This stage involves independently working through the practice book, with the teachers providing feedback and encouragement. This is a key feature of LRF!, as the principles of the intervention suggest that pupils need to independently and repeatedly practice decoding to develop the automaticity needed for fluent reading. However, W/C teachers explained that it was not feasible to observe the reading of all pupils in their section, and C/U resource room teachers stated it was too challenging for C/U pupils to read independently, with many pupils feeling frustrated. This resulted in teachers adapting delivery (see Adaptations). C/U pupils also explained that they found this section the most challenging and preferred the choral reading.



Moreover, coaches, C/U resource room and W/C teachers thought the volume of content was too much to cover. Teachers acknowledged that, while coaches advised them the lessons could take place over a couple of sessions, they sensed it would take around three or four sessions to cover a single lesson.

Parents, coaches and teachers perceived LRF! to be too difficult for pupils in Grade 1 or those struggling with literacy. They consistently attributed this to pupils being unable to recognise letters and so subsequently finding it impossible to read words or sentences.

“By the end of the term, G1 [Grade 1] pupils are expected to have learned a sentence made of two-three words, mostly short syllables or with vowels, according to the national curriculum [...] Why is there a sentence, in the book ‘Let’s Read Fluently!’, made of eight or nine words?” - Coach

C/U resource room teachers commented that the pace of content was too fast for struggling pupils, as they were unable to finish more than a single page of content in each session. C/U resource room teachers also stated that they were unable to adapt the pace of the content to match the needs of the pupils and still complete the syllabus within the curriculum. This was particularly the case for pupils who had missed sessions, as there was limited time to recap previous sessions. W/C classroom teachers had an added challenge, as the size of their sections meant they were unable to evaluate all pupils in each session.

“You make an effort to read four long pages, that you can’t finish, and you have no time to go around each pupil to evaluate them individually. It will be a random reading of such letter, word or sentence. I have to go to the next pupil who may be struggling” - W/C classroom teacher

Furthermore, parents and teachers were divided about who LRF! is most suitable for. Parents and W/C classroom teachers said LRF! could be suitable for pupils in Grades 2/3, as they knew letters and could combine them to read words. However, C/U classroom teachers thought even high-achieving Grade 2 pupils would find LRF! too difficult. Additionally, coaches acknowledged that there should be different books available by grade.

In contrast, in the teacher survey, the majority (80%) of W/C teachers either agreed or strongly agreed that LRF! was easy to implement. This was similar for C/U, where, in both semesters, at least five of the eight teachers agreed it was easy to implement (63% Grades 2/3 and 75% Grade 1).

LRF! practice book

Sixty-five percent of W/C teachers surveyed agreed the practice book was a useful aid in the classroom, with 20% disagreeing with this. This was similar to what C/U resource room teachers reported for their Grade 1 pupils (63%). C/U resource room teachers were more positive about the practice book for their Grades 2/3 pupils, with 100% agreeing that it was a useful aid. However, from the FGDs it is not clear why C/U resource room teachers were more favourable about the practice book in the first semester. C/U resource room teachers stated that they would continue to use the LRF! practice book after the pilot study, but they would adapt the delivery. No information was provided on what adaptations they would make.

Generally, survey findings showed W/C and C/U pupils were positive about reading or working from the practice book both at school (95% for W/C and 79% for C/U) and at home (93% for W/C



and 66% for C/U). However, C/U pupils were less likely than W/C pupils to feel favourably about reading or working from the practice book. Furthermore, while there is a small decline in positivity between school and home for W/C pupils (2 percentage points), the decline is larger for C/U pupils (13 percentage points).

FGDs with respondents involved in C/U and W/C model of intervention were less positive about the practice book and focused on how it could be improved.

Parents in both intervention arms were critical of the physical size of the practice book. They perceived the practice book's large size to be off-putting for pupils, reporting their children were initially enthusiastic about the intervention but were discouraged by the large practice book. Parents also commented that the practice book was too heavy for pupils. Similarly, W/C teachers found the size of the large-print practice book difficult to work with. They stated the large-print practice book was too heavy to carry and too big to balance on the board or wall. This resulted in both parents and teachers suggesting the practice book be divided into multiple smaller books.

The FGDs also covered how the content in the practice book (see Introduction) was presented. Key concerns included:

- **The page content was too dense.** Parents from both intervention arms and C/U classroom and resource room teachers stated there were too many words on each page of the practice book. Coaches suggested additional spacing between words. Additionally, parents stated the words were too similar, which made it confusing and difficult to read.
- **Unclear text and typos.** W/C parents and teachers outlined examples of unclear text and typos: incorrect diacritic marks on the letter 't' in 'Tatloo', absence of the sukoon mark on 'food', and blurred letter 'm'. Although, W/C and C/U parents and resource room teachers described the font as large and clear. Coaches noted that Arabic letters go above and below the line, but the LRF! practice book only has letters above the line. They perceived this to be confusing for pupils.

There were also contrasting views on whether the practice book should have **coloured font**. Coaches and W/C parents criticised the black font, with coaches suggesting colored pages can help pupils with sight problems. This differed from C/U resource room teachers who thought coloured text would distract pupils. Pupils did not report any issues with the font colour.

There were similar contrasting views on whether the practice book should contain **pictures**. Pupils described the practice book as 'boring', and W/C parents compared the practice book to an instruction manual without colors and images. These parents, C/U parents, and teachers from both intervention arms, suggested pictures would aid pupils to remember words and better engage them with the intervention. However, in line with LRF! principles (see Introduction) other C/U classroom teachers and resource room teachers believed pupils would simply guess words from the pictures rather than learning to read.

The inclusion of **invented words** (see Introduction) in the practice book was contentious. C/U resource room teachers, W/C parents and W/C teachers described how pupils would enquire about the meaning of invented words. This created an additional burden for teachers when teaching LRF! and made it harder for pupils, as they could not conceptualise the meaning of invented words. Parents stated this demotivated their children, as they struggled most with the invented words. Moreover, parents noted that by teaching invented words instead of real words,



there was a missed opportunity to extend pupils' vocabulary. Conversely, another group of W/C parents and C/U classroom teachers perceived the invented words to be useful, as they consisted of three-four syllables and trained pupils to pronounce letters quickly.

Across both intervention arms, participants thought the practice book was too difficult for pupils and cited various reasons, including:

- **The difficulty of words in the practice book.** In both intervention arms, coaches, teachers, and parents stated pupils struggled to read and pronounce words in the practice book, as it did not account for the Arabic language sound system. Coaches gave the specific examples of 'mustashfa' and 'madrassa' and explained they are more challenging because they include consonants.
- **The progression from letters, to syllables, to words.** C/U resource room teachers and parents thought the practice book progressed too quickly from letters, to syllable, to words. They perceived this to negatively affect pupils' ability to combine letters and use diacritics. C/U resource room teachers stated words and sentences should only be taught once pupils know letters.
- **Inclusion of distinct diacritical marks.** W/C and C/U parents liked that LRF! included distinct diacritical marks, as it is not part of the curriculum books. However, C/U classroom teachers were less favourable, as they explained it created an additional challenge for struggling pupils and contributed to LRF! having too much content.

C/U and W/C pupils were divided on whether the practice book was too difficult or easy, although W/C pupils said their Arabic book, used in non-LRF! lessons, was easier. W/C parents also reported the practice book was appropriate for their children's abilities, although they did not provide reasons why.

Coaches and C/U resource room teachers explained how, overall, they did not feel the book was suited to resource room delivery. They explained how the aim of the resource room is to provide tailored support to pupils, but the LRF! practice book is a 'one-size fits all' approach. They noted that different books for different abilities would be a solution to this.

Contextual factors

In FGDs, participants outlined several contextual factors that affected delivery of LRF! These are grouped into two categories, school-based factors, which affected delivery for all pupils within a school, and pupil-based factors, which affected delivery for individual pupils.

School-based factors

- **Staffing challenges.** Coaches reported staffing challenges in schools. These included absences of LRF! delivery teachers through ill health, and changes in school leadership. Coaches reflected how the changes in school leadership resulted in new senior school staff having limited understanding of LRF! and this disrupted the school's support for teachers delivering LRF!.
- **Quality of teaching.** Coaches reflected how the competencies and experience of teachers varied considerably. They explained how some were more experienced and better prepared to deliver LRF!, while others struggled with delivery; for example, reading quietly



or making mistakes when reading. Coaches explained that W/C teachers were stronger at LRF! delivery, as they have more experience teaching lower grades.

- **Large section size.** W/C teachers outlined how having large sections made implementation difficult, particularly in the ‘You do’ section (see Dosage). Coaches disagreed with this and stated that they did not think class size was a problem, but that teachers and teaching methods were the determining factor.
- **Challenging national curriculum.** Parents and classroom teachers, for both W/C and C/U, described the national curriculum as challenging in terms of content and time needed for delivery. This made it difficult for teachers to conduct LRF! lessons alongside the national curriculum, and so some C/U classroom teachers sought additional support from the resource room teachers to cover curriculum content. Similarly, parents perceived national curriculum homework to be burdensome for pupils and so were reluctant to take on the extra burden of LRF!.

Pupil-based factors

- **Pupil absenteeism.** Coaches and teachers reported pupil absenteeism as a barrier to LRF! delivery. W/C and C/U classroom teachers noted local sickness outbreaks, such as chickenpox, meant high levels of pupil absence. Additionally, there were seasonal absences. Coaches stated some pupils did not attend school in cold weather and C/U resource room teachers explained how families working in sheep farming removed their children from school for up to three months.
- **Maternal illiteracy.** C/U resource room teachers noted maternal illiteracy was a common barrier for parental engagement with LRF!.
- **Negative perceptions of the resource room.** Coaches and C/U resource room teachers reported that parents held negative perceptions of the resource room, often seeing it as a place that signifies poor attainment. For this reason, C/U parents were reluctant for their children to engage with LRF!.
 - ***“I was concerned that my daughter goes to the resource room...those who go there are not clever”*** – Parent

Furthermore, the pilot study was conducted during the COVID-19 pandemic.⁴¹ This caused added disruption to intervention delivery.

- **Missed kindergarten education.** Due to school closures prior to the pilot study, Grade 1 pupils had missed their kindergarten education. Teacher and parents for both intervention arms explained how this made LRF! additionally challenging for Grade 1 pupils as they lacked the foundational skills and knowledge.
- **Attitude towards school.** Parents reported school closures in 2020 and 2021 had negatively affected pupils’ attitudes towards school. For example, pupils were reluctant to attend school or complete homework.

⁴¹ The pilot of the LRF! program was conducted between September 2021 and May 2022. During this time, schools closed for an extended winter break due to the outbreak of COVID-19 cases. Additionally, due to the pandemic, schools were closed for large parts of 2020 and 2021.



Adaptation

RQ2: How do principals, teachers, parents and pupils perceive the intervention and any changes that it has delivered?

RQ4: Is there any evidence of unintended consequences (negative or positive) as a result of the implementation of LRF!?

RQ5: Was LRF! delivered as intended in terms of dosage, nature and quality? What modifications were made, with what implications?

RQ6: What is the learning about teacher's use of the Coarse-Grained Diagnostic (RAMP) tool? How successful is, it in use, at identifying the most appropriate pupils for the Literacy C/U implementation model (C/U model only)?

RQ8. To what extent do teachers develop sufficient skills and confidence through the training and coaching?

RQ9: What do we know about how teachers need to be supported (coached) during delivery?

RQ10: Are there any key contextual factors that appear to facilitate or impede successful implementation of LRF!?

This section covers the reported adaptations from the FGDs. While these adaptations demonstrate how implementation deviated from the intended design, these adaptations did not occur in all school settings.

Diagnostic Tool

The diagnostic tool was applied inconsistently across schools. Adaptations were driven by:

- **Pupil ability.** C/U classroom teachers described pupils having difficulty completing the assessment on their own, with some pupils finding the assessment confusing. These teachers adapted the implementation by administering the tool individually with pupils, rather than having the whole class complete the assessment concurrently.
- **Section size.** C/U classroom teachers described how in larger sections pupils would copy each other's answers. This meant the diagnostic scores did not necessarily represent ability and so in some cases C/U classroom teachers selected pupils for the C/U intervention based on their own judgement instead.
- **Teachers' knowledge of pupils.** C/U classroom teachers stated they selected pupils for the C/U intervention, as they felt their personal judgement of pupils' reading ability was more accurate than the results of the diagnostic tool.
 - ***"We have also taught them for a whole term, so we know those who learned the letters and those who didn't."*** - C/U Classroom teacher

In other instances, coaches reported C/U classroom teachers had exchanged notes with resource room teachers to explain why certain pupils should not participate in the intervention based on the 'full picture' of their academic performance.



Delivery

As covered in the dosage section, schools struggled to cover all elements of the lesson content within the designated number of sessions. This resulted in teachers adapting delivery to cover LRF! content. These adaptations included:

- **Reducing content.** C/U resource room teachers explained they did not recap the previous lesson at the start of LRF! sessions. They believed this made it particularly challenging for pupils who had been absent for the previous lesson. This finding contrasts with observation data (see Reach and Dosage).
- **Adjusting pace.** W/C classroom teachers explained they taught at a slower pace and ensured pupils were of an adequate standard before proceeding to the next lesson. They explained this resulted in added pressure to complete all the LRF! content.
- **Lengthening sessions.** W/C classroom teachers lengthened LRF! sessions and reduced the length of other curriculum lessons.

Teachers across both intervention arms also deviated from the 'I do', 'we do', 'you do' pedagogical approach. These deviations were driven by:

- **Pupils' reading ability.** C/U resource room teachers reported pupils having difficulty reading independently during the 'you do' section. This led resource room teachers to individually coach pupils to sound out letters, as they were not able to read the words without this support.
- **Class size.** For the W/C model, the number of pupils in each section made it difficult for the classroom teacher to hear pupils and assess whether they needed support. Because of this, W/C teachers suggested that they would have to go around the class to hear each pupil practice individually, however time did not always allow for this.

Resource room teachers in the C/U intervention arm also adapted delivery to improve pupil engagement with the intervention. They described how they incentivised pupils to complete tasks by playing games or giving pupils rewards. For example, they would ask pupils to find a word in the practice book and give out stars or stickers to the pupil who was able to do so the fastest.



RESPONSIVENESS

RQ2: How do principals, teachers, parents and pupils perceive the intervention and any changes that it has delivered?

RQ5: Was LRF! delivered as intended in terms of dosage, nature and quality? What modifications were made, with what implications?

RQ7: What were the facilitators and barriers to engagement in the teacher training, teacher coaching and Supervisor and Principal orientation sessions?

This section discusses teachers, pupils, parents and coaches' engagement with the LRF! intervention.

Training

As outlined earlier (see Quality), FGDs showed coaches perceived the training to go less well for C/U resource room teachers. Coaches thought the training added challenges for resource room teachers and resulted in the teachers having a reluctant and defeatist attitude.

Moreover, teachers from both arms of the intervention identified two key barriers to effective training: distance and duration. Teachers reflected how the training day required a very early start to travel to the venue, which they found tiring. W/C teachers stated the training should have been in closer proximity to, or even in, their schools, rather than requiring travel. Additionally, teachers suggested a longer training session would have been preferable as they needed more time to cover the practice book.

"I think it should have been longer than one day [...] When we began applying the program, we had many question[s] and queries. We couldn't learn everything on the same day or session." – Resource room teacher

Coaching

Facilitators for engaging coaches

Coaches described **relationship-building** as a key facilitator to their work with teachers. This included engaging directly with teachers but also building effective working relationships with principals. This is because principals hold influence in terms of motivating their staff team to engage with the intervention and the coach:

"If principals have a strong presence, and understand the program and its objective, they can support me by sitting with teachers to explain to them and facilitate our relations with teachers. [...] When principals know the full picture, they support us, and teachers observe what their principals say to them." - Coach

Clear lines of communication were also important for coaches to be able to fulfil their role effectively. An example where this did not happen was when a coach travelled to a school to find the pupils absent. The coach noted they did not receive any advanced communication from the school, and this had led to a wasted journey.

Teachers did not identify barriers to engaging with coaching.



Facilitators and barriers for engaging teachers

Coaches described three barriers to teacher's engagement, a constrained timeframe, culture of scrutiny, and a reluctance to deviate from usual practice:

A constrained coaching timeframe, resulting from a shift to the intervention timetable, meant coaching sessions were closer together. Coaches suggested this reduced their effectiveness, as teachers would not have sufficient gaps between to notice pupils' progress (see Dosage).

A culture of scrutiny, inspection and judgement resulted in a persistent perception among teachers that coaches were inspecting and monitoring their work, rather than assessing the intervention method. Coaches believed teachers saw their observations as similar to inspections from the ministry, and this perception created tension and stress for teachers. W/C teachers discussed rushing to complete lessons ahead of coaches' visits, adding strain to their wider workload.

A reluctance among teachers to deviate from their usual teaching style and approaches. For example, calling only on specific children to read rather than giving all a turn. Coaches described this as a 'mindset that can't be changed'; this made it challenging to engage teachers with the coaching sessions.

Delivery

LRF! Content

The survey findings showed a high proportion of teachers for W/C (65%) and C/U Grades 2/3 (75%), who thought pupils were engaged with LRF!. However, for C/U Grade 1, only one in four (25%) of resource room teachers thought pupils were engaged.

The teacher survey showed that engagement with parents was high for both the W/C and C/U models. Across both intervention arms, only one C/U resource room teacher stated they were not engaging with parents. The most popular method of communication was WhatsApp (at least 87% across both intervention arms), with most teachers using WhatsApp a few times a week (at least 87% for both W/C and C/U). Additionally, except for a single teacher, both C/U resource room and W/C teachers commented that parents were always aware if pupils were taking home the practice book.

Across the lesson

Across both intervention arms, almost all pupils had the practice book open at the right page during observed lessons. During these lessons, nearly all teachers encouraged pupil participation, responded to pupils' questions, and showed positivity when acknowledging pupils. This suggests high engagement from teachers during the observed lessons.

I do

During observations, the engagement of pupils in the 'I do' section was high for both W/C and C/U. In almost all observations, pupils listened to the activities and looked at the teacher and the large print letters. However, some pupils were less engaged in the 'I do' section. Around a third of observations recorded some pupils being distracted or visibly frustrated. A small number of observations recorded disruptive behavior from some pupils during the 'I do' section. These observations were similar for W/C and C/U.



We do

For the 'We do' sections pupil engagement was also high. In the observed lessons, as with the 'I do' section, pupils watched and listened to the teacher and followed the teacher's instructions. Additionally, pupils verbalised along with the teacher; however, over half of observations for W/C and C/U recorded pupils verbalising mainly by copying their peers.

As with the 'I do' section, some observations found a group of pupils less engaged. For example, around a third of observations for both W/C and C/U noted some pupils did not verbalise in the 'We do' section. Additionally, some observations found pupils haphazardly repeating the letter sound. A small number of pupils were visibly frustrated or disruptive during the observations.

You do

Observations found similar levels of high engagement among W/C and C/U pupils with the 'You do' section. In almost all observations, pupils looked at the practice book, applying themselves to the task for at least one continuous minute, and following the text with their finger as they read.

Not many observations found pupils asking the teacher questions, although most observations showed pupils listening to the teacher's instructions and feedback.

While still a minority, a larger number of observations noted W/C and C/U pupils being distracted, disruptive, or showing frustration during the 'You do' section. As previously noted, FGDs indicated that teachers and pupils found this the most challenging aspect of LRF! (see Quality).

Teacher engagement

In FGDs, teachers from both intervention arms noted they had limited 'buy-in' to the LRF! method. They stated this was because they preferred alternative pedagogical approaches, for example, developing writing skills alongside reading to reinforce recall. Teachers specifically commented they did not like the use of invented words in the practice book.

Pupil engagement

FGDs revealed mixed views on pupil engagement with LRF!. Teachers from both intervention arms reported that pupils liked the LRF! method; this mirrored the views of pupils who explained how much they liked the practice book and the choral reading. However, other teachers disagreed with this. They thought pupils were disengaged and bored, evidenced by them daydreaming or requesting to leave the room to go to the toilet. Pupils also commented that the practice book was boring, and parents noted their children had been reluctant or unwilling to read the practice book.

Coaches, teachers, and parents indicated that teachers played a key role in motivating pupils to participate in and outside class; for example, by encouraging a pupil's sense of pride about developing their literacy. This was the case for both the W/C and C/U models. Additionally, C/U parents and teachers explained how the resource room enables teachers to give dedicated focus to struggling pupils, and this in turn can motivate pupils. They also reflected how the different syllabus in the resource room can relieve pressure off pupils, as they are not concerned about their grades.

Parents, teachers, and coaches provided various reasons for why pupils might not engage with LRF!. Some of these related specifically to LRF!:



- **Complexity.** As outlined already, (see Quality), all participant groups perceived LRF! as particularly difficult for pupils in Grade 1 and those struggling with literacy. Parents and teachers, for both W/C and C/U, perceived the complexity to be a barrier to pupils engaging with LRF!.
- **Practice book.** Parents and W/C teachers stated pupils found the practice book ‘dense’ and intimidating.
- **Repetition.** Coaches, W/C teachers, and parents noted the repetitive content could induce boredom. Coaches witnessed pupils’ focus diminishing when the teacher read the same content for a third time. This differs from other findings (see Usual practice) which indicated repetition was useful for learning.
- **Use of invented words.** Parents and teachers in both intervention arms believed that the use of invented words discourages pupils. They thought that pupils struggled most with these words, and some would just laugh at them.
- **Presentation of content in the practice book.** As outlined previously (see Quality), some teachers and parents perceived the lack of colours and pictures in the practice book to be disengaging for pupils. However, there were contrasting views from coaches, parents and teachers who thought pictures and colours would be distracting.

There were also contextual factors that hindered pupils’ engagement with the intervention.

- **Individual characteristics.** Teachers and coaches regarded pupils’ individual characteristics as a barrier to engagement. For example, if a pupil is shy, they may not take part in the ‘We do’ or ‘You do’ sections, or if a pupil has eyesight difficulties, they might struggle to participate and engage.
- **Teaching style and skill.** Where teachers were less experienced or skilled at reading themselves, coaches suggested this negatively impacted on class engagement with the lesson.
- **Negative perceptions of the resource room.** C/U resource room teachers mentioned a reluctance among some pupils and parents for their children to attend the resource room (see Quality for more information).

Parental engagement

Engagement with the school

Teachers and coaches explained how schools informed parents about LRF! either via WhatsApp or at an introductory information evening at the start of the semester. This included sharing information on the aims and intended outcomes of LRF!, and the use of invented words, diacritic marks, letter combination and nunation.⁴² Parents commented that they regularly received WhatsApp messages from teachers on information and resources (such as page numbers and explanatory videos) to support home learning. W/C teachers commented that they supplied resources in accessible formats, for example video or voice recordings, to support parents’ involvement.

Despite the high level of communication between schools and parents, coaches and teachers outlined challenges with parental communications. These included:

⁴² Unvocalised suffix -n, which is pronounced but not written at the end of nouns.



- **Digital exclusion.** C/U resource room teachers reported not all parents had access to the internet or electronic devices.
- **Illiteracy.** Teachers noted some parents, more often mothers, were illiterate. This made it difficult to communicate with parents and limited parents' ability to support their children.
- **Gender.** Coaches described difficulties in communication between some male staff and mothers. They believed this was because not all male staff felt comfortable speaking to mothers. This meant male teachers needed to find a way to engage with fathers. Similarly, in rural communities, mothers would not communicate with their son's all-male school.

Furthermore, coaches and teachers in both intervention arms stated parental engagement with schools was very limited, with only a small number of parents interacting with teachers. This included attending the introductory information evening. One view among teachers in both intervention arms was the parents who engaged with school communications tended to be the same as those who supported children with wider education. Findings from parent FGDs also suggested mixed engagement, with variation between W/C parents on whether they had asked the school questions about the intervention.

Engagement at home

Parents described a range of approaches to support their children to read at home. These can be grouped into three categories:

- **No awareness.** Parents said they were unaware of the practice book. In some instances, this was because their spouse took responsibility for supporting children's home learning.
- **Aware but disengaged.** Parents were aware of the intervention, practice book, and school communications, but had not followed up with their children about this.
- **Actively engaged.** Parents regularly supported their children using the practice book in a variety of ways. For example, taking on a supervisory role, listening to their children read rather than offering any guidance, or actively teaching their children the practice book content.

Coaches described parents' lack of engagement as 'creating considerable problems' for delivery. Coaches and teachers in both intervention arms viewed parental follow-up at home as an important part of the intervention, noting that children's reading improved when parents followed up. Four aspects facilitated at home support.

Capacity. Parents stated they needed both time and skills to support their children. For example, a father reported he had limited time or energy after work and so his wife took responsibility for their child's schoolwork. Coaches and teachers noted that parents who had unstable living arrangements had limited capacity to support their children. This included parents who moved a lot for seasonal work, worked long hours, or were illiterate.

Sense of responsibility. Some parents explained how they should contribute to their child's learning. They acknowledged the responsibility for their children's learning should not exclusively be with the teacher or school.

"They [pupils] need help. Without it, they won't respond or study. When he [son] comes home, I ask him what he learned. If he doesn't find anyone checking on him, he will forget, and the next day, he will forget what he learned." – C/U Parent



Other parents disagreed with this sentiment and perceived it to be the school's responsibility. C/U classroom teachers reported this too, noting that some parents expected children to learn everything from teachers.

Perceived value of the LRF! intervention. Parents commented that the homework needed to be worthwhile, otherwise they would not support their children with it. This aligns with W/C teacher perceptions that parents focused on examined work and saw LRF! as an extracurricular activity (see Usual Practice). Some parents preferred alternative literacy approaches (like dictation and handwriting) and so focused on these with their children. No information was provided on why these parents preferred alternative practices.

Pupil enthusiasm. Parents believed their ability to support their children depended on their children's enthusiasm. For example, parents commented that it was easier to support those children who enjoy school and homework. Parents whose children did not enjoy reading or completing homework reported avoiding 'pressuring' their children.

USUAL PRACTICE/INTERVENTION DIFFERENTIATION

RQ1: 1. In what ways, and to what extent, does LRF! affect school, teacher, pupil, and parental practice as compared to business-as-usual teaching and learning?

This section explores the extent to which LRF! activities differ from other existing practices associated with learning to read.

LRF! practice book versus the curriculum book

In FGDs, participants made comparisons between the LRF! practice book and the Arabic curriculum book.

Participants reported that the practice book and the Arabic curriculum book differ on the order in which letters are introduced. For this reason, W/C teachers and parents felt the two books did not align, which made it difficult for LRF! to reinforce curriculum learning. Similarly, coaches and C/U classroom teachers thought the difference in letter order between the LRF! practice book and the national curriculum book was particularly problematic for Grade 1 pupils who are still learning to recognise letters. They explained the different sequential order made it challenging to reinforce learning and coaches found it challenging for teachers to teach.

Moreover, C/U classroom teachers explained that pupils were confused by learning different letters in the resource room and the main classroom. They also noted how the technique of breaking up syllables and teaching phonetics differed, which created additional confusion for pupils. However, C/U resource room teachers disagreed with this and thought the different content and techniques aided pupil engagement.

Additionally, parents, pupils, teachers, and coaches considered the LRF! practice book to be more challenging than the Arabic curriculum. Parents suggested that the practice book be used as a source of additional material; for example, for those Grade 1 pupils particularly strong in literacy, or as an additional stage between Grade 1 and Grade 2, as this is where pupils begin to read paragraphs and so need to recognise all letters. When reflecting on whether the two books should be merged, W/C pupils responded that the two books should be separate because the national Arabic book was easier. In contrast, W/C parents thought the books should be merged, so pupils would learn from a single book.



Business-as-usual teaching practice

FGDs identified several techniques as business-as-usual teaching practice in control schools.

Repetition is a technique used in business-as-usual teaching and LRF!. In business-as-usual teaching for example, control classroom teachers described how they would begin by saying a letter or word aloud. This would then be repeated by their pupils (first with the teacher and then without), and then the pupils would repeat the letter or word individually. This is very similar to the 'I do, We do, You do' structure. The same was noted for less able pupils, as C/U classroom teachers described how they typically introduced a concept to pupils before instructing them and then leaving them to work independently. Control classroom teachers used repetition exercises because they saw it as a key technique and a key part of the RAMP method and the MoE approach.

Control schools used competitive techniques to engage pupils in reading. These included getting pupils to try and beat the length of time it takes to read a section of text or by getting pupils to compete against each other. Examples of this included 'Revolving Bird', where the teacher will begin by reading and then pupils take it in turn to follow on. Pupils are encouraged to follow the text and be ready to start reading when it is their turn. Control classroom teachers saw this as an effective way of engaging all pupils; it ensured they focused on the lesson while competing with their classmates. Like the repetitive techniques, control classroom teachers associated these practices with RAMP.

Control schools taught pupils how to write or formulate the letter shapes to help them learn how to read. They used a variety of materials, including pencils, paste and sticks, or threading beads. Control teachers and parents perceived learning to write as an important part of learning to read, as it reinforces learning. Control schools would often take a staggered approach to introducing literacy concepts, for example by beginning with letters, followed by syllables. When pupils recognise the letters and their combinations, they would then move on to reading a line or two. Some intervention schools also did this as standard practice. A C/U classroom teacher taught from a book called Elbonyaan. It starts with letters and the 'fatha' (short "a"), then it moves on to 'damma' (short "uu"), before eventually covering three letter words.

Reading stories and improving reading comprehension was another business-as-usual technique. Control classroom teachers believed it helped pupils develop a sense of curiosity and imagination. They would also give specific exercises based on stories, such as asking pupils to read a story from a book and then asking them questions about it.

Control teachers and parents saw positive reinforcement as essential when teaching pupils to read. Control teachers regarded this as a way to encourage pupils to continue with learning to read. Often teachers and control parents used rewards as positive reinforcement, such as teachers giving pupils badges, certificates, and sharing photos with parents.

“When they read a paragraph, I say to them...‘You are today the exemplary pupil’, ‘You are a young teacher, and they need to learn from you’, or I am going to take a photo of you and send it to your mum, or post on the group’. This is a great way to encourage them.” - Control Teacher

Control parents motivated their children to complete their homework by rewarding them with gifts, money, or phone time.



Barriers to business-as-usual teaching

In FGDs, control teachers and parents outlined barriers to teaching pupils to read. These were similar to the barriers and contextual factors reported by intervention participants (see Quality).

- **Low pupil engagement.** Control teachers noted pupils struggling with literacy or speech could be disengaged with reading. Control parents also explained their children might not like reading or Arabic homework, and so it would be difficult to engage or motivate them.
- **Busy curriculum.** Control classroom teachers commented how the busy curriculum is a challenge. This is because it puts pressure on both parents and teachers on what pupils need to learn and what they should focus on. Moreover, it makes it difficult for control classroom teachers to have enough time to support struggling pupils. Control classroom teachers did not think additional resource would help with this challenge; they simply need more time in the day.
- **Section size.** Control classroom teachers mentioned the size of their section can impact their ability to interact with pupils one on one. They reflected how this increases the burden on the parents to support the teacher.

Struggling pupils

In FGDs, control teachers outlined the various ways they identify pupils with struggling literacy. For Grade 1 pupils, this included getting classroom teachers to identify pupils with no reading experience at the start of the school year, and then providing additional support to these pupils. Control classroom and resource room teachers reflected how, due to their limited schooling, it could be difficult identifying Grade 1 pupils who are struggling with literacy. They also stated that due to the inexperience of Grade 1 pupils, the standardised tests would be inappropriate.

For Grades 2/3, schools monitored pupil progress throughout the school year. For example, they ask pupils to read a storybook every month, evaluate their progress, and then tailor the support accordingly. Alternatively, control teachers complete annual paperwork covering the number of stories read by each pupil, what books they have taken home and when they are returned to school, and any information the pupil was able to recall about the story. They then use this to identify struggling pupils.

Control schools used different approaches to identify pupils for the resource room. This included pupils selected based on assessments, using general and special tools. Alternatively, schools asked the classroom teacher to create a shortlist of pupils (based on previous school grades and information from the pupil's parents), from which the resource room teacher would choose a few pupils. Generally, in control schools, those Grades 2/3 pupils identified with additional need attended the resource room for support with literacy. In FGDs, the number of pupils sent to the resource room varied between control schools, from only a single pupil up to four pupils.

Control schools adopted a wide range of approaches for their struggling pupils. Classroom teachers aimed to make lessons easier for struggling pupils by either shortening the lesson or using a curriculum designed for younger pupils. They also mentioned using a variety of resources with pupils, including workpapers, flashcards, and notebooks. In resource rooms, teaching techniques for struggling pupils were similar to the business-as-usual approaches. For example, they focused on repetitive learning, with the purpose of ingraining the standard Arabic curriculum content. Both control classroom and resource room teachers reflected that struggling pupils could see themselves as different from their peers, and so positive reinforcement and support is vital.



Control teachers had mixed reflections on the resource room. For C/U classroom teachers, they perceived resource room delivery to be a positive, as they found it challenging to teach struggling pupils with the same techniques as the other pupils in the section. They noted how the resource room purposefully adopts different activities and games to support struggling pupils in their learning. Furthermore, C/U classroom teachers described how struggling pupils could not learn at the same speed as their peers. So, the multiple resource room sessions per week, provided the needed additional time and support.

Despite these positive reflections, control teachers reflected on the downsides of resource room delivery. They explained the resource room lacked the equipment to support struggling pupils and pupils missed other lessons to attend the resource room.

Homework

Like LRF!, in control schools literacy homework consisted of assigned readings. Teachers gave a wide-range of examples, including: assigning pupils a set of words or books to read at home and then checking these had been completed, asking pupils to read a storybook and then draw a picture, giving writing or recitation exercises. Additionally, like the QR code in the LRF! practice book, some control teachers shared voice recordings of reading lessons with parents to assist them with the homework.

Control classroom teachers reported pupils would not complete homework or hand it in, and for this reason some teachers opted not to assign homework. Teachers saw poor parental engagement as a contributing factor for pupils not completing homework.

Parental engagement

In FGDs, control and intervention teachers described how parental engagement was integral for teaching pupils to read. They reflected how it reinforced learning and enabled pupils to have individual attention and focus, something not often possible in large sections. However, control parents did not believe they had a primary role as educators and believed teachers were more experienced and better suited to teaching pupils to read.

Despite these disagreements, control parents explained they often read with their children. They noted they use a range of exercises to support their children's literacy. Examples included, reading stories and using shadow puppets or asking their children to describe the pictures in story books with the diacritics. Also, they would get their children to read signs when walking or spell out words on their hands. Some control parents involved siblings when teaching their children to read, either by getting older siblings to help or by asking their child who is learning to read to teach younger siblings.

Control teachers regularly communicated with parents through videos, WhatsApp messages, and in-person at school. Control teachers noted parental communication via meetings and WhatsApp groups were vital for ensuring alignment between school and home teaching and giving parents instructions on what teaching methods to use.

Overall, control teachers thought that parents did not sufficiently support their children with learning to read. They saw this as a barrier to pupils' progress. Teachers perceived parental disengagement to be due to maternal illiteracy or parental focus on examination grades. This is similar to the findings for the intervention group (see Responsiveness). Control teachers also thought this was more common among struggling pupils. Some control schools have sought to



overcome this by running a course for parents on teaching methods at the start of the academic year.

Library usage

Across intervention arms, the proportion of pupils who stated they could take storybooks home from school varied substantially between baseline and endline surveys. This is potentially due to pupils not being certain about whether they could borrow books, and at endline, intervention pupils might have inaccurately recalled taking home the LRF! practice book. Generally, at the endline survey for both control and intervention schools, around four in ten pupils stated they could borrow books (46% W/C, 44% W/C control, 39% C/U, 33% C/U control). However, due to the variation in scores between baseline and endline, it is too difficult to make any further judgements.

Similarly, between baseline and endline surveys, there was considerable variation in the proportion of pupils who were asked the follow up question on taking books home from school⁴³. Therefore, it again makes it difficult to make any reasonable judgments from the data. The only summary findings are:

- For W/C control and intervention, the proportion of pupils taking storybooks home at least once a week increased between baseline and endline (21% at baseline to 39% at endline for W/C, 18% at baseline to 38% at endline for W/C control).
- For C/U control and intervention, there was an increase in the proportion of pupils taking storybooks home at least once a week between baseline and endline (0% at baseline to 20% at endline for C/U, 13% at baseline to 49% at endline for C/U control).

In FGDs, control teachers reported some schools had limited library access and so pupils were unable to borrow books. Additionally, control parents stated they did not take their children to libraries as they felt their children were too young for the books on offer.

Reading at home

According to the pupil survey, almost seven in ten W/C pupils (68%) read storybooks at home at least once a week. This compares to around half (50%) of the W/C control pupils and is up 20-percentage points on baseline. Typically, these pupils read for between 15-30 minutes (47%) and most read with their parents (65%), siblings (25%), or by themselves (37%). These results were similar to the W/C control group (43% read for 15-30 minutes, 71% with parents, 21% with siblings, or 29% by themselves).

Furthermore, the survey found an increase in the proportion of C/U pupils reading storybooks at home at least weekly between baseline and endline (17%). However, at endline, this only accounts for a third (33%) of C/U pupils, so far less than for W/C (68%). Like W/C pupils, most read with their parents (56%), siblings (22%), or by themselves (22%). However, C/U pupils typically read for less than 15 minutes (33%). Results for C/U control pupils were more similar to W/C pupils than the C/U intervention arm; 53% read storybooks at home at least weekly and generally for between 15-30 minutes (43%).

⁴³ The follow up question 'How often do you take books home from your school or a library?' was asked to those pupils who answered 'Yes' to 'Are you able to take story books home from your school or a library?'.



In FGDs, control parents described their children as being enthusiastic about reading. They evidenced this by explaining their children are excited to read new stories or wanted to read before bed.

Internet usage

In the endline survey, almost all pupils stated either they or their parents had access to an internet enabled device, most commonly a tablet (98% for W/C and 91% for C/U). Moreover, the majority of pupils reported having internet access at home (83% for W/C and 78% for C/U), and most of these pupils reported being allowed to use the internet (79% for W/C and 82% for C/U).

Pupils who have access to and can use the internet were asked if they read stories online. For W/C pupils, 60% did this, up 12 percentage points from the baseline. As with reading physical books, W/C pupils typically read online stories with parents (57%), by themselves (36%) or with siblings (27%). C/U pupils were less likely than those in the W/C arm to read stories online (37.73%). Additionally, there was a decline in the proportion of C/U pupils reading stories online between the baseline and endline surveys (-20 percentage points). Of those C/U pupils reading stories online at endline, as with other similar measures, most read online stories with parents (50%), by themselves (39%) or with siblings (33%).



PERCEIVED OUTCOMES

RQ3: Is there evidence to support the Logic Models?

This section explores the perceived outcomes of LRF! for pupils and teachers.

Pupils

Interim outcomes

The following reflections are from the observations that took place at two points in each semester. They cover what pupils were doing during the semester rather than pupil outcomes at the end of the LRF! intervention.

During observations, both W/C and C/U pupils were able to accurately verbalise the letter sound and diacritic during the 'We do' and 'You do' sections. Pupils also attempted to decode words.

In most W/C observations pupils were able to accurately read three letter words with the lesson letter and read two to three letter words that were not associated with the lesson letter. However, observations with C/U pupils showed they were more likely to struggle with this. Only around half of the observations saw C/U pupils reading the words.

W/C and C/U pupils found it challenging to read sentences, with only a small number of observations recording pupils able to do so. Additionally, in many observations, teachers did not cover the sentences in the lessons.

Literacy abilities

Just over half (55%) of W/C teachers surveyed are positive LRF! achieved the following pupil outcomes: improved reading comprehension, greater engagement in reading, and more confident in reading. Around two out of three (65%) W/C teachers thought LRF! improved pupils' literacy. For all three of these measures, disagreement that LRF! achieved these outcomes ranged from 10-20%. The remainder of W/C teachers (25-30%) neither agreed nor disagreed (see Figure 26 below)

The above findings correspond with the findings from the FGDs. W/C teachers and parents found pupils' reading ability and speed had improved following LRF! sessions. For example, W/C teachers noted pupils were now spelling out the letters and syllables within words, which in turn was improving their reading ability.

"I have some girls whose reading has improved and is now faster... and even go up to the board and read." - W/C classroom teacher

Parents argued the approach of spelling out letters and syllables had slowed their children's reading. They noted their children had previously read flowingly but were stuttering across the letters and syllables following LRF! sessions. In other instances, parents reported their children linking letters to form words following LRF! sessions. However, they also acknowledged that, due to the challenging LRF! content, their children were still struggling to link letters to form longer words. W/C parents also found their children had improved pronunciation of letters and words, better diacritics, and greater confidence when reading. They also reported their children had begun tracing or writing letters on their own.

However, despite the reported success in literacy, W/C teachers indicated not all pupils had seen improvements in literacy. They offered conflicting perspectives about who had benefitted most from the intervention. For example, it had an impact on the 'weaker' pupils and limited effects on



pupils with strong literacy abilities. Conversely, the ‘weaker’ pupils had struggled with LRF! and therefore the only improvements in literacy outcomes were for pupils with average grades or higher.⁴⁴

For C/U resource room teachers, the perceived outcomes were more mixed. For Grades 2/3, the survey found that C/U resource room teachers mostly agreed the following pupil outcomes had improved: literacy (88%), reading comprehension (75%), and confidence when reading (63%). Half of C/U resource room teachers (50%) agreed pupils in Grades 2/3 were more engaged in reading due to LRF!. Disagreement across all four of these measures was non-existent (0%). In FGDs, C/U resource room teachers stated they had observed improvements in Grade 3 pupil performance. Coaches also reflected that Grades 2/3 pupils were more confident, as the repetitive approach of LRF! had ingrained learning and so pupils were less nervous when reading (see Figure 26 below).

However, for Grade 1 C/U, resource room teachers were far less positive in their survey responses. Only a quarter (25%) agreed Grade 1 pupils were more engaged with reading, more confident when reading, and had improved reading comprehension. Just one in eight (13%) saw improvements in pupils’ literacy. Disagreement with these statements ranged from 50-63% (see Figure 26). In FGDs, C/U classroom teachers suggested LRF! had limited impact on literacy outcomes because the content was too challenging for Grade 1 pupils.

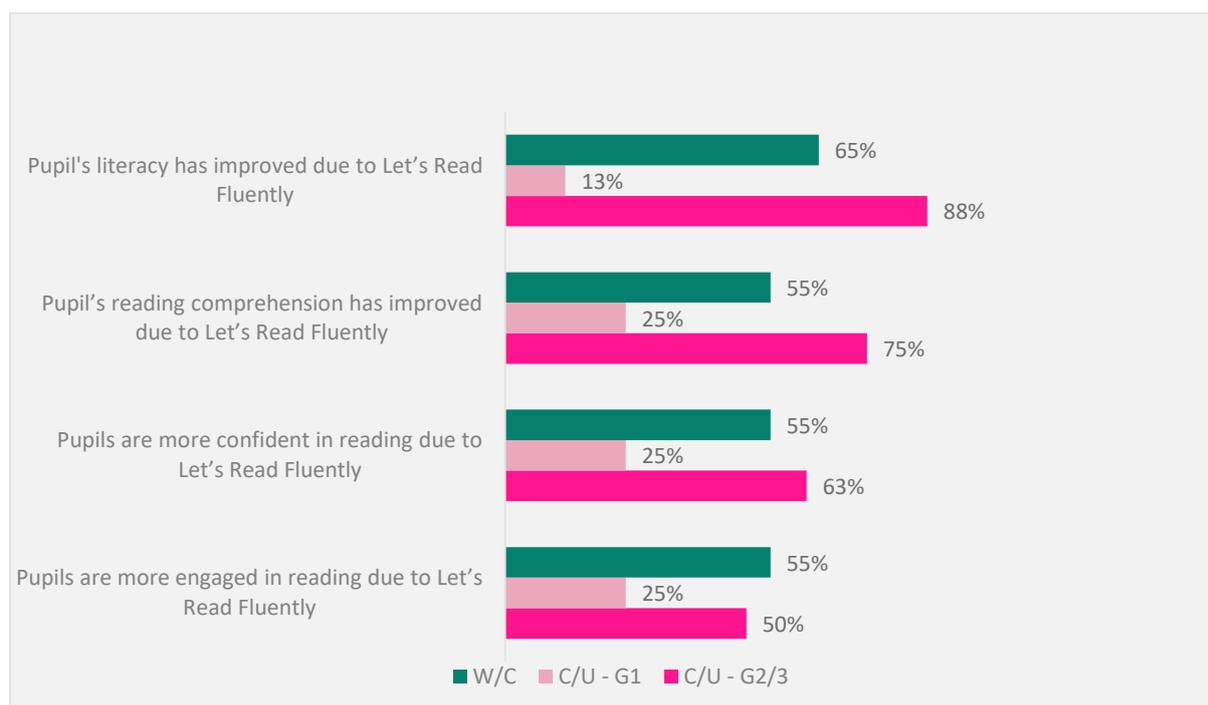
Overall, C/U classroom teachers saw improvements in pupils’ reading ability, evidenced by their ability to recognise letters, syllables, and four-letter words following the intervention. C/U pupils also stated they were better at reading following LRF!, although they gave no specific reasons for this. Additionally, both C/U classroom teachers and parents observed improvements in pupils’ writing abilities.

However, as with the W/C model, C/U classroom teachers did not report improvements among all pupils. They noted pupils who had a high number of absences from sessions were least likely to have improved.

⁴⁴ On the other hand, our descriptive sub-group analysis explained in the Appendix A shows that in the W/C model lower achieving pupils, on average, increased their EGRA+pre-lit score more compared to higher achieving pupils from baseline to endline. Please see Appendix A for more detail.



Figure 26: Teachers Perceptions of Pupil Outcomes - Agreement



Source: Teacher survey, W/C n=20, C/U Grade 1 n=8 and C/U Grades 2/3 n=8

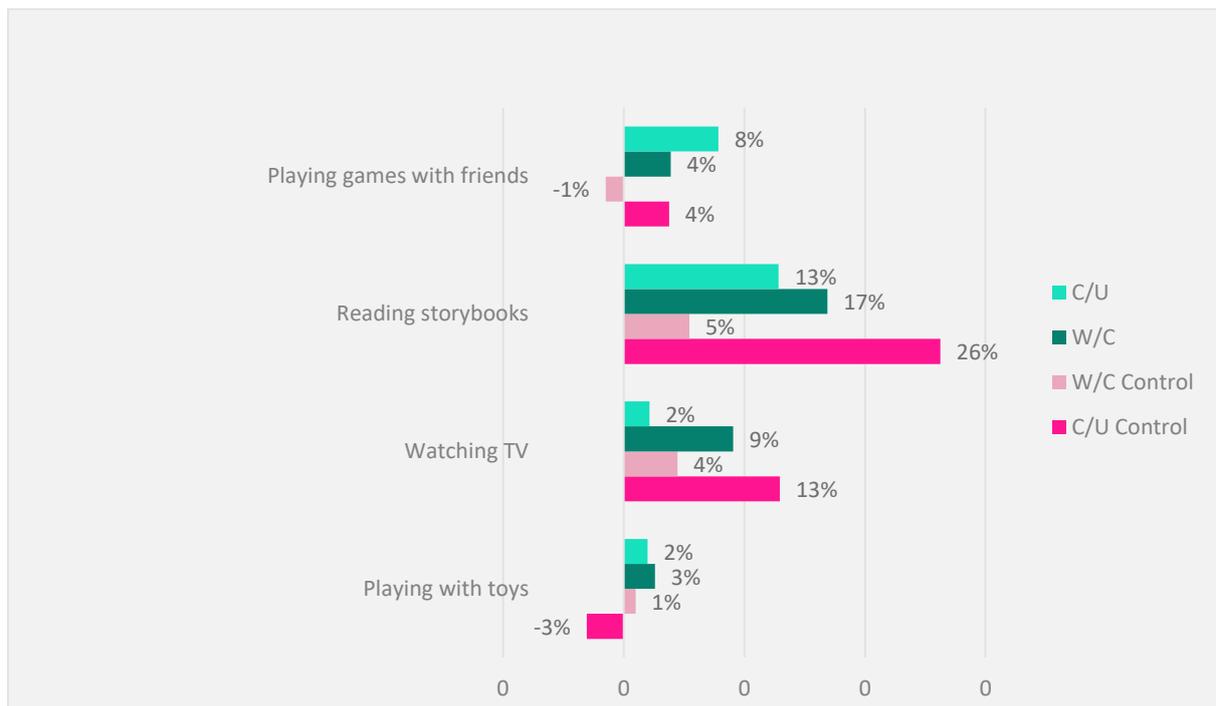
Literacy engagement

Across both W/C and C/U, pupil positivity about reading storybooks increased between the baseline and endline surveys (see Figure 27)⁴⁵. Notably, the increase in favourability for reading storybooks was higher than any other activity for both W/C and C/U pupils. When compared to the control group, the increase in favourability was higher for W/C pupils (17% compared with 5%). However, the results were less clear for the C/U model. C/U pupils saw an increase in favourability of 13% but the control group for C/U scored 26%. It is not clear why control C/U pupils saw such increased levels of favourability, although this could be due to the difference in control sample between baseline and endline testing⁴⁶.

⁴⁵ This data is a percentage change in favourability between baseline and endline. Pupils were asked to select how they felt about each of the activities based on three faces: one sad, one neutral, and one happy. These numbers are the percentage change in the proportion of pupils who selected the happy face between baseline and endline.

⁴⁶ The pupil survey was still in development when baseline testing of pupils in Grades 2/3 (for both intervention and control) was taking place. Therefore, the baseline data only contains Grade 1 pupils, whereas the endline data contains Grade 1 pupils and those in Grades 2/3. Subsequently, the endline sample is larger compared to the baseline and any change in results between the two surveys may be attributable to the variation in sample.

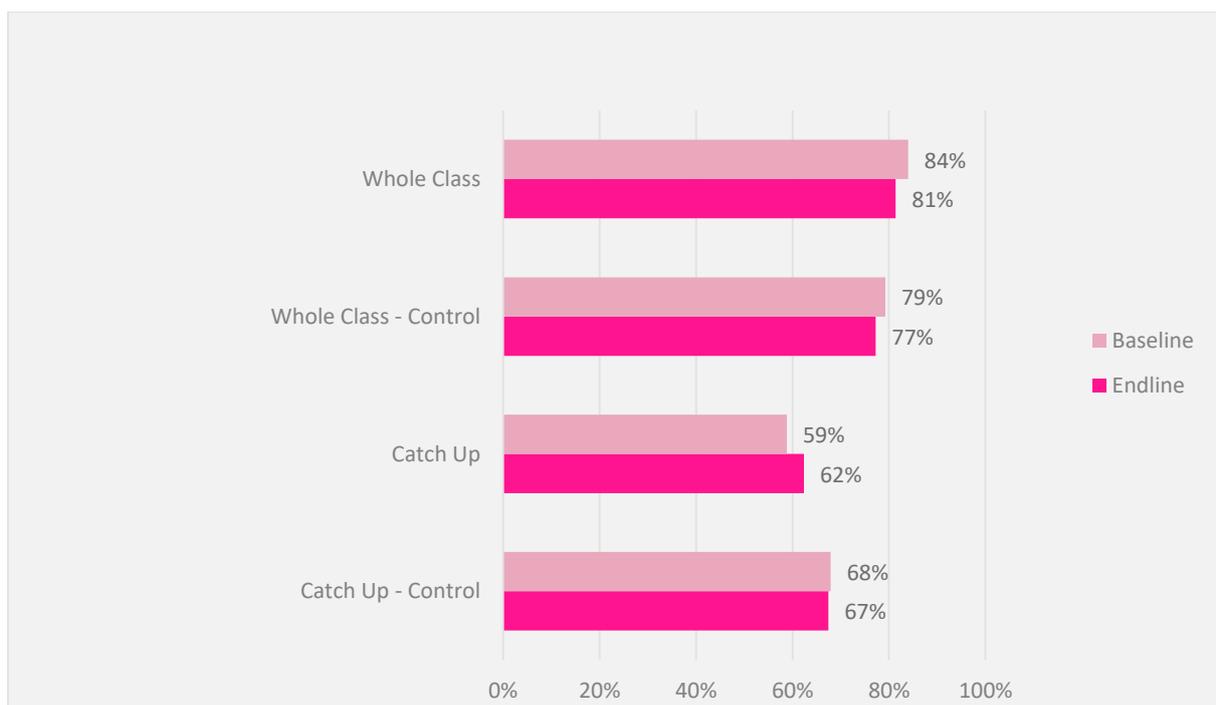
Figure 27: Pupil perceptions of activities – Percentage change in Favourability



Source: Pupil survey, W/C n=221-231 and C/U n=34-77

In baseline and endline surveys, the majority of pupils across all three arms of the trial agreed they would like to have more storybooks at home (see Figure 28). However, unlike the increased positivity around reading storybooks, there was limited change between baseline and endline in the proportion of pupils who would like to have more books at home.

Figure 28: Pupils who would like to own more storybooks at home - Agreement (Pupil survey)



Source: Pupil survey, W/C n=231, W/C – Control n=202, C/U n=34-77 and C/U – Control n=53-132



TEACHERS

Teaching practices

W/C teachers surveyed believed they would continue to use the skills and approach of LRF! in their teaching going forward (70%), although one in four (25%) disagreed with this statement. As with other survey measures, C/U resource room teachers gave similar responses with regard to their Grade 1 pupils (75% agreement) but were overwhelmingly positive about Grades 2/3 (100%).

Teacher burden

In FGDs, coaches reported their visits to schools and the observations of LRF! sessions had caused stress and tension for teachers. They described how both W/C classroom teachers and C/U resource room teachers were under pressure as they perceived the visits to be assessments, like those conducted by Ministry officials. Coaches noted, despite their reassurances, teachers continued to feel anxious during visits.

“Even during the visits, I felt that I needed to repeat myself. One teacher asked me how much I’d score her out of 10... I tell them I am not an inspector and am independent from the ministry. Many of them couldn’t get rid of this idea.” - Coach



UNINTENDED OUTCOMES

Pupil engagement with other subjects

In FGDs, coaches and parents reported improved pupil engagement with other subjects, an unintended positive outcome attributed to the LRF! intervention. Both groups evidenced this by how pupils contributed in class. For example, coaches stated pupils had learned the routine of lessons and so they know when to listen and when to participate. Coaches perceived this had also instilled discipline within pupils. Moreover, a W/C parent commented their child contributed more in other school subjects and so was more engaged with their learning. Not all parents agreed with this though, as they perceived that LRF! had resulted in pupils reducing focus on other school subjects, although they did not provide any reasons why.

Pupil psychological and social outcomes

FGDs found C/U pupils had improved psychological and social outcomes. This was another positive outcome which was not part of the logic model. Coaches reported Grade 1 C/U pupils benefitted from having targeted and personalised support in the resource room. Furthermore, C/U classroom teachers acknowledged pupils were more motivated and encouraged following resource room sessions. They found pupils returned to the main classroom happier and in a better mood.

Despite these reported psychological improvements for C/U pupils, coaches also witnessed negative impacts for Grade 1 pupils. They perceived that administering the diagnostic tool with pupils when they have just started school can feel disruptive. Also, coaches noted it can be disheartening for pupils who find the diagnostic tool challenging.



CONCLUSIONS

The pilot evaluation of LRF! aimed to assess evidence of promise, feasibility of intervention, feasibility of trial and readiness for trial of two different delivery approaches: W/C and C/U.

Below we reflect on some of our main findings, drawing out some suggestions for future delivery.

Evidence of promise

The impact evaluation employed a RCT design in preparation for a future trial, not with the aim to establish evidence of the causal effect of LRF! on the intended outcomes. Therefore, the findings from the impact evaluation should be treated with extreme caution, as the study was underpowered to detect differences between the intervention and the control arms.

Evidence from the impact evaluation analysis suggests that children in schools that implemented the W/C approach have made greater improvements in literacy based on the EGRA+pre-lit assessment in comparison to children in control schools. Similar, perceived outcomes for the W/C approach, as reported through the teacher survey and FGDs suggested improvements in pupils' reading comprehension, engagement and increased confidence in reading. This was corroborated by coaches and parents of children in W/C, who perceived positive changes in children's literacy. Overall, findings from the EGRA assessment, the teacher survey and the FGDs suggest that W/C can improve literacy attainment.

The evidence on whether the C/U approach improved literacy were mixed. The impact analysis does not indicate that C/U has the potential to improve literacy attainment for pupils in Grade 1, 2, and 3. The evaluation identified differences in the perceived impact of C/U on older pupils (Grade 2 and 3) in comparison to younger pupils (Grade 1). Teachers perceived improvements in pupil performance and confidence in reading amongst Grade 3 pupils and to a lesser extent for Grade 2 pupils. However, for Grade 1 pupils, resource room teachers were far less positive in their survey responses. Similar to teachers, coaches also observed more positive impact for pupils in Grades 2 and 3. The evidence suggests that positive changes in literacy due to C/U are more likely for pupils in Grades 2 and 3 than for pupils in Grade 1.

The evaluation also identified some potential unintended outcomes for pupils in Grade 1 assigned to the C/U model. There were reports from coaches that LRF! could have some negative psychological impacts for Grade 1 C/U pupils. This seemed to be a result of the difficulty of the material. Coaches and teachers indicated that the material in the practice book was too difficult for pupils in Grade 1 C/U, and pupils were particularly frustrated during the 'I do' session. Similarly, findings from parents and pupil reports suggest that LRF! could have been too difficult for children in C/U Grade 1.

Feasibility of Intervention

Overall, LRF! was described as feasible to deliver and was implemented as planned in terms of the key ingredients being delivered. However, analysis of the more detailed IPE data has revealed the need for modifications in the following broad areas:

Training and coaching sessions

The vast majority of teachers were positive about the training and coaching sessions. Teachers across both intervention arms attended the training and they felt that the training equipped them with the skills and knowledge needed to deliver LRF!. Even though most teachers found the



training useful, some teachers from both the W/C and C/U models also felt that the training was insufficient to cover all the material. Coaches expressed concerns over the suitability of the training for C/U resource room teachers. They felt that W/C were more experienced and that the training did not equip C/U resource room teachers with the needed skills to deliver LRF!. Coaches felt that C/U resource room teachers required additional support to attend the training and deliver LRF!. Coaches also reported that resource room teachers were absent from training sessions due to logistical issues.

LRF! sessions

According to the observation data, teachers delivered the 'I do, We do, You do' elements of the lesson content in most LRF! sessions. Findings from the FGDs with teachers and coaches indicate that the 'You do' process of LRF! was most challenging to follow across both intervention arms. Moreover, coaches, C/U resource room and W/C teachers thought the volume of content was too much to cover. This resulted in teachers adapting parts of delivery. To reduce the content some C/U resource teachers did not recap the previous lesson. To support pupils that struggled with the 'You do' process C/U resource teachers offered additional support to pupils. Resource room teachers in the C/U intervention arm also adapted delivery to improve pupil engagement with the intervention. To help them cover all the content, W/C teachers lengthened LRF! sessions and reduced the length of other curriculum lessons. To improve engagement among pupils with lower language ability, W/C teachers taught LRF! at a slower pace and ensured all pupils were of an adequate standard before proceeding to the next lesson.

Practice book

The practice book was seen as a useful aid and was used in most LRF! sessions as intended, but perceptions on its content were less positive. There are themes that emerged in the analysis of the FGDs in terms of how the practice book could be improved. First, parents and teachers from both C/U and W/C expressed concerns about the size of the book. They commented that the large size was off-putting for pupils, too heavy and the large-print was difficult to work with. Second, participants in the FGDs expressed concerns about its content. In particular, it was perceived that the page content was too dense, or they outlined examples of unclear text and typos. Third, there were a number of suggestions about how the presentation of the content could be improved through better use of coloured front, pictures or invented words. Fourth, the practice book differed from the Arabic curriculum book, which made it challenging to reinforce learning. Last, the evidence from the FGDs suggests that the book was not suited to resource room delivery. The current LRF! book offers 'one-size fits all' that does not take into account the different levels of ability among pupils. Coaches and teachers thought that there should be different versions of the book that would take into account differences in ability.

Diagnostic tool

Even though many teachers and coaches thought that the diagnostic tool was successful in identifying the most appropriate pupils in Grades 2 and 3 for the C/U model, they felt that the tool could not capture Grade 1 pupils appropriately. Some teachers also believed that the tool was not necessarily useful in identifying the lowest performing pupils, while some coaches felt that the tool does not account for learning difficulties. Coaches and teachers also voiced concerns about the test environment of the tool proving unsettling for some pupils, and this could have made pupils underperform. This suggests that a future evaluation of C/U may want to reconsider the process for selecting pupils.



Barriers and facilitators to successful delivery

Participants reported school-based and pupil factors that impacted the delivery. In terms of the barriers schools encountered, many teachers and coaches mentioned staffing challenges including absences of teachers or changes in school leadership. Another barrier, mentioned by many coaches, was the varying experience among teachers. One of the most frequently reported barriers hindering delivery was the national curriculum. Coaches and teachers did not think that LRF! is aligned with the national curriculum. This, coupled with the fact that the national curriculum is very challenging in terms of content and homework for pupils, made it difficult to conduct LRF! lessons alongside the national curriculum.

Another element external to LRF! that could have impacted on the success of the intervention was pupil absenteeism due to a local sickness outbreak, and the fact that the pilot study was conducted right after school closures due to the COVID-19 outbreak. As a result of school closures prior to the pilot study, Grade 1 pupils had missed their kindergarten education. This made LRF! challenging for Grade 1 pupils as they lacked the foundational skills and knowledge. Parents also expressed that the COVID-19 outbreak negatively affected pupils' attitudes towards school.

Feasibility of the efficacy trial

On the whole, the process of recruiting schools to the pilot evaluation was well planned and executed. Evidence from take up data showed that it was more challenging to recruit schools to LRF! in Amman, middle excluding Amman region and near north region in comparison to near south region. The future recruitment process should consider regional differences when preparing for recruitment. Only two C/U schools dropped shortly after randomisation, and the majority of pupils completed the baseline and endline assessment in W/C Grade 1 (i.e. 89% of pupils across the treatment and control group) and C/U Grade 1,2, 3 (i.e. 70% of pupils across the treatment and control group). We observed differences in attrition between treatment and control arms for C/U Grade 1, however the size of the difference is mainly driven by the small sample sizes. Overall, the level of attrition in the pilot evaluation was at a normal level, and it was driven mainly by pupils not attending on the testing day due to local sickness outbreaks, i.e. chickenpox, or other illness. In these cases, a mop-up testing was organised on a second day, asking those pupils who were absent to be tested. The timing of conducting testing before Eid may have also had an impact on attrition rates. This suggests that a future evaluation should take into account the time of the year when planning data collection activities.

As outlined in the method section the primary outcome in the pilot evaluation was the EGRA+pre-lit tool. Analysis of the pre-test data showed that the EGRA+pre-lit tool captured a range of learner proficiencies.

Assessing readiness for trial

The evidence from the pilot evaluation suggests that the key intervention inputs, activities, and outputs as outlined in the logic model for both approaches (i.e. the W/C and C/U) are acceptable to schools and they were implemented as intended with some adaptations. The current evaluation is not able to draw firm conclusions on how well the proposed logic model works in practice, that is, whether LRF! changes pupils' literacy levels considering that this is pilot study with small sample sizes. Nonetheless, given the inconclusive findings from the impact evaluation and the evidence from the survey and FGDs suggesting that the intervention was largely delivered as intended, the pilot evaluation did not yield any rationale to update or adapt the logic model.



Regarding the evaluation design, this evaluation demonstrated successful piloting of the key procedures for RCT, such as recruitment, engagement with data collection activities, and randomisation. This means that the evaluation is sufficiently well-specified to be delivered at scale.

As part of the workshops undertaken for this pilot evaluation, we explicitly assessed the readiness of the implementation and delivery partner to deliver LRF! at scale. The partners have the capacity and knowledge to deliver and evaluate LRF! at scale. It was agreed that allowing enough time for planning and undertaking some of the suggested modifications will be vital to the evaluation success.

Overall, the evidence from the pilot evaluation suggests that W/C is ready for trial conditional on modifications to the time allocated for LRF! delivery, and the content of the practice book. The conclusion is shared by coaches and teachers that future delivery may need to allow for more time for delivery (e.g. to extend the delivery to two semesters).

The inconsistencies in evidence of promise and the suggestions for improvements in relation to the feasibility of C/U suggest that the C/U intervention is not ready for trial. This conclusion is shared by teachers, coaches and parents who responded to the surveys and FGDs, suggesting that substantial changes are required. Training for resource room teachers was seen as useful in parts, but the evidence suggested that the training would require improvement, but also additions in terms of the support that was offered. The evaluation indicated that there was resourcing challenge among resource room teachers that needs to be addressed. The evidence also suggested that, for many C/U pupils, LRF! was too challenging and would require substantial further design work before it is suitable for pupils with lower ability. The selection process based on the diagnostic tool, was not considered particularly useful, and future evaluation may consider different approaches for selecting pupils. Taking all these into account, at least one more year of C/U development seems appropriate before it is reassessed for readiness for trial.

Limitations

There are key limitations to this study that readers should be aware of:

The pilot evaluation provided an assessment of evidence of promise rather than conclusive evidence of impact. Evidence on impact was limited given the small sample sizes, and the study was not powered to detect any meaningful differences in outcomes. While the statistical analysis is comprehensive and presents useful assessment of changes on the intended outcomes over time, it should be considered as exploratory only. This means that any changes in literacy cannot be confidently attributed to LRF!. While a range of schools took part in the LRF! pilot study, the sample of schools that engaged with the pilot study could not be viewed as a representative of schools in Jordan.

A further limitation is that the evaluation was impacted by COVID-19 and as a result some of the data collection activities were delayed. For instance, data collection for C/U Grade 2 and 3 was delayed for three months.

Our findings build upon evidence provided only by those teachers, coaches, parents and pupils that engaged in the surveys, FGDs interviews. We cannot rule out the possibility that those teachers or parents that engaged with the survey or FGDs have been more engaged with LRF!.



While the evaluation captures different views at different time points, the findings from the teacher survey are based on a relatively small number of respondents, and there may be some bias in the type of teachers that choose to provide answers to the survey.

While the evaluation undertook FGDs with various stakeholders, the FGDs provided limited detail. This means that future FGDs need to explore in more depth some of the barriers and facilitators to successful delivery of LRF!.

Baseline data on the pupil survey is missing for C/U Grade 2 and 3.

Considerations for further development of LRF! and recommendations for future scaled evaluation

As a test for feasibility of LRF!, the pilot demonstrated that it is feasible to implement the intervention. Overall, schools adhered to the prescribed delivery, however, a number of adjustments are needed for future successful delivery.

Below we summarise our recommendations for the further development of the LRF! intervention, and key considerations for future evaluation.

Considerations for the development of LRF!

The evidence from the pilot evaluations suggests that considerations for adaptations may be needed in the following areas:

- **Training for resource room teachers:** An additional coaching for C/U resource room teachers may be beneficial. Similarly, adaptations to the training materials for C/U teachers may be required as they work solely with struggling pupils and they are less experienced than classroom teachers.
- **Time needed for delivering LRF! sessions for the W/C and C/U models:** A future delivery of LRF! may want to consider extending the time allocated for each session, or offering LRF! for an extended period of time.
- **Suitability for C/U across different grades:** There is an indication that the C/U model may have been more suitable for children in higher grades. This may be driven by the fact that pupils in higher grades in the C/U model have higher language ability. Changes to the content for pupils in Grade 1 may be appropriate to increase potential impact at this grade.
- **Practice book:** The practice book was perceived to be more suitable for the W/C model, but there were important suggestions for improvements. These included changes to its size, format of presentation and minor changes to its content. There were also a number of recommendations for how to ensure better alignment with the curriculum book that should be considered carefully before future delivery (see Feasibility of Intervention for more detail).⁴⁷
- **Selecting pupils for the C/U model:** The evidence suggests that there is a need to reconsider the process for selecting pupils. In particular, future delivery of the C/U model should provide clear guidance on what is an appropriate test environment, and on the process of testing pupils with learning difficulties.
- **Engagement with LRF!:** The evidence suggests that there is scope to improve engagement of teachers and parents with the LRF! intervention. Future delivery should consider strategies to improve teacher 'buy-in' to the LRF! method. Similarly , since parental

⁴⁷ The national curriculum is currently being revised and will be more similar to the LRF curriculum.



engagement has been seen as a key factor for successful delivery, it is important to consider further strategies in addition to the information session that can facilitate higher perceptions of the added value of the LRF! intervention among parents.

Considerations for future scaled evaluation

We recommend that a scaled evaluation of LRF! should carefully consider and plan in the following domains:

- **Recruitment of schools:** While recruitment of schools was successful for the pilot, suggesting that LRF! was seen as an attractive intervention for schools, it also revealed some challenges. There were important regional differences in the level of engagement, which should be considered during the design of a larger evaluation. From a practical perspective, it is important to obtain accurate contact information for all schools for recruitment.
- **Assessment:** The pilot evaluation revealed a high level of absenteeism among pupils, suggesting that there is need to plan for multiple rounds of assessment at endline to reduce attrition.
- **Outcome measures:** Whilst the EGRA+pre-lit assessment is seen as an appropriate outcome measure, evidence from consultations with enumerators suggests that enumerators and/or learners would benefit from clearer instructions on the assessment. We observed large floor effects for the secondary outcome measures (the sub-domains of the EGRA assessment). This was expected given the age of the pupils. The scores of the primary outcome (EGRA+pre-lit) were roughly normally distributed suggesting that it is important to administer the pre-literacy items in addition to the EGRA.
- **Systematic data collection on usual literacy practice:** Although there is some evidence from the pilot on business-as-usual practice, there would be value in collecting the information in a systematic way. This will be helpful in contextualising the impact evaluation findings.



REFERENCES

- Abadzi, H. (2013) 'Literacy for All in 100 Days: A research-based strategy for fast progress in low-income countries', *GPE Working Paper Series on Learning No. 7*.
- Abadzi, H., & Martelli, M. (2014). Efficient reading for Arab students: Implications from neurocognitive research. World Summit of Innovation in Education (WISE).
- Brombacher, A. et al. (2012) EdData II: National Early Grade Literacy and Numeracy Survey–Jordan Intervention Impact Analysis Report. Research Triangle Park, North Carolina, US: Research Triangle Institute (RTI) & USAID.*
- Castles, A., Rastle, K. and Nation, K. (2018) 'Ending the reading wars: Reading acquisition from novice to expert', *Psychological Science in the Public Interest*, 19(1), pp. 5–51.
- Eckert, M., Wilson, E., Abadzi, H., & Jeon, S. (2020). Improving Arabic reading fluency: Results from Iqra, an early-grade reading intervention in Ras Al Khaimah (Policy Paper No. 39). Sheikh Saud bin Saqr Al Qasimi Foundation for Policy Research. <http://dx.doi.org/10/18502/aqf.0142>
- EEF (2021), Phonics, Available at: <https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit/phonics>
- EEF (2018b) 'Statistical Analysis Guidance for EEF Evaluations', London: Education Endowment Foundation:
https://educationendowmentfoundation.org.uk/public/files/Grantee_guide_and_EEF_policies/Evaluation/Writing_a_Protocol_or_SAP/EEF_statistical_analysis_guidance_2018.pdf
- EEF (2019) 'Implementation and Process Evaluation Guidance for EEF evaluations', London: Education Endowment Foundation:
https://educationendowmentfoundation.org.uk/public/files/Evaluation/Setting_up_an_Evaluation/IPE_guidance.pdf
- Goldstein, H., & Healy, M. J. (1995). The graphical presentation of a collection of means. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 158(1), 175-177.
- QRF; 'Exploring Jordan's performance' (September 2020), <https://www.qrf.org/en/what-we-do/research-and-publications/pisa-2018-exploring-jordan%E2%80%99s-performance>
- RTI, Aug. 2012, Pupil Performance in Reading and Mathematics, Pedagogic Practice, and School Management in Jordan, <https://earlygradereadingbarometer.org/files/EGRA%20in%20Jordan.pdf>
- RTI, Nov. 2018, Early Grade Reading and Mathematics Initiative Lot Quality Assurance Sampling Assessment.
<https://ierc-publicfiles.s3.amazonaws.com/public/resources/Jordan%20RAMP%20LQAS%20Summary%20Report%20FINAL.pdf>
- RTI International. (2018). 2018 Early Grade Reading and Mathematics Initiative (RAMP) Lot Quality Assurance Sampling Assessment (p. 9). RTI International.*



Seidenberg, M. (2017) Language at the Speed of Sight: How we read, why so many can't, and what can be done about it. New York, NY: Basic Books

UNICEF, 'The Impact of COVID-19 on children in the Middle East and North Africa' (November 2020),

<https://www.unicef.org/mena/media/10231/file/Impact%20of%20COVID%20on%20Children-Snapshot%20report.pdf%20.pdf>

World Bank (2019) Ending learning poverty: what will it take? Washington, D.C.: World Bank.

Available at:

[https://openknowledge.worldbank.org/bitstream/handle/10986/32553/142659.pdf?sequence=7.](https://openknowledge.worldbank.org/bitstream/handle/10986/32553/142659.pdf?sequence=7)



APPENDICES

Appendix A: Additional Descriptive Subgroup Analysis

Low, medium and high ability pupils

Table 18 below shows the mean endline EGRA+pre-lit scores by the baseline EGRA+pre-lit categories (i.e. low, medium, and high) for the W/C pupils. As expected, the pupils in the high achieving group based on their baseline assessment had, on average, a higher endline score than those in the low and medium achieving groups. This was the case for both the intervention and control groups. Furthermore, the differences in the mean endline EGRA+pre-lit score between the intervention and control group decrease from the lower achieving group to the higher achieving group (Table 18). For instance, while the difference in the mean endline EGRA+pre-lit score was 61 in the lower achieving group, this difference decreased to 36 and 30 for the medium and higher achieving groups, respectively.

Table 18: Mean endline EGRA+pre-lit score by baseline attainment level

Terciles	W/C Intervention		W/C Control	
	Mean	Std	Mean	Std
Low	501	82	440	63
Medium	565	79	529	82
High	660	83	630	86

We further examined how much the EGRA+pre-lit score changed on average, from baseline to endline. Table 19 below shows the mean improvement in the EGRA+pre-lit score by the baseline EGRA+pre-lit categories for both the W/C intervention and control groups. As with the earlier descriptive analysis, low, medium, and high achieving pupils in the intervention group had, on average, higher improvement in the EGRA+pre-lit score than those in the control group.

Moreover, Table 19 indicates differential improvement in the EGRA+pre-lit score by the baseline EGRA+pre-lit categories. While the average change in the EGRA+pre-lit measure was 86.1 for the lower achieving pupils in the intervention group, this change decreased to 73.5 and 61.5 for the medium and high achieving pupils in the intervention group, respectively.

Table 19: Mean endline EGRA+pre-lit score by baseline attainment level

Terciles	W/C Intervention		W/C Control	
	Mean	Std	Mean	Std
Low	86.1	74.1	39.8	64.3
Medium	73.5	80.1	37.5	78.6
High	61.5	72.6	29.1	73.9



Gender

Table 20 below shows the gender distribution for the W/C intervention and control groups. It clearly shows that the two groups are not balanced in terms of gender. While 67% of the W/C intervention group were girls, only 25% of the W/C control group were girls.⁴⁸

Table 20 also shows the mean endline EGRA+pre-lit scores by gender for the W/C pupils. While we do not see a large difference in the endline EGRA+pre-lit score between boys and girls in the intervention group, girls, on average, had a higher EGRA+pre-lit score than boys in the control group. More specifically, the mean EGRA+pre-lit score was 501 for boys in the control group, this was 566 for girls in the control group.

Table 20: Mean EGRA+pre-lit score by gender

Gender	Number of pupils	W/C Intervention		Number of pupils	W/C Control	
		Mean	Std		Mean	Std
Boys	77 (33%)	573	105	152 (75%)	501	101
Girls	154 (67%)	578	103	50 (25%)	566	102
Total	231 (100%)			202 (100%)		

⁴⁸ The imbalances in some characteristics are possible in the pilot evaluations due to their small sample size.

APPENDIX B: ZERO SCORE ANALYSIS FOR SECONDARY OUTCOME MEASURES

We undertook a zero score analysis for the secondary outcome measures as part of the additional analyses. The percentage of pupils scoring zero was calculated for every secondary outcome measure. These percentages were calculated for the intervention and control groups at baseline and endline for both approaches. A graphical representation of these figures can be found below.

It is a common practice to add a 95% confidence interval for each quantity and check whether the confidence intervals of two quantities fail to overlap.⁴⁹ Failure to overlap is interpreted as indicating that the two quantities are significantly different at the 5% level.

Overall, our findings from the zero score analysis for the secondary outcome measures point to a statistically significant difference in percentages of pupils scoring zero in every QRF sub-domain of EGRA between intervention and control groups for the W/C model only.

Oral Reading Fluency (ORF)

Figure 37 shows the percentages of pupils scoring zero in the ORF sub-domain of EGRA by intervention and control groups for the W/C model. There were fewer pupils scoring zero in the intervention group at endline. Fifty-nine percent of pupils in the control group scored zero and 38% of pupils scored zero in the intervention group. It is clear from Figure 37 that these percentages are significantly different at the 5% level as their 95% confidence intervals fail to overlap.

Figure 38 shows the percentages of pupils scoring zero in the ORF sub-domain of EGRA by intervention and control groups for the C/U model. At endline, 78% of pupils in the control group scored zero and 76% of pupils in the intervention group scored zero. Although the percentage of pupils scoring zero looks slightly lower in the intervention group, their 95% confidence intervals overlapped indicating that they are not significantly different at the 5% level.

Letter Sound Identification (LSI)

Figure 39 shows the percentages of pupils scoring zero in the LSI sub-domain of EGRA by intervention and control groups for the W/C model. There were less pupils scoring zero in the intervention group than those in the control group at endline. Thirty percent of pupils in the control group scored zero while 12% of pupils scored zero in the intervention group. These percentages are significantly different at the 5% level as their 95% confidence intervals fail to overlap (Figure 39).

Figure 40 shows the percentages of pupils scoring zero in the LSI sub-domain of EGRA by intervention and control groups for the C/U model. At endline, 46% of pupils in the control group scored zero and 42% of pupils in the intervention group scored zero. Although the percentage of pupils scoring zero looks slightly lower in the intervention group, their 95% confidence intervals overlap, indicating that they are not significantly different at the 5% level.

⁴⁹ Goldstein and Healy, 1993, recommend using ± 1.39 (standard error of mean) to create the 95% confidence interval. This adjustment of confidence intervals allows obtaining the required significance level for the non-overlap criterion.



Syllable Identification (SI)

Figure 41 shows the percentages of pupils scoring zero in the SI sub-domain of EGRA by intervention and control groups for the W/C model. At endline, 45% of pupils in the control group scored zero while 23% of pupils scored zero in the intervention group. These percentages are significantly different at the 5% level as their 95% confidence intervals fail to overlap (Figure 41).

Figure 42 shows the percentages of pupils scoring zero in the SI sub-domain of EGRA by intervention and control groups for the C/U model. At endline, 64% of pupils in the control group scored zero and 63% of pupils in the intervention group scored zero. As it is clear from Figure 42, their 95% confidence intervals overlap, indicating that they are not significantly different at the 5% level.

Reading Comprehension (RC)

Figure 43 shows the percentages of pupils scoring zero in the RC sub-domain of EGRA by intervention and control groups for the W/C model. There were less pupils scoring zero in the intervention group than those in the control group at endline. Eighty-three percent of pupils in the control group scored zero and 68% of pupils scored zero in the intervention group. Figure 43 shows that their 95% confidence intervals fail to overlap, indicating they are significantly different at the 5% level.

Figure 44 shows the percentages of pupils scoring zero in the RC sub-domain of EGRA by intervention and control groups for the C/U model. At endline, 85% of pupils in the control group scored zero and 89% of pupils in the intervention group scored zero. Although the percentage of pupils scoring zero looks slightly lower in the control group than in the intervention group, their 95% confidence intervals overlap, indicating that they are not significantly different at the 5% level.



Figure 37: Proportion of pupils that scored zero in Oral Reading Fluency (W/C Model)

Base: 202, 231 pupils

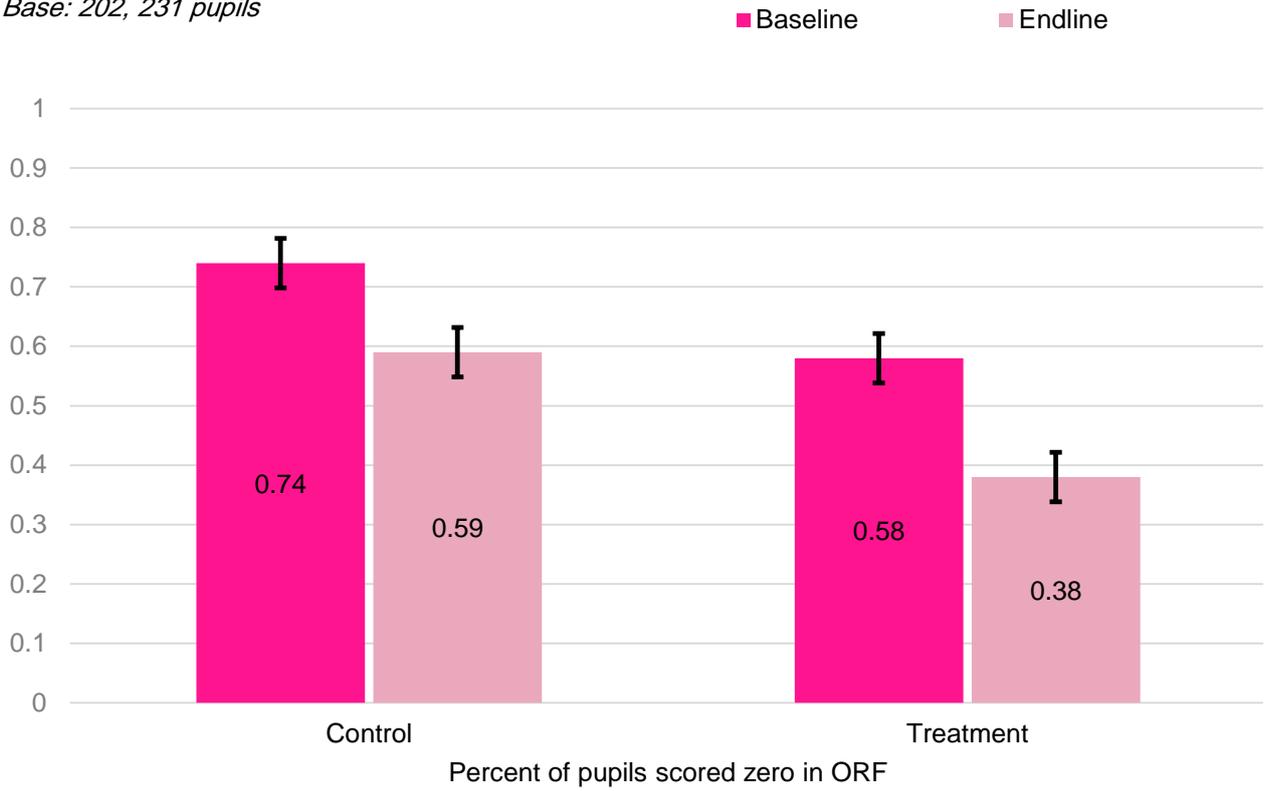


Figure 38: Proportion of pupils that scored zero in Oral Reading Fluency (C/U Model)

Base: 133, 76 pupils

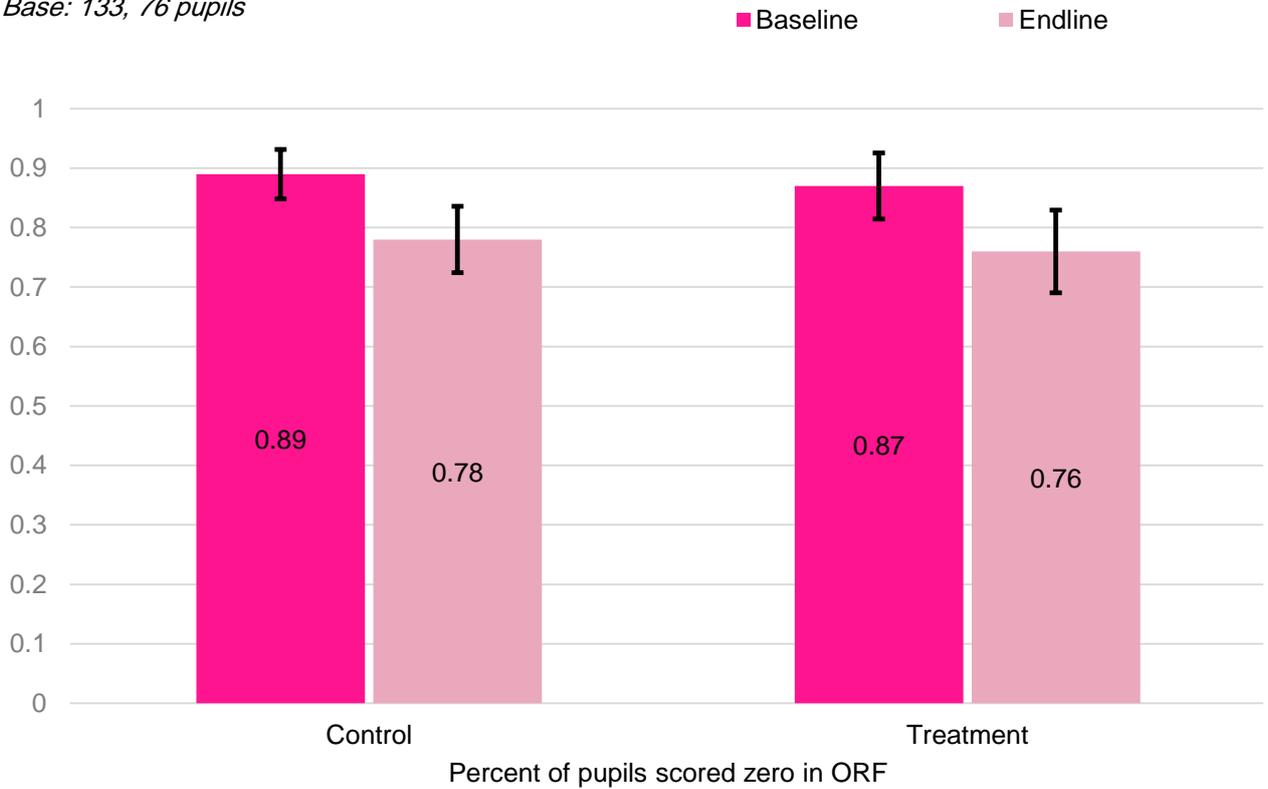


Figure 39: Proportion of pupils that scored zero in Letter Sound Identification (W/C Model)

Base: 202, 231 pupils

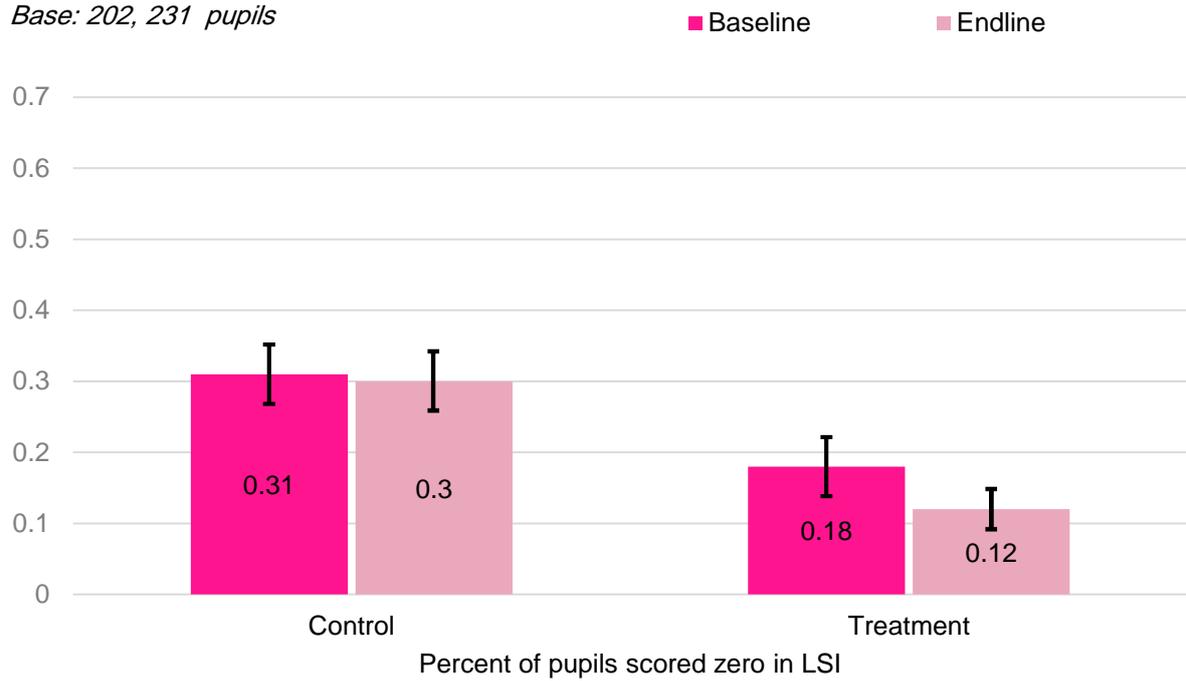


Figure 40: Proportion of pupils that scored zero in Letter Sound Identification (C/U Model)

Base: 133, 76 pupils

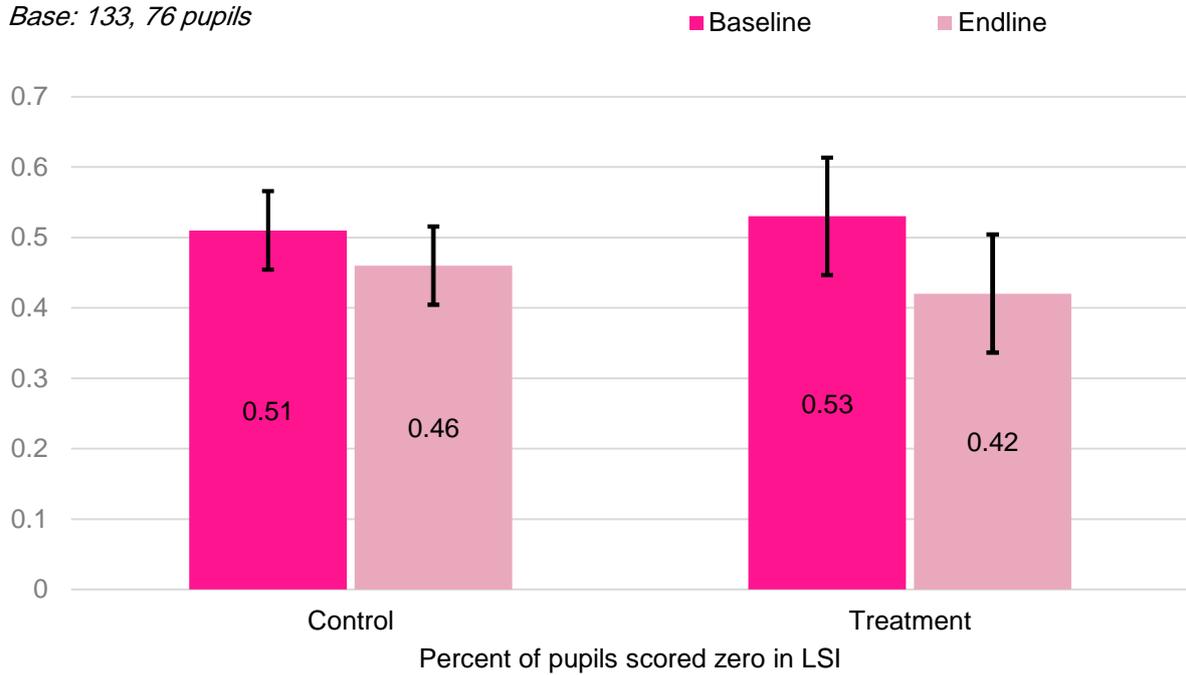


Figure 41: Proportion of pupils that scored zero in Syllable Identification (W/C Model)

Base: 202, 231 pupils

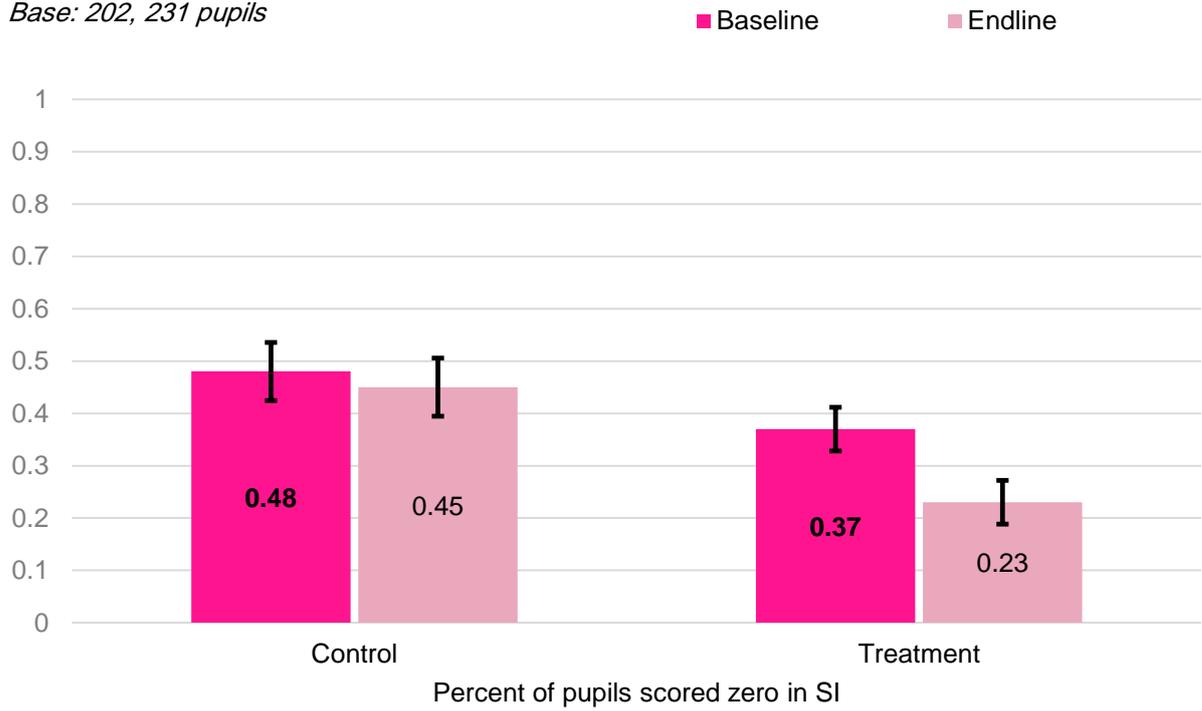


Figure 42: Proportion of pupils that scored zero in Syllable Identification (C/U Model)

Base: 133, 76 pupils

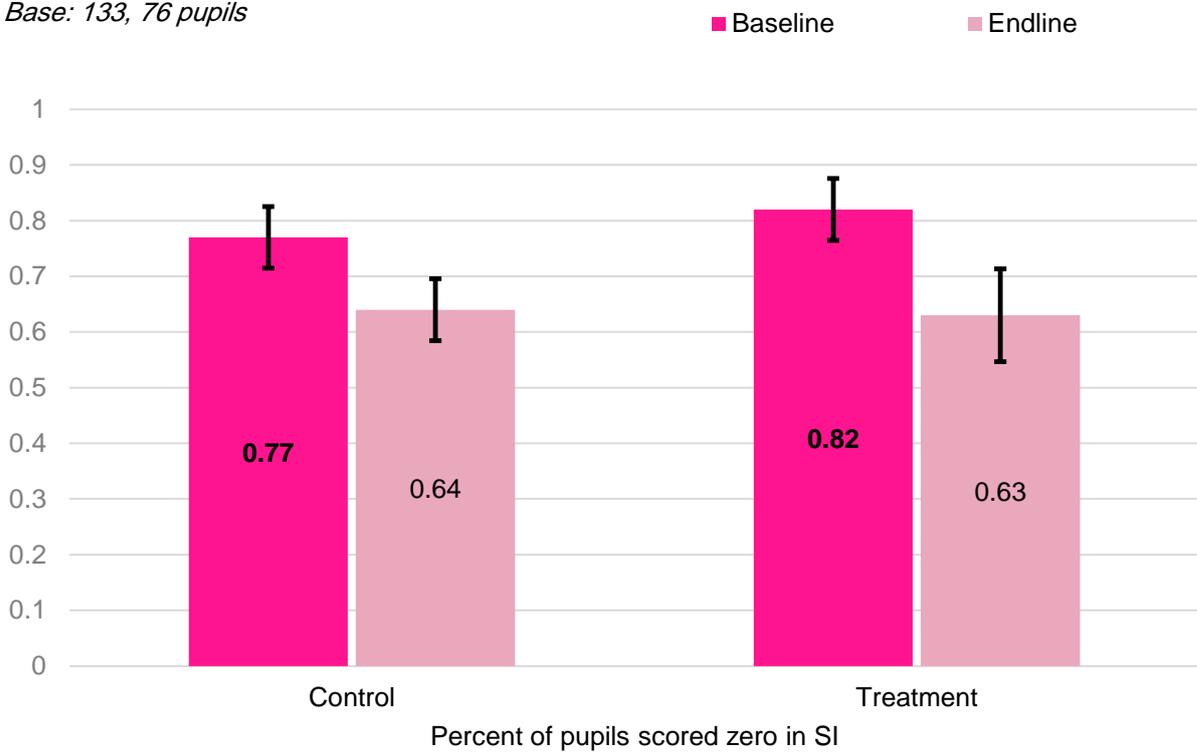


Figure 43: Proportion of pupils that scored zero in Reading Comprehension (W/C Model)

Base: 202, 231 pupils

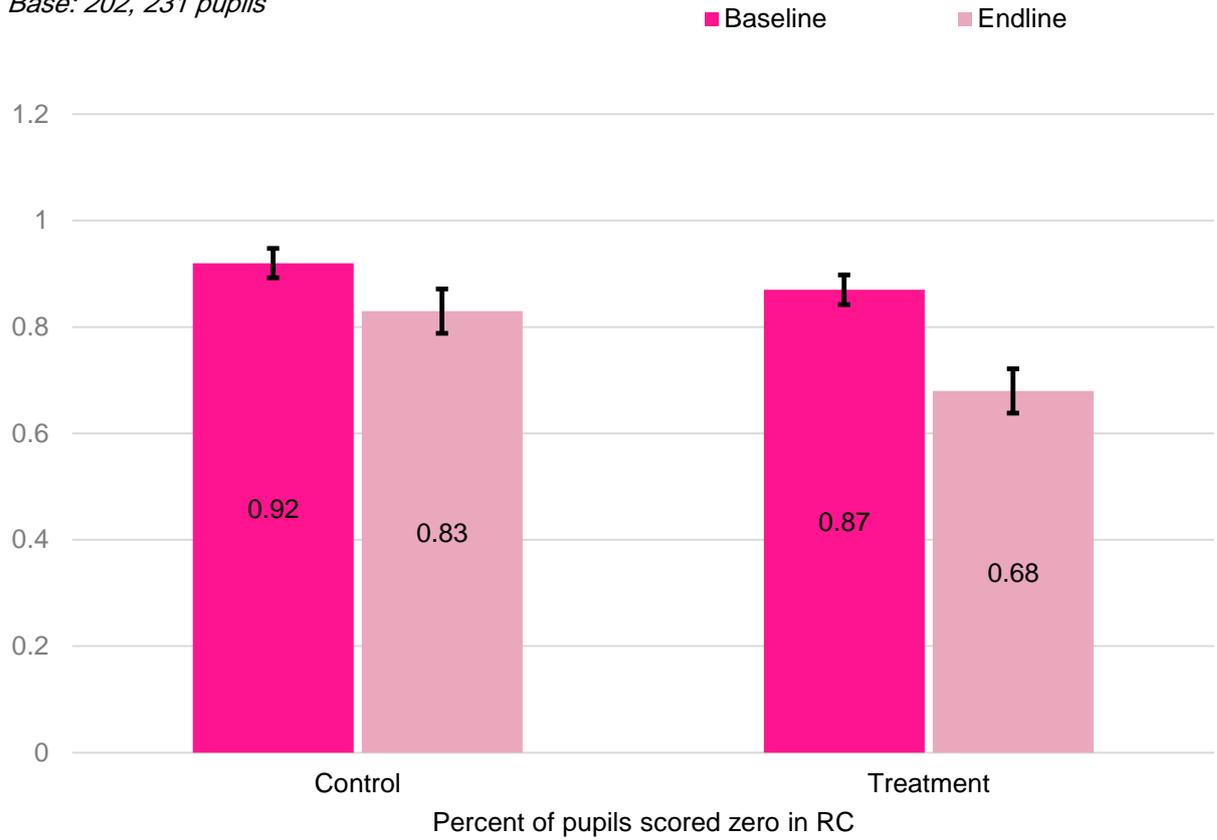
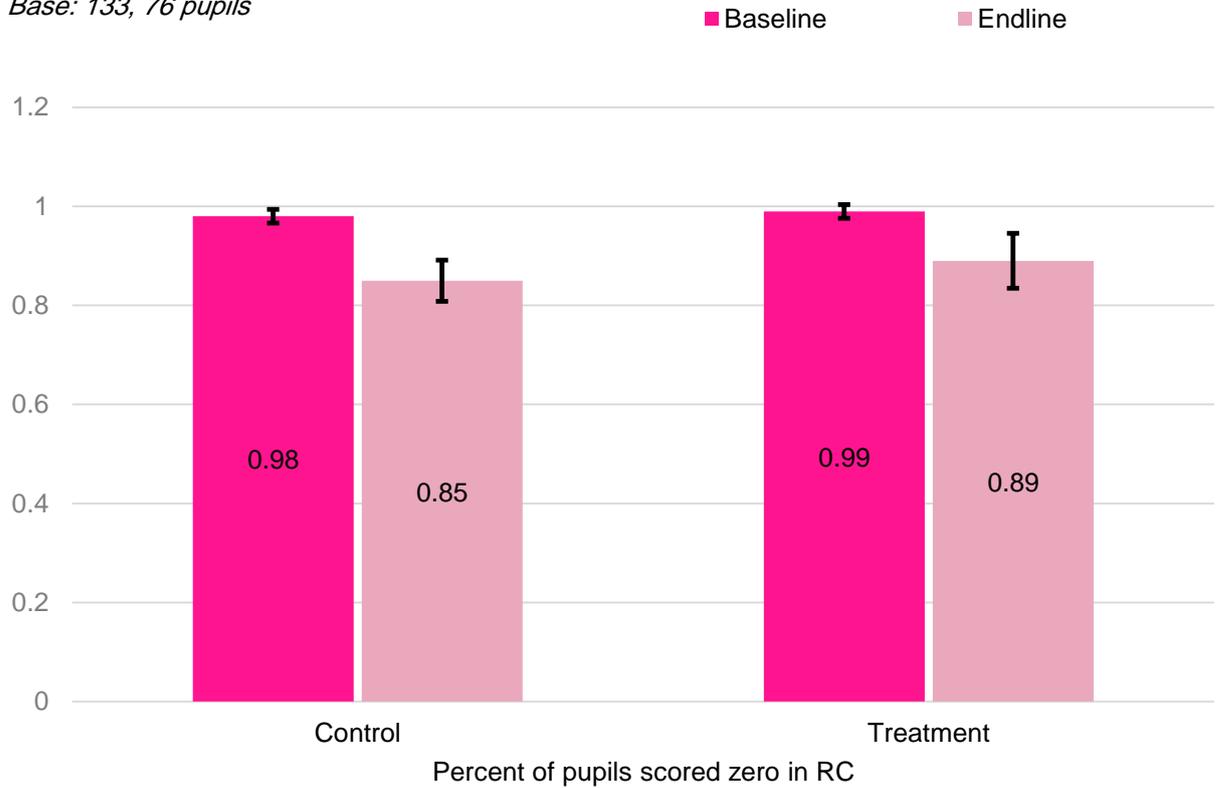


Figure 44: Proportion of pupils that scored zero in Reading Comprehension (C/U Model)

Base: 133, 76 pupils



APPENDIX C: DATA COLLECTION TOOLS

ADAPTED EGRA WITH ADDITIONAL PRELITERACY ITEMS

EGRA+prelit -Reading Assessment Tool

General instructions:

It is important that you create an atmosphere of fun with the child being evaluated by starting with them a simple conversation about topics that interest them (see example below)

Let them feel that this assessment is like a game so they will enjoy it and is not a difficult task.

It is very important that you ONLY read the content of the boxes, aloud clearly and slowly

Good morning. My name is ____ I live in _____. I want to talk to you about myself, I have of children, their age; I have at home..... the sports I do..... etc.]

1. Tell me about yourself and your family? [wait for response; If the student is not excited to talk, ask him/her question number 2. If he/she speaks comfortably, move to the verbal consent paragraph].
2. What game do you like?

- Allow me to tell you why I am with you today. I work for the Ministry of Education, and I try to understand how children learn to read. You have been randomly selected to do this test.
- I would love for you to cooperate with me in this process. But if you don't want to share, you can.
- We will play a reading game where I will ask you to read some letters, some words and a short story out loud.
- I will use this watch to calculate the time you need to read.
- This is not an exam, and it has no effect on your school scores.
- I will ask you some other questions about your family.
- I will not write your name on the test paper. No one will see your answers to them.
- Again, you are under no obligation to participate if you don't want to, and if we start and you don't answer a question, that's fine.

Do you have a question? Are you ready?

Verbal consent

If you get the child's oral consent, put an (X) in this box Yes .1

If you do not get approval, thank the child and move on to the next child and use the same for



1. Date of Assessment:	Day: _____ Month: _____ Year: _____
2. Governate:	
3. MOE Field Directorate	
4. School Name	
5. National ID for School	
6. Student's Shift	<input type="radio"/> One shift <input type="radio"/> Morning Shift <input type="radio"/> Evening Shift
7. Name of Evaluator	
8. Evaluator Code	
9. Grade	<input type="radio"/> Second Grade <input type="radio"/> Third Grade
10. Division	
11. Child Number	
12. Child's Date of Birth	Month: _____ Year: _____
13. Child's Gender	<input type="radio"/> Girl <input type="radio"/> Boy
14. Exam Start Time:	_____ : _____ Choose One time slot: <input type="radio"/> Morning <input type="radio"/> Evening



Section 1 Print awareness	60 Seconds
<p>— This is a book. Can you take it from me and put it into your hands? Then I’m going to ask some questions about the book.</p> <p>Ensure the book is in the hands of the child.</p>	
<p>— Let's Begin</p> <p>— With the book in your hands, can you show me the front of the book?</p> <p>[Include here instructions to the enumerator to indicate if the child correctly or incorrectly identified the front of the book].</p> <p>— Thank you. Now can you open the book to the first page and point to where we can begin reading the story?</p> <p>[Include here instructions to the enumerator to indicate if the child correctly or incorrectly identified (a) the first page of the book and (b) where to begin reading.</p>	<p>Take the book back from the child before moving on to the next section</p>

Section 2 Oral vocabulary	60 Seconds
<p>Let’s play a few more naming games now. Think about the different things that you can eat. Name as many things that you can eat as you can.</p> <p>Clearly put a tick (/) in the box for each correct word.</p> <p>MULTI-SELECT</p> <p>01 Word one is correct 02 Word two is correct 03 Word three is correct 04 Word four is correct 05 Word five is correct 06 Word six is correct 07 Word seven is correct 08 Word eight is correct 09 Word nine is correct 10 Word ten is correct 11 Child was unable to say any correct words</p>	<p>After 30 seconds, you will tell the child to 'stop'.</p> <p></p> <p>The Early stop rule:</p> <p>If the child hesitates to name things you can eat after 5 seconds, say “thank you” and stop the exercise.</p>
<p>Now I would like you to name as many animals as you can.</p>	<p>After 30 seconds,</p>

<p>Clearly put a tick (/) in the box for each correct word.</p> <p>MULTI-SELECT</p> <p>01 Word one is correct</p> <p>02 Word two is correct</p> <p>03 Word three is correct</p> <p>04 Word four is correct</p> <p>05 Word five is correct</p> <p>06 Word six is correct</p> <p>07 Word seven is correct</p> <p>08 Word eight is correct</p> <p>09 Word nine is correct</p> <p>10 Word ten is correct</p> <p>11 Child was unable to say any correct words</p>	<p>you will tell the child to 'stop'.</p> <p></p> <p>The Early stop rule:</p> <p>If the child hesitates to name animals after 5 seconds, say “thank you” and stop the exercise.</p>
--	--

Section 3 Recognise Letter Names	30 Seconds
<p>☑☑We will do some alphabet letter games now. Do you see these letters? I would like you to tell me the name of each letter. It’s ok if you don’t know all of them.</p> <p>☑ Now let's do this exercise: tell me the name of this letter [and point to the first letter]:</p> <p>If a child gets stuck for more than 5 seconds, mark as incorrect and encourage the child to continue, pointing to the next letter and say: “now let’s try this one.”</p> <p>[[دار - زيت - خيوط - ملعب - شتاء]]</p> <p>Clearly put a tick (/) on any mistake the child makes.</p> <p>In the event that the child corrects him/herself, circle the sign (/) that you previously made for him/her.</p> <p>Put a tick (/) on the last letter the child identifies.</p>	<p>If the child hesitates to name the letter for more than 3 seconds, point to the next letter and say: “Let’s continue, please.”</p> <p></p>

	<p>The Early stop rule:</p> <p>If you mark the first three answers as wrong and the child does not correct any mistakes, say “thank you” and stop the exercise.</p>
--	---

Section 4: Read high frequency words	
<p>— This is a sheet that includes words. I’d like you to read as many of them as you can. For example, we read this word [point to the word “[add example word]” as in the word “[example]”.</p> <p>— Let's Begin</p> <p>— Can you try the next word?</p> <p>[add 5 high frequency words, use one high frequency word as the example].</p> <p>Clearly put a tick (/) on any mistake the child makes.</p> <p>In the event that the child corrects him/herself, circle the sign (/) that you previously made for him/her.</p>	

Section 5: Recognize Letter Sounds	60 Seconds
<p>☐ This is a sheet of Arabic letters and movements. Read as many of them as you can (read the letter's sound, not its name). For example, the sound of this letter [indicate the letter "l"] is "for," as in the word "playing."</p> <p>☐ Now let's do this exercise: tell me the sound of this letter [and point to the letter K]:</p> <p>Good, the sound of this letter is "k."</p> <p>The sound of this letter is "K".</p> <p>☐ Let's try another example: Tell me the sound of this movement [point to the aperture]:</p> <p>Well done, the sound of this movement is "–"</p> <p>The sound of this movement is "—"</p> <p>☐ Did you understand what is required from you?</p> <p>When I tell you "Let's get started," read the sound of the letters as accurately and as quickly as possible. We'll start from here and continue this way [point to the first letter on the first line, and trace it with your finger on the letters in the entire first line]. are you ready?</p> <p>☐ Let's Begin</p>	<p>After 60 seconds, you will tell the child to 'stop.'</p> <p>If the child hesitates to read the letter for more than 3 seconds, point to the next letter and say: "Let's continue, please."</p> <p>☐</p> <p><u>The Early stop rule:</u></p>
<p>Clearly put a tick (/) on any mistake the child makes.</p> <p>In the event that the child corrects himself, circle the sign (/) that you previously made for him.</p> <p>Put a tick (/) on the last letter the child reads.</p> <p>Example: for K</p>	<p>If you mark all the answers in the first line as wrong and the child does not correct any of his mistakes, say "thank you" and stop the exercise. Put an (X) in the box at the bottom of the page and go to the next exercise.</p>

	10	9	8	7	6	5	4	3	2	1
(10)	ذ	ه	ح	ة	ج	ة	ه	ف	ت	ب
(20)	ـ	و	ق	ز	س	خ	ه	ف	ة	ب
(30)	ق	ص	ب	غ	ق	و	ز	خ	ة	ص
(40)	ذ	ظ	ي	ن	و	ه	ط	ش	ض	ع
(50)	ء	خ	ش	ق	ن	ي	غ	د	ب	غ
(60)	م	ن	ط	س	ف	ث	ح	ذ	ض	ب
(70)	ع	ط	ث	ق	ج	ه	ض	ه	ح	ت
(80)	و	ص	ج	د	ذ	خ	ه	ج	ظ	ل
(90)	س	ل	ذ	خ	ث	ء	م	خ	ء	ز
(100)	خ	ه	ه	ض	ع	ح	و	ذ	ب	ـ

Remaining time of exercise time (number of seconds)

Check this box (X) in case you have left this part of the assessment

Because the child did not read any of the words in the first line correctly

Section 6: Read the character syllable	60 Seconds
<p><input checked="" type="checkbox"/> This is a sheet that includes Arabic syllables and movements, read as many of them as you can (read the passage). For example, we read this passage [point to the syllable "a'a"] as in the word "aa".</p> <p><input checked="" type="checkbox"/> Now let's do this exercise: read this passage [point to the syllable "ra"]: Good, we read this passage like this "Ra" We read this passage "Ra"</p> <p><input checked="" type="checkbox"/> Let's try another example: read this passage [point to the passage]: Well done, we read this passage like this "C"</p>	<p>After 60 seconds, you will tell the child to 'stop.'</p> <p>If the child hesitates to read the letter for more than 3 seconds, point to the next letter and say:</p>

<p>The sound of this movement is "Su"</p> <p>☒ Did you understand what is required from you?</p> <p>When I tell you "Let's get started," read the syllable accurately and as quickly as possible. We'll start from here and continue this way [point to the syllable in the first line, and trace it with your finger on the syllable in the entire first line]. are you ready?</p> <p>☒ Let's Begin</p>	<p>"Let's continue, please."</p> <p>☒</p> <p><u>The Early stop rule:</u></p> <p>_If you mark all the answers in the first line as wrong and the child does not correct any of his mistakes, say "thank you" and stop the exercise. Put an (X) in the box at the bottom of the page and go to the next exercise.</p>
<p>Clearly put a tick (/) on any mistake the child makes.</p> <p>In the event that the child corrects himself, circle the sign (/) that you previously made for him.</p> <p>Put a tick (/) on the last letter the child reads.</p> <p>Example: for K</p>	

	10	9	8	7	6	5	4	3	2	1
(10)	دي	فو	بِ	هِ	مي	زِ	وَقِي	جِب	تِي	ظَ
(20)	ءِ	زُ	كُ	قو	رس	قِي	رِ	الِ	قَب	هُ
(30)	مَ	إِ	كو	ذي	ظَ	جا	حو	دي	هَ	يَخِ
(40)	طو	لِي	ةِ	ذا	حَا	دَا	رو	تِ	رِ	هَمِ
(50)	نِ	دَ	بِ	كث	مِخ	نَ	مِنَ	صَوُ	دِ	دا
(60)	ها	لَتَ	قِي	ضِي	عِنَ	را	ةِ	جا	دا	فِي
(70)	ضِي	صَوُ	أَلِ	أَنَ	عُصَ	ةِ	سُ	حَتَ	مُ	زُ
(80)	قا	يِنَ	رِي	خَ	قَةِ	دَ	عُ	ظَ	هَ	خِي
(90)	بِخِ	عَيِ	رِ	كِنَ	إِ	تُ	مَزُ	مَنَ	عِنَ	عا
(100)	عا	رخ	حا	أَجِ	تَنَ	فِي	هق	كو	با	دز

Remaining time of exercise time (number of seconds)

Check this box (X) in case you have left this part of the assessment

Because the child did not read any of the words in the first line correctly

Section 7 Part A: Read a text orally	Section 7 Part B: Reading comprehension
<p>☒ This is a short story, focus well and read it correctly, aloud and as quickly as possible.</p> <p>- When you're done, I'll ask you some questions about what you've read. Did you understand what is required of you? - When I tell you, "Let's begin."</p> <p>- Start reading. ready? Let's Begin</p>	<p>Pull the text of the story in front of the child and ask them the questions below.</p> <p>Leave the child maximum 15 seconds to answer each question.</p> <p>Ask the question corresponding to each line the child has read until you reach the line with the mark (), which indicates where the child stopped reading.</p>
<p>After 60 seconds, you will tell the child to 'stop.'</p> <p>If the child hesitates to read the letter for more than 3 seconds, point to the next letter and say: "Let's continue, please."</p> <p>☒ <u>The Early stop rule:</u></p>	<p>☒☒ I will now ask you some questions about the story that I read. Answer the questions correctly.</p>

<p>_If you mark all the answers in the first line as wrong and the child does not correct any of his mistakes, say “thank you” and stop the exercise. Put an (X) in the box at the bottom of the page and go to the next exercise.</p>				
<p>Clearly put a tick (/) on any mistake the child makes. Tick the last word the child reads with (l)</p>	<p>Put a tick (X) in the box that corresponds to the child's answer, and then move on to the next question.</p>			
		<p>No answer</p>	<p>Not Correct</p>	<p>Correct</p>
<p>Dima is a student in the third grade. She likes to read books and writing stories 10</p>	<p>What does Dima like? <u>Reading books and writing stories</u></p>			
<p>Dima went with her classmate Farah to the school library 18</p>	<p>Where did dima go with her classmate? <u>To the school library</u></p>			
<p>Farah read a book about space, and Dima chose a story about birds 28</p>	<p>What did Farah read? <u>A book about space</u></p>			
<p>Farah asked: Why do you like reading stories? 34</p>	<p>What did Farah ask her classmate? <u>Why do you like reading stories?</u></p>			
<p>Dima answered confidentially: I dream about becoming a writer for children 42</p>	<p>Why does Dima dream about becoming a writer for children?</p>			

	<u>Because she</u> <u>likes reading</u> <u>and writing,</u> <u>because she</u> <u>wants to be</u> <u>famous, to</u> <u>have more</u> <u>networks,</u> <u>because she</u> <u>likes children</u>			
--	---	--	--	--

Remaining time of exercise time (number of seconds):

Put a tick (X) in this box if you stop this part of the assessment because the child did not read any word in the first line correctly.

LRF! - Pupil survey

ENUMERATOR TO COMPLETE, SINGLE CODE

1. Is this a control or intervention school?
 - a. Control
 - b. Intervention

ASK ALL, SINGLE CODE PER ROW

2. How do you feel about...?

Playing with toys			
Watching TV			
Reading story books			
Playing games with friends			

ASK ALL, SINGLE CODE

3. Do you have story books at home?
 - a. Yes
 - b. No
 - c. Don't know **(Do not read out)**

ASK ALL, [more]=Q3=A, SINGLE CODE

4. Would you like to have [more] story books at home?
 - a. Yes
 - b. No
 - c. Don't know **(Do not read out)**

ASK ALL, SINGLE CODE

5. Are you able to take story books home from your school or a library?
 - a. Yes
 - b. No
 - c. Don't know **(Do not read out)**

ASK IF Q5 = a. yes, SINGLE CODE

6. How often do you take books home from your school or a library?
- a. Every day
 - b. Every week
 - c. Less than every week
 - d. Never
 - e. Don't know **(Do not read out)**

ASK ALL, SINGLE CODE

7. How often do you read story books at home?
- a. Every day
 - b. Every week
 - c. Less than every week
 - d. Never
 - e. Don't know

ASK Q7=A OR B, MULTI CODE, DO NOT READ OUT AND SELECT AS APPROPRIATE

8. Who do you read with at home?
- a. Parents
 - b. Grandparents
 - c. Brothers/Sisters
 - d. Other relatives
 - e. Friends
 - f. No one - by myself

ASK Q7=A or B, SINGLE CODE

9. How long do you usually read for at home? **If asked, this is amount of time per day.**
- a. Less than 15 minutes
 - b. 15-30 minutes
 - c. More than 30 minutes
 - d. Don't know **(Do not read out)**

ASK ALL, MULTI CODE

10. Which of these do you or your family have at home?
- a. A TV
 - b. A computer/laptop
 - c. A tablet (like an iPad)
 - d. A phone that connects to the internet
 - e. None of these **(SINGLE CODE)**

ASK ALL, SINGLE CODE

11. Can your family access the internet at home?

- a. Yes
- b. No
- c. Don't know (**Do not read out**)

ASK Q11=A, SINGLE CODE

12. Are you allowed to use the internet at home, either by yourself or with other family members?

- a. Yes
- b. No
- c. Don't know (**Do not read out**)

ASK Q12=A, SINGLE CODE

13. Do you ever use the internet to read stories? **If yes, ask if this is on a phone that connects to the internet.**

- a. Yes – on a phone that connects to the internet
- b. Yes – on something else
- c. No
- d. Don't know (**Do not read out**)

ASK Q13=A OR B, MULTI CODE, DO NOT READ OUT AND SELECT AS APPROPRIATE

14. Who do you read these stories with?

- a. Parents
- b. Grandparents
- c. Brothers/Sisters
- d. Other relatives
- e. Friends
- f. By myself

FOR ENDLINE ONLY

ASK Q1=B, SINGLE CODE

15. Do you have a Let's Read Fluently practice book?

- a. Yes
- b. No
- c. Don't know (**Do not read out**)

ASK Q15=A, SINGLE CODE

16. Do you take the practice book home?

- a. Yes
- b. No

- c. Don't know **(Do not read out)**

ASK Q16=A, SINGLE CODE

17. Do you read the practice book at home?
- a. Yes
 - b. No
 - c. Don't know **(Do not read out)**

ASK Q17=A, MULTI CODE, DO NOT READ OUT AND SELECT AS APPROPRIATE

18. Who do you read the practice book with?
- a. Parents
 - b. Grandparents
 - c. Brothers/Sisters
 - d. Other relatives
 - e. Friends
 - f. By myself

ASK Q15=A, SINGLE CODE PER ROW

19. How do you feel about...?

Reading or working from the practice book at school			
Reading or working from the practice book at home			

LRF! - Teacher Survey

SECTION 1: Background

ASK ALL, SINGLE CODE

1. Which version of LRF are you delivering?
 - a. Whole class
 - b. Catch-up

SECTION 2: Training

ASK ALL, SINGLE CODE

2. Did you attend the Let's Read Fluently training ran by QRTA?
 - a. Yes
 - b. No

Q2=A, SINGLE CODE PER ROW

3. To what extent do you agree or disagree with the following:

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
The training was relevant to my work	A	B	C	D	E
The training was delivered well	A	B	C	D	E
The training gave me a good understanding of Let's Read Fluently	A	B	C	D	E
The training equipped me with the skills and knowledge to deliver Let's Read Fluently	A	B	C	D	E
The training was the right length	A	B	C	D	E
I have been able to apply the training in lessons	A	B	C	D	E

ASK Q2/5=D or E, SINGLE CODE

4. Was the training...?
 - a. Too long
 - b. Too short

Q2=B, MULTICODE

5. Did you receive any training on Let's Read Fluently? If yes, who from?

- a. Yes – The School Principal
- b. Yes – Other teachers
- c. Yes – Someone else (please specify)
- d. No, I did not receive training **SINGLE CODE**

SECTION 3: Coaching

ASK ALL, SINGLE CODE

6. How many coaching sessions have you received?
- a. 4 or more
 - b. 3
 - c. 2
 - d. 1
 - e. None

ASK Q6=A, B, C, or D, SINGLE CODE PER ROW

7. Did the following happen at the last coaching session?

	Yes	No
The full lesson was observed	A	B
Clear action points were provided and agreed upon	A	B

Q6=A, B, C, or D, SINGLE CODE PER ROW

8. To what extent do you agree or disagree with the following:

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
The coaching sessions were delivered well	A	B	C	D	E
The coaching sessions were useful	A	B	C	D	E
The number of coaching sessions was sufficient	A	B	C	D	E
The coaching sessions supported my teaching of Let’s Read Fluently	A	B	C	D	E

SECTION 4: Implementation

ASK ALL, SINGLE CODE

9. On average, how many Let’s Read Fluently sessions did you deliver per week?

- a. 4 or more
- b. 3
- c. 2
- d. 1
- e. None

Q9=C,D OR E, MULTICODE

10. Is there anything that prevented you from delivering more sessions?

- a. Not enough time in the school day
- b. Other lessons are more important
- c. Needed more time to plan Let’s Read Fluently sessions
- d. Children didn’t have their practice books
- e. The school/students shifted to hybrid or online learning
- f. Something else (please specify)
- g. Nothing prevented me from delivering more sessions **SINGLE CODE**

ASK ALL, SINGLE CODE PER ROW

11. To what extent do you agree or disagree with the following:

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
The practice book is a useful aid in the classroom	A	B	C	D	E
Students were engaged with the Let’s Read Fluently sessions	A	B	C	D	E
Let’s Read Fluently is easy to implement	A	B	C	D	E

SECTION 5: Parental engagement

ASK ALL, MULTI CODE

12. How, if at all, are you engaging with parents on Let’s Read Fluently?

- a. Through WhatsApp
- b. Parental orientation session
- c. Talking to parents about LRF before and after school
- d. In another way (please specify)
- e. I’m **not** engaging with parents **SINGLE CODE**

ASK Q12=A, SINGLE CODE

13. How often do you send parents messages about Let’s Read Fluently on WhatsApp?

- a. Every day

- b. A few times a week
- c. Once a week
- d. Less than once a week
- e. Not at all

ASK ALL, SINGLE CODE

14. How often, if at all, are children taking the practice book home?
- a. Every day
 - b. A few times a week
 - c. Once a week
 - d. Less than once a week
 - e. Not at all

ASK Q14=D AND E, MULTICODE

15. Why are children not taking the practice book home?
- a. There are not enough copies of the book
 - b. They forget to bring the book back
 - c. They won't use the book at home
 - d. Another reason (please specify)

ASK Q14=A,B,C, OR D, SINGLE CODE

16. On average, how many pages are you assigning as homework in the practice book?
- a. 5 or more
 - b. 3-4 pages
 - c. 1-2 pages
 - d. I'm not assigning any pages

ASK Q14=A,B, C,OR D, SINGLE CODE

17. What proportion of students are generally completing the assigned pages?
- a. All of them
 - b. Over half
 - c. Around half
 - d. Less than half
 - e. None of them
 - f. Not sure

ASK Q14=A,B,C, OR D, SINGLE CODE

18. Do parents know the children are taking the practice book home?
- a. Always
 - b. Most of the time
 - c. Sometimes
 - d. Never

e. Not sure

SECTION 6: Overall thoughts and perceived outcomes

ASK ALL, SINGLE CODE PER ROW

19. To what extent do you agree or disagree with the following:

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
Children's literacy has improved due to Let's Read Fluently	A	B	C	D	E
Children's reading comprehension has improved due to Let's Read Fluently	A	B	C	D	E
Children are more engaged in reading due to Let's Read Fluently	A	B	C	D	E
Children are more confident in reading due to Let's Read Fluently	A	B	C	D	E
I will continue to use the skills and approaches of Let's Read Fluently in my teaching after this semester	A	B	C	D	E

Focus Group Discussion Guide: Pupils

Let's Read Fluently!' (LRF) is designed to address issues with literacy among Arabic students. The intervention centers around a student practice book designed to reflect evidence about how Arabic reading fluency is best acquired. Teachers receive one-day training from the Queen Rania Teacher Academy (QRTA) and up to three follow-up coaching sessions to support them to use the practice book.

Teaching and Learning approach (W/C) and Literacy Catch-Up (C/U). Both run for one semester. W/C targets students in Grade 1 and is delivered to classes via three weekly 30-minute classes. The teacher adopts a 'I do', 'we do', 'you do' pedagogical approach using the practice book. C/U targets the 20% of lowest-achieving students in Grades 1 to 3 and is delivered through small group tuition (three weekly sessions).

The Queen Rania Foundation has commissioned a consortium which includes NatCen and Integrated to carry out a pilot evaluation of LRF. The aim of the evaluation is to explore delivery, understand the feasibility of different evaluation methods and inform learning for the planning of an effectiveness trial.

INTRODUCTION

Aim: to explain the aims of the research, how the interview will be conducted and how the data will be used.

- Introduction to researcher [NAME]. Thank you for helping us today.
- We will discuss:
 - Your thoughts and feelings about Let's Read Fluently lessons
 - Your thoughts and feelings about the practice book
- We will not share what you say with your teachers or parents.
- You can say if you don't want to take part.
- We would like to audio-record the discussion.
- The FGD will last up to 40 minutes
- Any questions?

Date of FGD:

Implementation model (C/U or W/C):

School/s:

1. Children's background [10 minute]

Aim: To ease participant into conversation, get pupils talking about themselves and their schoolwork

- Background of children
 - Age/Grade
 - Favourite lessons at school
 - Whether likes reading/ listening to story books/ favourite stories

2. Experience and opinion of LRF lessons [10 minutes]

Aim: To get pupils talking about LRF lessons and what they think of them

- What do you like about Let's Read Fluently lessons at school?
 - PROBE: Why?
- What do you not like?
 - PROBE: Why?
- Do you find Let's Read Fluently lessons easy or difficult?
 - What parts are most difficult? (PROBE: Why?)
- How are they different from your normal literacy lessons?

3. Use of the practice book [5 minutes]

Aim: To get pupils talking about the LRF practice book and how they use it

- Whether has and uses the LRF practice book?
 - Do you have your own copy of the LRF practice book?
 - PROBE: Tell me about when/ where you use it?
 - Where do you like to use it most? (PROBE: home/school?)
 - If at home; does anyone help you? (PROBE: parents, siblings, friends, grandparents)
 - How do they help?

4. Opinion of the practice book [10 minutes]

Aim: To understand pupils' thoughts about the LRF practice book

- What do you like about the practice book?
 - PROBE: Why?
- What do you not like about the practice book?
 - PROBE: Why?
- (IF TIME) What do you find difficult about the practice book?
 - PROBE: Which part of it is difficult?
- (IF TIME) What do you find easy about the practice book?
 - PROBE: Which part is easy?

5. Progress in learning [5 minutes]

Aim: To see if pupils think their literacy is improving and what is helping them learn to read

- What helps you with learning to read?
 - How does that help?
 - PROBE: Does LRF help you learn to read?
 - Why/Why not?

Focus Group Discussion Guide: Parent (treatment group - whole class model)

Let's Read Fluently!' (LRF) is designed to address issues with literacy among Arabic students. The intervention centers around a student practice book designed to reflect evidence about how Arabic reading fluency is best acquired. Teachers receive one-day training from the Queen Rania Teacher Academy (QRTA) and up to three follow-up coaching sessions to support them to use the practice book.

Teaching and Learning approach (W/C) and Literacy Catch-Up (C/U). Both run for one semester. W/C targets students in Grade 1 and is delivered to classes via three weekly 30-minute classes. The teacher adopts a 'I do', 'we do', 'you do' pedagogical approach using the practice book. C/U targets the 20% of lowest-achieving students in Grades 1 to 3 and is delivered through small group tuition (three weekly sessions).

The Queen Rania Foundation has commissioned a consortium which includes NatCen and Integrated to carry out a pilot evaluation of LRF. The aim of the evaluation is to explore delivery, understand the feasibility of different evaluation methods and inform learning for the planning of an effectiveness trial.

INTRODUCTION

Aim: to explain the aims of the research, how the interview will be conducted and how the data will be used.

- Introduction to researcher [NAME]. Thank you for agreeing to take part.
- Introduction to Integrated - we have been commissioned by the QRF (with NatCen) to conduct an independent evaluation of the pilot intervention 'Let's Read Fluently'.
- Explanation of research – in addition to speaking to them and parents at other participating schools, we are speaking to pupils, teachers and other stakeholders. Have also conducted classroom observations.
- Explanation of FGD, we will discuss:
 - Their understanding of the program
 - Their usage of the LRF practice book and general help with schoolwork
 - Their insights into perceived impacts
- Participation is voluntary – they can choose not to discuss any issue and can withdraw from the interview at any point.
- The information they provide will be used to write a findings report at the end of this school year for QRF.
- We are conducting this research independently and all information shared will be treated confidentially.
 - Caveats to anonymity – although we will take steps to ensure anonymity and no individual or organisation will be named in the report, it may be that because of their

the fairly small sample of schools and the small number of pupils in the program, they may recognise them from what they have said.

- We would like to audio-record the discussion, so we have an accurate record of what is said.
- The FGD will last up to 60 minutes
- Any questions?

Date of FGD:

School/s:

1. Introduction to pupils and involvement in schoolwork [10 minutes]

Aim: To ease participant into conversation, background of their children and involvement in their schoolwork.

Background

- How old are your children? What grades are they in?
- Do your children like reading?

Involvement in schoolwork

- Do you help with your children's schoolwork or homework?
 - What do you do? (PROBE: Is there anything particular they need help with?)
 - How do you get involved? (I.e. helping them when they're stuck, reviewing their homework.)
 - How often do you do this?
 - How do you find it? (PROBE: What do you find easy? What do you find difficult?)
 - IF NOT, why do you not get involved? (PROBE: Is it too difficult? Do you have time?)

2. LRF program [10 minutes]

Aim: To explore awareness and understanding of the LRF program.

Knowledge

- Are you aware of the Let's Read Fluently program?
 - What do you know about it?
 - Do you know how it is different from standard literacy lessons?
 - IF YES: How is it different?
 - How did you find out about it?
 - Were you invited to attend an orientation session? (PROBE: If yes, did you attend?)
 - IF ATTENDED: What did you think of the orientation session?
- Is there anything you like about the program?

- Is there anything you dislike about the program?

3. LRF practice book [20 minutes]

Aim: To explore awareness, usage, and perceptions of the LRF practice book.

Awareness

- Are you aware of the Let's Read Fluently practice book?
 - Have you seen the workbook?
 - What do you like about it?
 - What do you not like about it?

Usage

- Do your children bring the practice book home from school?
- How, if at all, do you use the practice book with your children?
 - What do you do? Can you describe using the book to me?
 - How many pages do you typically do? (PROBE: Why this number?)
 - How long do you typically do this for?
 - How often do you do this?
- What information, if any, did you receive on how to use the practice book?
 - PROBE: Information sheet, training session, spoke to teacher, WhatsApp, QR code
 - IF USED WHATSAPP: How have you used WhatsApp?
 - IF USED QR CODE: How have you used the QR code?
 - What information did you receive?
 - Have you done anything differently? (PROBE: Why?)
 - Is there any other information you would like?
 - IF NOT, what made you decide to use the practice book in the way you have?

Perceptions

- How easy or difficult do you find the practice book to use? (PROBE: Why?)
 - Is there anything you find difficult about using the practice book?
 - Is there anything that you would change?
- Overall, what are your thoughts on the practice book?

4. Impacts of LRF [20 minutes]

Aim: To explore perceived impacts of the LRF program.

General impacts

- Have you noticed any changes in your children since the start of the LRF program?
- Why do you think this?
- When have you noticed this? (PROBE: When using the practice book or reading more generally?)

Impacts on reading

- Has the LRF program had any impact on your children's reading?
- Why do you think this?
- When have you noticed this? (PROBE: When using the practice book or reading more generally?)

- Has LRF changed the way your children feel about reading?
- PROBE: More confident? More enthusiastic? Any negative effects?
- Why do you think this?
- When have you noticed this? (PROBE: When using the practice book or reading more generally?)

Other impacts

- (IF TIME) Has the LRF program had any impact on your children's writing?
- Why do you think this?
- When have you noticed this?
- Has the LRF program had any impact on your children's education?
- Any impacts on their other subjects?
- Any change in how they feel about school?

- Thanks and close.

Focus Group Discussion Guide: Teachers (treatment group – whole class model)

Let's Read Fluently!' (LRF) is designed to address issues with literacy among Arabic students. The intervention centers around a student practice book designed to reflect evidence about how Arabic reading fluency is best acquired. Teachers receive one-day training from the Queen Rania Teacher Academy (QRTA) and up to three follow-up coaching sessions to support them to use the practice book.

Teaching and Learning approach (W/C) and Literacy Catch-Up (C/U). Both run for one semester. W/C targets students in Grade 1 and is delivered to classes via three weekly 30-minute classes. The teacher adopts a 'I do', 'we do', 'you do' pedagogical approach using the practice book. C/U targets the 20% of lowest-achieving students in Grades 1 to 3 and is delivered through small group tuition (three weekly sessions).

The Queen Rania Foundation has commissioned a consortium which includes NatCen and Integrated to carry out a pilot evaluation of LRF. The aim of the evaluation is to explore delivery, understand the feasibility of different evaluation methods and inform learning for the planning of an effectiveness trial.

INTRODUCTION

Aim: to explain the aims of the research, how the interview will be conducted and how the data will be used.

- Introduction to researcher [NAME]. Thank you for agreeing to take part.
- Introduction to Integrated - we have been commissioned by the QRF (with NatCen) to conduct an independent evaluation of the pilot intervention 'Let's Read Fluently'.
- Explanation of research – in addition to speaking to them and teachers at other participating schools, we are speaking to pupils, parents and other stakeholders. Have also conducted classroom observations.
- Explanation of FGD, we will discuss:
 - Their understanding of the program
 - Their experiences of delivering the program to date
 - Their insights into perceived impacts
- Participation is voluntary – they can choose not to discuss any issue and can withdraw from the interview at any point.
- The information they provide will be used to write a findings report at the end of this school year for QRF.
- We are conducting this research independently and all information shared will be treated confidentially.
 - Caveats to anonymity – although we will take steps to ensure anonymity and no individual or organisation will be named in the report, it may be that because of their

role and the fairly small sample of schools, others familiar with the intervention may recognise them from what they have said.

- At the end of the interview, let us know if they would be happy for us to include quotes or would like to have anything removed.
- We would like to audio-record the discussion, so we have an accurate record of what is said.
- The FGD will last up to 60 minutes
- Any questions?

Date of FGD:

Implementation model (C/U or W/C):

School/s:

1. Participant background [5 minutes]

Aim: To ease participant into conversation, explore role and responsibilities, their understanding of the intervention and examine the school context

Role/responsibilities

- Brief overview of role/responsibilities (PROBE: class teacher or resource room teacher)
 - At school
 - In relation to the pilot

School context

- Overview of school
 - Size
 - Number of pupils/ staff

Understanding of program

- How they would describe the program
 - Key aims as they understand them
 - Initial expectations for the program

2. Reflections on training [5 minutes]

Aim: To explore experiences and usefulness of LRF training sessions. Understanding what worked well/ could be improved.

Training

- Feedback on QRTA's LRF training session
 - Format and delivery of training (PROBE: what worked well/ less well)

- Any issues with the training session (PROBE: issues with trainers, practical arrangements, materials etc.)
- Met expectations? (PROBE: anything not covered that should have been)
- PROBE: reasons for non-attendance if relevant

Application in school

- Application of training in classroom setting
 - Elements of training that worked well in practice (PROBE: why)
 - Elements that didn't work/ were less successful in the classroom (PROBE: why)

Reflections on training

- Anything that should be changed about the training? (PROBE: what and why)

3. Reflections on coaching [5 minutes]

Aim: To explore experiences and usefulness of LRF coaching sessions. Understanding what worked well/ could be improved.

Coaching

- Feedback on QRTA's LRF coaching sessions
 - Format and delivery of coaching (PROBE: how many coaching sessions have they received, what worked well/ less well)
 - Any issues with the coaching sessions (PROBE: issues with coaches, practical arrangements, materials etc.)
 - Met expectations? (PROBE: anything not covered that should have been)
 - Anything that should be changed about the coaching sessions? (PROBE: what and why)

4. Reflections on LRF in practice [10 minutes]

Aim: To explore how LRF was implemented in the classroom and any issues encountered by teachers, with either the approach or the practice book. Also, to explore pupil engagement with LRF and the practice book.

Implementing the LRF model in the classroom

- **Implementation in class**
 - How has the LRF model been implemented in the classroom? (PROBE: frequency of sessions – delivered 3 times a week, followed 'I do', 'we do', 'you do')
 -
 - Differences between LRF and usual literacy teaching (PROBE: positive and negative comparisons to usual practice)
 - Anything about the model that doesn't work well in the classroom (PROBE: level of individualisation, issues with pupil engagement)
- **Using the practice book in class**

- (How) has the practice book been used in class? (PROBE: are students able/ keen to use the practice book)
- What about the practice book works well? (PROBE: any early reflections on impact for pupils)
- Any issues with the practice book? (PROBE: anything missing, issues with layout/ format, issues with content etc.)
- (How) could the practice book be improved for use in the classroom?

- **Pupil engagement**

- How have students responded to the LRF approach? (PROBE: reaction of students in the classroom)
- How engaged are students in the program? (PROBE: any examples teachers can give of engagement or lack of engagement).
- Any difference between pupils engagement (PROBE: subset of learners for whom LRF was more/ less effective, any students struggling more than others)?
- What about the program has encouraged/discouraged student engagement? (PROBE: what elements of the program do students respond to/ engage with the most?)
- Anything about LRF that could be changed to improve student engagement? (PROBE: why).

5. Engaging with parents [5 minutes]

Aim: To explore whether teachers were able to engage parents in LRF, and what engagement strategies were more/ less successful.

Contact with parents and home working

- **Making contact with parents**

- What has been done to try to engage parents of LRF pupils?
- Success and failures in terms of contacting/ engaging parents (PROBE: challenges connecting with parents – any effect positive or negative of LRF on engagement with parents)
- How do teachers feel about use of WhatsApp to contact parents? (PROBE: levels of success)

- **Use of practice book at home**

- (How) has the practice book been used at home (PROBE: how has that gone, have parents been supportive?)
- (How) have teachers supported home working (PROBE: have there been checks on work at home, how has that been done, how have pupils/parents responded?)
- What challenges have there been with practice book transfer between home and school?

6. Perceived impacts [5 minutes]

Aim: To explore perceived benefits of programme for students and schools

- Perceived benefit of the programme for pupils

Any effects of LRF on:

- Pupils' reading ability
 - Confidence with reading
 - Engagement with literacy learning
 - Any difference between pupils (PROBE: subset of learners for whom LRF was more/ less effective)?
- Perceived benefits of the programme for schools
 - Any unanticipated effects

7. Feedback on LRF program [5 minutes]

Aim: To explore perceived benefits of programme for students and schools

- Anything about the program that will be taken forward after the implementation period/ changed approach to teaching?
- Anything that should be done differently about the classroom delivery?
- Anything that should be changed in terms of the practice book?
- Any changes to parental engagement?
- Any recommendations to improve the approach overall?

Focus Group Discussion Guide: Parent (treatment group - C/U model)

Let's Read Fluently!' (LRF) is designed to address issues with literacy among Arabic students. The intervention centers around a student practice book designed to reflect evidence about how Arabic reading fluency is best acquired. Teachers receive one-day training from the Queen Rania Teacher Academy (QRTA) and up to three follow-up coaching sessions to support them to use the practice book.

Teaching and Learning approach (W/C) and Literacy Catch-Up (C/U). Both run for one semester. W/C targets students in Grade 1 and is delivered to classes via three weekly 30-minute classes. The teacher adopts a 'I do', 'we do', 'you do' pedagogical approach using the practice book. C/U targets the 20% of lowest-achieving students in Grades 1 to 3 and is delivered through small group tuition (three weekly sessions).

The Queen Rania Foundation has commissioned a consortium which includes NatCen and Integrated to carry out a pilot evaluation of LRF. The aim of the evaluation is to explore delivery, understand the feasibility of different evaluation methods and inform learning for the planning of an effectiveness trial.

INTRODUCTION

Aim: to explain the aims of the research, how the interview will be conducted and how the data will be used.

- Introduction to researcher [NAME]. Thank you for agreeing to take part.
- Introduction to Integrated - we have been commissioned by the QRF (with NatCen) to conduct an independent evaluation of the pilot programme 'Let's Read Fluently'.
- Explanation of research – in addition to speaking to them and parents at other participating schools, we are speaking to pupils, teachers and other stakeholders. Have also conducted classroom observations.
- Explanation of FGD, we will discuss:
 - Their understanding of the program
 - Their usage of the LRF practice book and general help with schoolwork
 - Their insights into perceived impacts
- Participation is voluntary – they can choose not to discuss any issue and can withdraw from the interview at any point.
- The information they provide will be used to write a findings report at the end of this school year for QRF.
- We are conducting this research independently and all information shared will be treated confidentially.
 - Caveats to anonymity – although we will take steps to ensure anonymity and no individual or organisation will be named in the report, it may be that because of their

the fairly small sample of schools and the small number of pupils in the program, they may recognise them from what they have said.

- We would like to audio-record the discussion, so we have an accurate record of what is said.
- The FGD will last up to 60 minutes
- Any questions?

Date of FGD:

School/s:

1. Introduction to pupils and involvement in schoolwork [10 minutes]

Aim: To ease participant into conversation, background of their children and involvement in their schoolwork.

Background

- How old are your children? What grades are they in?
- Do your children like reading?

Involvement in schoolwork

- Do you help with your children's schoolwork or homework?
 - What do you do? (PROBE: Is there anything particular they need help with?)
 - How do you get involved? (I.e. helping them when they're stuck, reviewing their homework.)
 - How often do you do this?
 - How do you find it? (PROBE: What do you find easy? What do you find difficult?)
 - IF NOT, why do you not get involved? (PROBE: Is it too difficult? Do you have time?)

2. LRF program [10 minutes]

Aim: To explore awareness and understanding of the LRF program.

Knowledge

- Are you aware of the Let's Read Fluently program?
 - What do you know about it?
 - Do you know how it is different from standard literacy lessons?
 - IF YES: How is it different?
 - How did you find out about it?
 - Were you invited to attend an orientation session? (PROBE: If yes, did you attend?)
 - IF ATTENDED: What did you think of the orientation session?

- Is there anything you like about the program?
- Is there anything you dislike about the program?

Catch up model

- Are any of your children involved in the catch-up version of the Let's Read Fluently program?
PROMPT: The catch-up version offers extra support to a small number of pupils with their literacy.
- What do you know about the catch-up version?
- Is there anything you like about the catch-up version?
- Is there anything you dislike about the catch-up version?

3. LRF practice book [20 minutes]

Aim: To explore awareness, usage, and perceptions of the LRF practice book.

Awareness

- Are you aware of the Let's Read Fluently practice book?
 - Have you seen the workbook?
 - What do you like about it?
 - What do you not like about it?

Usage

- Do your children bring the practice book home from school?
- How, if at all, do you use the practice book with your children?
 - What do you do? Can you describe using the book to me?
 - How many pages do you typically do? (PROBE: Why this number?)
 - How long do you typically do this for?
 - How often do you do this?
- What information, if any, did you receive on how to use the practice book?
 - PROBE: Information sheet, training session, spoke to teacher, WhatsApp, QR code
 - IF USED WHATSAPP: How have you used WhatsApp?
 - IF USED QR CODE: How have you used the QR code?
 - What information did you receive?
 - Have you done anything differently? (PROBE: Why?)
 - Is there any other information you would like?

- IF NOT, what made you decide to use the practice book in the way you have?

Perceptions

- How easy or difficult do you find the practice book to use? (PROBE: Why?)
 - Is there anything you find difficult about using the practice book?
 - Is there anything that you would change?
- Overall, what are your thoughts on the practice book?

4. Impacts of LRF [20 minutes]

Aim: To explore perceived impacts of the LRF program.

General impacts

- Have you noticed any changes in your children since the start of the LRF program?
 - Why do you think this?
 - When have you noticed this? (PROBE: When using the practice book or reading more generally?)

Impacts on reading

- Has the LRF program had any impact on your children's reading?
 - Why do you think this?
 - When have you noticed this? (PROBE: When using the practice book or reading more generally?)
- Has LRF changed the way your children feel about reading?
 - PROBE: More confident? More enthusiastic? Any negative effects?
 - Why do you think this?
 - When have you noticed this? (PROBE: When using the practice book or reading more generally?)

Other impacts

- (IF TIME) Has the LRF program had any impact on your children's writing?
 - Why do you think this?
 - When have you noticed this?
- Has the LRF program had any impact on your children's education?
 - Any impacts on their other subjects?
 - Any change in how they feel about school?

Thanks and close.

Focus Group Discussion Guide: Classroom Teachers (treatment group - C/U model)

Let's Read Fluently!' (LRF) is designed to address issues with literacy among Arabic students. The intervention centers around a student practice book designed to reflect evidence about how Arabic reading fluency is best acquired. Teachers receive one-day training from the Queen Rania Teacher Academy (QRTA) and up to three follow-up coaching sessions to support them to use the practice book.

Teaching and Learning approach (W/C) and Literacy Catch-Up (C/U). Both run for one semester. W/C targets students in Grade 1 and is delivered to classes via three weekly 30-minute classes. The teacher adopts a 'I do', 'we do', 'you do' pedagogical approach using the practice book. C/U targets the 20% of lowest-achieving students in Grades 1 to 3 and is delivered through small group tuition (three weekly sessions).

The Queen Rania Foundation has commissioned a consortium which includes NatCen and Integrated to carry out a pilot evaluation of LRF. The aim of the evaluation is to explore delivery, understand the feasibility of different evaluation methods and inform learning for the planning of an effectiveness trial.

INTRODUCTION

Aim: to explain the aims of the research, how the interview will be conducted and how the data will be used.

- Introduction to researcher [NAME]. Thank you for agreeing to take part.
- Introduction to Integrated - we have been commissioned by the QRF (with NatCen) to conduct an independent evaluation of the pilot program 'Let's Read Fluently'.
- Explanation of research – in addition to speaking to them and teachers at other participating schools, we are speaking to pupils, parents and other stakeholders. Have also conducted classroom observations.
- Explanation of FGD, we will discuss:
 - Their understanding of the program
 - Their experiences of delivering the diagnostic tool and engaging with pupils and parents
 - Their insights into perceived impacts
- Participation is voluntary – they can choose not to discuss any issue and can withdraw from the interview at any point.
- The information they provide will be used to write a findings report at the end of this school year for QRF.
- We are conducting this research independently and all information shared will be treated confidentially.

- Caveats to anonymity – although we will take steps to ensure anonymity and no individual or organisation will be named in the report, it may be that because of their role and the fairly small sample of schools, others familiar with the program may recognise them from what they have said.
- At the end of the interview, let us know if they would be happy for us to include quotes or would like to have anything removed.
- We would like to audio-record the discussion, so we have an accurate record of what is said.
- The FGD will last up to 30 minutes
- Any questions?

Date of FGD:

Implementation model (C/U or W/C):

School/s:

1. Participant background [5 minutes]

Aim: To ease participant into conversation, explore role and responsibilities, their understanding of the program and examine the school context

Role/responsibilities

- Brief overview of role/responsibilities (PROBE: class teacher or resource room teacher)
 - At school
 - In relation to the pilot

School context

- Overview of school
 - Size
 - Number of pupils/ staff

Understanding of program

- How they would describe the program
 - Key aims as they understand them
 - Initial expectations for the program

2. Reflections on LRF in practice [15 minutes]

Aim: To explore how LRF was implemented including use of the diagnostic tool. Also, to explore how classroom teachers perceive implementation and pupil/parental engagement with LRF.

Implementing the diagnostic tool

- **Implementation of the diagnostic tool**
 - How was the diagnostic tool implemented in the classroom? (PROBE: what worked well/less well)

- Is there anything that would have been better at identifying students.

Implementing the LRF model in the classroom

- **Implementation in class**

- Are you aware of how the LRF model has been implemented in the classroom? (PROBE: frequency of sessions – delivered 3 times a week, followed ‘I do’, ‘we do’, ‘you do’)
- Aware of the differences between LRF and usual literacy teaching? (PROBE: positive and negative comparisons to usual practice)

- **Pupil engagement**

- How have students responded to the LRF approach?
- Anything about LRF that could be changed to improve student engagement? (PROBE: why).

- **Parent engagement**

- Have you engaged with parents on LRF? (PROBE: any successes/challenges)
- Anything about LRF that could be changed to improve parental engagement? (PROBE: why).

3. Perceived impacts [5 minutes]

Aim: To explore perceived benefits of program for students and schools

- Perceived benefit of the program for pupils

Any effects of LRF on:

- Pupils’ reading ability
- Confidence with reading
- Engagement with literacy learning
- Any difference between pupils (PROBE: subset of learners for whom LRF was more/ less effective)?

- Perceived benefits of the program for schools

- Any unanticipated effects

4. Feedback on LRF program [5 minutes]

Aim: To explore feedback of the program for students and schools

- Anything about the program that will be taken forward after the implementation period/ changed approach to teaching?
- Any recommendations to improve the approach overall?

Focus Group Discussion Guide: Resource Room Teachers (treatment group – Catch Up model)

Let's Read Fluently! (LRF) is designed to address issues with literacy among Arabic students. The intervention centers around a student practice book designed to reflect evidence about how Arabic reading fluency is best acquired. Teachers receive one-day training from the Queen Rania Teacher Academy (QRTA) and up to three follow-up coaching sessions to support them to use the practice book.

Teaching and Learning approach (W/C) and Literacy Catch-Up (C/U). Both run for one semester. W/C targets students in Grade 1 and is delivered to classes via three weekly 30-minute classes. The teacher adopts a 'I do', 'we do', 'you do' pedagogical approach using the practice book. C/U targets the 20% of lowest-achieving students in Grades 1 to 3 and is delivered through small group tuition (three weekly sessions).

The Queen Rania Foundation has commissioned a consortium which includes NatCen and Integrated to carry out a pilot evaluation of LRF. The aim of the evaluation is to explore delivery, understand the feasibility of different evaluation methods and inform learning for the planning of an effectiveness trial.

INTRODUCTION

Aim: to explain the aims of the research, how the interview will be conducted and how the data will be used.

- Introduction to researcher [NAME]. Thank you for agreeing to take part.
- Introduction to Integrated - we have been commissioned by the QRF (with NatCen) to conduct an independent evaluation of the pilot program 'Let's Read Fluently'.
- Explanation of research – in addition to speaking to them and teachers at other participating schools, we are speaking to pupils, parents and other stakeholders. Have also conducted classroom observations.
- Explanation of FGD, we will discuss:
 - Their understanding of the program
 - Their experiences of delivering the program to date
 - Their insights into perceived impacts
- Participation is voluntary – they can choose not to discuss any issue and can withdraw from the interview at any point.
- The information they provide will be used to write a findings report at the end of this school year for QRF.
- We are conducting this research independently and all information shared will be treated confidentially.
 - Caveats to anonymity – although we will take steps to ensure anonymity and no individual or organisation will be named in the report, it may be that because of their

role and the fairly small sample of schools, others familiar with the program may recognise them from what they have said.

- At the end of the interview, let us know if they would be happy for us to include quotes or would like to have anything removed.
- We would like to audio-record the discussion, so we have an accurate record of what is said.
- The FGD will last up to 60 minutes
- Any questions?

Date of FGD:

Implementation model (C/U or W/C):

School/s:

1. Participant background [5 minutes]

Aim: To ease participant into conversation, explore role and responsibilities, their understanding of the program and examine the school context

Role/responsibilities

- Brief overview of role/responsibilities (PROBE: class teacher or resource room teacher)
 - At school
 - In relation to the pilot

School context

- Overview of school
 - Size
 - Number of pupils/ staff

Understanding of program

- How they would describe the program
 - Key aims as they understand them
 - Initial expectations for the program

2. Reflections on training [5 minutes]

Aim: To explore experiences and usefulness of LRF training sessions. Understanding what worked well/ could be improved.

Training

- Feedback on QRTA's LRF training session
 - Format and delivery of training (PROBE: what worked well/ less well)

- Any issues with the training session (PROBE: issues with trainers, practical arrangements, materials etc.)
- Met expectations? (PROBE: anything not covered that should have been)
- PROBE: reasons for non-attendance if relevant

Application in school

- Application of training in classroom setting
 - Elements of training that worked well in practice (PROBE: why)
 - Elements that didn't work/ were less successful in the classroom (PROBE: why)

Reflections on training

- Anything that should be changed about the training? (PROBE: what and why)

3. Reflections on coaching [5 minutes]

Aim: To explore experiences and usefulness of LRF coaching sessions. Understanding what worked well/ could be improved.

Coaching

- Feedback on QRTA's LRF coaching sessions
 - Format and delivery of coaching (PROBE: how many coaching sessions have they received, what worked well/ less well)
 - Any issues with the coaching sessions (PROBE: issues with coaches, practical arrangements, materials etc.)
 - Met expectations? (PROBE: anything not covered that should have been)
 - Anything that should be changed about the coaching sessions? (PROBE: what and why)

4. Reflections on LRF in practice [10 minutes]

Aim: To explore how LRF was implemented in the classroom and any issues encountered by teachers, with either the approach or the practice book. Also, to explore pupil engagement with LRF and the practice book.

Implementing the diagnostic tool

- **Implementation of the diagnostic tool**
 - How was the diagnostic tool implemented in the classroom? (PROBE: what worked well/less well)
 - Is there anything that would have been better at identifying students.

Implementing the LRF model in the classroom

- **Implementation in class**
 - How has the LRF model been implemented in the classroom? (PROBE: frequency of sessions – delivered 3 times a week, followed 'I do', 'we do', 'you do')
 - *For C/U only PROBE: where are classes delivered? Any issues with the practicalities of delivery?*

- Differences between LRF and usual literacy teaching (PROBE: positive and negative comparisons to usual practice)
- Anything about the model that doesn't work well in the classroom (PROBE: level of individualisation, issues with pupil engagement)
- **Using the practice book in class**
 - (How) has the practice book been used in class? (PROBE: are students able/ keen to use the practice book)
 - What about the practice book works well? (PROBE: any early reflections on impact for pupils)
 - Any issues with the practice book? (PROBE: anything missing, issues with layout/ format, issues with content etc.)
 - (How) could the practice book be improved for use in the classroom?
- **Pupil engagement**
 - How have students responded to the LRF approach? (PROBE: reaction of students in the classroom)
 - How engaged are students in the program? (PROBE: any examples teachers can give of engagement or lack of engagement, any differences between pupils).
 - What about the program has encouraged/discouraged student engagement? (PROBE: what elements of the program do students respond to/ engage with the most?)
 - Anything about LRF that could be changed to improve student engagement? (PROBE: why).

5. Engaging with parents [5 minutes]

Aim: To explore whether teachers were able to engage parents in LRF, and what engagement strategies were more/ less successful.

Contact with parents and home working

- **Making contact with parents**
 - What has been done to try to engage parents of LRF pupils?
 - Success and failures in terms of contacting/ engaging parents (PROBE: challenges connecting with parents – any effect positive or negative of LRF on engagement with parents)
 - How do teachers feel about use of WhatsApp to contact parents? (PROBE: levels of success)
- **Use of practice book at home**
 - (How) has the practice book been used at home (PROBE: how has that gone, have parents been supportive?)
 - (How) have teachers supported home working (PROBE: have there been checks on work at home, how has that been done, how have pupils/parents responded?)
 - What challenges have there been with practice book transfer between home and school?

6. Perceived impacts [5 minutes]

Aim: To explore perceived benefits of program for students and schools

- Perceived benefit of the program for pupils
Any effects of LRF on:
 - Pupils' reading ability
 - Confidence with reading
 - Engagement with literacy learning
 - Any difference between pupils (PROBE: subset of learners for whom LRF was more/ less effective)?
- Perceived benefits of the program for schools
- Any unanticipated effects

7. Feedback on LRF program [5 minutes]

Aim: To explore feedback of program for students and schools

- Anything about the program that will be taken forward after the implementation period/ changed approach to teaching?
- Anything that should be done differently about the classroom delivery?
- Anything that should be changed in terms of the practice book?
- Any changes to parental engagement?
- Any recommendations to improve the approach overall?

Focus Group Discussion Guide: Parents (control group)

Let's Read Fluently! (LRF) is designed to address issues with literacy among Arabic students. The intervention centers around a student practice book designed to reflect evidence about how Arabic reading fluency is best acquired. Teachers receive one-day training from the Queen Rania Teacher Academy (QRTA) and up to three follow-up coaching sessions to support them to use the practice book.

Teaching and Learning approach (W/C) and Literacy Catch-Up (C/U). Both run for one semester. W/C targets students in Grade 1 and is delivered to classes via three weekly 30-minute classes. The teacher adopts a 'I do', 'we do', 'you do' pedagogical approach using the practice book. C/U targets the 20% of lowest-achieving students in Grades 1 to 3 and is delivered through small group tuition (three weekly sessions).

The Queen Rania Foundation has commissioned a consortium which includes NatCen and Integrated to carry out a pilot evaluation of LRF. The aim of the evaluation is to explore delivery, understand the feasibility of different evaluation methods and inform learning for the planning of an effectiveness trial.

INTRODUCTION

Aim: to explain the aims of the research, how the interview will be conducted and how the data will be used.

- Introduction to researcher [NAME]. Thank you for agreeing to take part.
- Introduction to Integrated - we have been commissioned by the QRF (with NatCen) to conduct an independent evaluation of the pilot program 'Let's Read Fluently'. This includes speaking to those who are not delivering 'Let's Read Fluently'.
- Explanation of research – in addition to speaking to them and teachers at other participating schools, we are speaking to pupils, teacher and other stakeholders.
- Explanation of FGD, we will discuss:
 - Their typical approach for supporting their children to learn to read
 - Their perceptions of enablers and barriers children face when learning to read
 - Their insights into how teaching literacy could be improved
- Participation is voluntary – they can choose not to discuss any issue and can withdraw from the interview at any point.
- The information they provide will be used to write a findings report at the end of this school year for QRF.
- We are conducting this research independently and all information shared will be treated confidentially.
 - At the end of the interview, let us know if they would be happy for us to include quotes or would like to have anything removed.

- We would like to audio-record the discussion, so we have an accurate record of what is said.
- The FGD will last up to 50 minutes
- Any questions?

Date of FGD:

School/s:

1. Introduction and involvement in schoolwork [10 minutes]

Aim: To ease participant into conversation, background of their children and involvement in their schoolwork.

Background

- How old are your children? What grades are they in?
- Do you like reading?
 - PROBE: Why/Why not?
- Do your children like reading?
 - PROBE: Why/Why not?
- How important do you think learning to read is for children?
 - PROBE: Why/Why not?

2. Reflections on supporting children to learn to read [30 minutes]

Aim: To explore experiences of supporting their children to learn to read

Typical approach

- Does the school set any specific homework on learning to read?
 - What is the homework?
 - How often is homework on reading set?
 - Does your child usually complete the homework?
 - IF NO, why not?
 - IF YES, how often need help completing the homework?
 - How do you get involved? (I.e. helping them when they're stuck, reviewing their homework.)
 - How often do you do this?
 - How do you find it? (PROBE: What do you find easy? What do you find difficult?)
 - IF NOT, why do you not get involved? (PROBE: Is it too difficult? Do you have time?)
- Do you have any direct communication with the teacher about the classwork, homework, or the progress of your child?

- IF YES,
 - How do you communicate?
 - What's app
 - Letters sent home with child
 - Phone call
 - How often do you communicate?
 - What is the communication about?
 - Is this communication helpful? Why/why not?
- IF NO,
 - Would you find direct communication useful?
- Do your children have access to books/reading materials at home?
 - What types of books/reading materials?
 - Where do you get these books/reading materials from? E.g. Buy them, from school, from the library
 - Do your children read these? / Do you read these to your children?
 - If yes: When, how often and who with? (e.g., parent, grandparent, older sibling?)
 - If no: Why do they not read these?
 - Does your child have a favourite reading book?
 - Do you or anyone else read to your children at bedtime?

(In addition to any homework) How do you typically support your children to learn to read?

- Any particular exercises?
- Any particular books/reading materials?
- How often?
- PROBE: Why this way?
- IF NO ADDITIONAL SUPPORT: Why not? E.g. too difficult, don't know what to do, they wouldn't be interested

Enablers to teaching to read

- What do you think are the biggest enablers when supporting your children to learn to read?
 - Time spent reading
 - At home or at school
 - Alone or with someone else – if someone else, who?
 - Access to books and reading materials
 - What books or materials
 - Access from where
 - Quality, level, and topic of books/reading materials (e.g., topics and stories your children find readable and interesting)
 - Motivators to read
 - Rewards

- Trip to library
- Anything else
- Why are these effective?

Barriers to teaching to read

- What do you think are the biggest barriers when supporting your children to learn to read?
 - Lack of motivation or confidence
 - Inability to focus
 - Availability of someone to sit and read with them
 - Time spent reading
 - Access to books and reading materials
 - What books or materials
 - Access from where
 - Anything else
 - Why are these barriers?
 - Are you able to overcome these?
 - If yes – How?
 - If no – why not?

3. Potential improvements [10 minutes]

Aim: To explore what changes could be made to better support children when learning to read.

- What do you think about literacy teaching in school?
 - The amount of literacy teaching
 - The exercises
- How do you think the school or teacher could better support your child in developing their reading skills?
- Is there anything else you think works well when supporting your children to learn to read?
 - Why does it work well?
- Is there anything else you would change about teaching children to read?
 - Why would you change this?
- Any final comments and questions?

Thank and close

Focus Group Discussion Guide: Teachers (control group)

Let's Read Fluently! (LRF) is designed to address issues with literacy among Arabic students. The intervention centers around a student practice book designed to reflect evidence about how Arabic reading fluency is best acquired. Teachers receive one-day training from the Queen Rania Teacher Academy (QRTA) and up to three follow-up coaching sessions to support them to use the practice book.

Teaching and Learning approach (W/C) and Literacy Catch-Up (C/U). Both run for one semester. W/C targets students in Grade 1 and is delivered to classes via three weekly 30-minute classes. The teacher adopts a 'I do', 'we do', 'you do' pedagogical approach using the practice book. C/U targets the 20% of lowest-achieving students in Grades 1 to 3 and is delivered through small group tuition (three weekly sessions).

The Queen Rania Foundation has commissioned a consortium which includes NatCen and Integrated to carry out a pilot evaluation of LRF. The aim of the evaluation is to explore delivery, understand the feasibility of different evaluation methods and inform learning for the planning of an effectiveness trial.

INTRODUCTION

Aim: to explain the aims of the research, how the interview will be conducted and how the data will be used.

- Introduction to researcher [NAME]. Thank you for agreeing to take part.
- Introduction to Integrated - we have been commissioned by the QRF (with NatCen) to conduct an independent evaluation of the pilot program 'Let's Read Fluently'. This includes speaking to those who are not delivering 'Let's Read Fluently'.
- Explanation of research – in addition to speaking to them and teachers at other participating schools, we are speaking to pupils, parents and other stakeholders.
- Explanation of FGD, we will discuss:
 - Their typical approach for teaching children to read
 - The enablers and barriers to teaching children to read
 - Their insights into how teaching literacy could be improved
- Participation is voluntary – they can choose not to discuss any issue and can withdraw from the interview at any point.
- The information they provide will be used to write a findings report at the end of this school year for QRF.
- We are conducting this research independently and all information shared will be treated confidentially.

- At the end of the interview, let us know if they would be happy for us to include quotes or would like to have anything removed.
- We would like to audio-record the discussion, so we have an accurate record of what is said.
- The FGD will last up to 45 minutes
- Any questions?

Date of FGD:

School/s:

1. Participant background [5 minutes]

Aim: To ease participant into conversation, explore role and responsibilities, their understanding of the program and examine the school context

Role/responsibilities

- Brief overview of role/responsibilities (PROBE: class teacher or resource room teacher)
 - Length of time as a teacher

School context

- Overview of school
 - Size
 - Number of pupils/ staff

2. Reflections on teaching children to read [15 minutes]

Aim: To explore experiences of teaching children to read, and to ascertain a better understanding of business-as-usual delivery

Typical approach

- How would you typically teach children in grades 1-3 to read?
 - Any particular exercises?
 - Any particular books/reading materials?
 - PROBE: Why do you teach this way?
- Do you think there is the standard approach across Jordan?
 - IF YES - What is the standard approach?
 - IF YES - Does your approach differ? (PROMPT IF YES: How?)

Supporting students with difficulty reading

- How would you support those students who are struggling with reading?
 - Any particular exercises?

- Any particular books/reading materials?
- Use of resource teacher/resource room?
 - How often and for how long for?
 - How many students at one time?
 - Any particular exercises?
 - Any particular books/reading materials?
- PROBE: Why do you do it this way?

3. Reflections on enablers and barriers to reading [20 minutes]

Aim: To better understand what enablers and barriers are common with teaching to read, including interactions with parents and homework.

Enablers to teaching to read

- What do you think are the biggest enablers when teaching children to read?
 - Time spent reading
 - At home or at school
 - Alone or with someone else – if someone else, who?
 - Access to books and reading materials
 - What books or materials
 - Access from where
 - Parental/at home support
 - Anything else
 - Why are these effective?

Barriers to teaching to read

- What do you think are the biggest barriers when teaching children to read?
 - Time spent reading
 - Access to books and reading materials
 - What books or materials
 - Access from where
 - Parental support
 - Anything else
 - Why are these barriers?
 - Are you able to overcome these?
 - If yes – How?
 - If no – why not?

Engagement with parents/homework

- Do you involve parents when teaching children to read?
 - If yes – how?
 - WhatsApp
 - If no – why not?
- Do you think involving parents makes a difference?
 - Why/why not?
- Do you set homework when teaching children to read?
 - If yes – what do you typically set as homework?
 - If yes – how often do you set homework?
 - If yes – Why do you set homework?
 - If no – Why do you not set homework?

4. Potential improvements [5 minutes]

Aim: To explore what changes could be made to better support children when learning to read.

- Is there anything else you think works well when teaching children to read?
 - Why does it work well?
- Is there anything else you would change about teaching children to read?
 - Why would you change this?
- Any final comments and questions?

Thank and close

Appendix D: Criteria for moving to trial phase

W/C

Dimension	Criteria	Data source	Key Conclusions
Evidence of promise 1. In what ways, and to what extent, does 'Let's Read Fluently!' (LRF!) affect teacher and student practice as compared to business-as-usual teaching and learning? What are the participants' views on the perceived impact of LRF!?	1.1 The evidence suggests that the intervention could improve oral reading fluency and specific sub-domains of literacy attainment as measured with the EGRA in either model.	EGRA Assessment	Pupils in W/C schools made more progress in literacy attainment, on average, compared to pupils in control schools. These findings hold for all sub-domains of the EGRA assessment. Given the small sample sizes the analysis is not able to detect relevant differences.
	1.2 Teachers, coaches, parents, and students perceive that there is a positive value in this intervention compared to/in addition to business-as-usual.	Teacher survey FGDs	-Most teachers (65%) who responded to the survey felt that pupils in W/C Grade 1 improved their literacy attainment due to LRF!. Majority (55%) also felt that pupils' reading comprehension improved. -Parents felt that the LRF! practice book increased their children's reading skills confidence and skills -Pupils felt that the LRF! material was useful and helped them read.
	1.3 $\geq 50\%$ of Pupils articulate their personal development (ability, confidence) as a result of being part of this intervention.	FGDs with pupils	Pupils thought LRF! was useful and it helped them improve their reading ability.
	1.4 $\geq 70\%$ of teachers agree that the book was a useful aid in the classroom.	Teacher survey	Majority (65%) of teachers who responded to the survey agreed that the practice book is useful aid in the classroom.
	1.5 $\geq 70\%$ of teachers agree that children are more confident in reading due to Let's Read Fluently.	Teacher survey	Majority of teachers (55%) who took part in the survey agreed that children are more confident in reading due to LRF!.
	1.6 Coaches agree that LRF! had a positive impact on pupils' learning.	FGDs with	Coaches thought that W/C pupils were better able to navigate the practice book and they felt that pupils in W/C made an improvement in their reading skills.

	1.7 Evidence suggests that unintended negative consequences as a result of the implementation of LRF! is zero or minimal.	EGRA assessment; FGD	<p>-The descriptive analysis based on the EGRA assessment suggested that there are no negative consequences on literacy as result of LRF!.</p> <p>-Most teachers, parents and coaches did not think that LRF! could lead to any unintended negative consequences.</p> <p>-However, some parents felt that LRF! was burden to their children and that it may have worsened pupil attainment in other subjects.</p> <p>-Some teachers also felt that LRF! doesn't align with the national Arabic curriculum, and that teaching LRF! in addition to the national curriculum could be too much, and therefore there is a need for more coordination between the two.</p>
Evidence of promise 2. To what extent do teachers develop sufficient skills and confidence through the training and coaching?	2.1 Teachers believe that they have the skills and confidence to successfully deliver the intervention approach following the training and coaching.	Teacher survey	Majority of teachers (65%) who responded to the survey felt that the training equipped them with the skills and knowledge to deliver LRF!.
	2.2 Coaches think that teachers have developed the required skills and confidence through the training and coaching to successfully deliver the intervention approach.	FGD	Coaches reported that W/C teachers were engaged with the training and they were well equipped to implement LRF!

Dimension	Criteria	Data source	Key Conclusions
Feasibility of intervention 3. Was LRF! delivered as intended in terms of <u>dosage</u> , nature and quality? What modifications were made, with what implications?	3.1 At least 2 out of the 3 sessions per week are delivered as intended (i.e. sessions are 30 min long, with tasks being completed in the recommended length),	Teacher survey	All teachers who responded to the survey delivered 3 sessions per week.
	3.2 At minimum, 50% of the lessons were covered from the practice book; and teachers adopt all pedagogical approaches in the session (“I do”, “we do” “you do”).	Observations	Based on the data from the observation, the majority of the lessons used the practice and they delivered all pedagogical approaches.
	3.3 \geq %70 of teachers attending training sessions.	Teacher survey	Majority of teachers who responded to the survey (90%) attended the training.
	3.4 \geq 70% of teachers agree that the training equipped them with the skills and knowledge required to deliver Let’s Read Fluently.	Teacher survey	Majority of teachers who responded to the survey (84%) agreed that the training equipped them with the skills and knowledge required to deliver LRF!.
	3.5 \geq %70 of pupils feel happy to undertake reading or working from the practice book at schools.	Pupil survey	Most pupils (95%) felt happy to undertake reading or working from the practice books at schools.
Feasibility of intervention 4. What is the learning about teacher’s use of the Coarse-Grained Diagnostic (RAMP) tool? How successful is it, in use, at identifying the most appropriate pupils for the Literacy Catch-Up implementation model (C/U model only)?	4.1 \geq 50% of teachers (and coaches) in the catch up model agree that the diagnostic tool was successful in identifying the most appropriate pupils for the Literacy Catch-Up implementation model (maybe with few changes).	FGD	Not applicable
	4.2. The evidence suggests that the results of the diagnostic tool were consistent with the baseline (pre) EGRA assessment tool i.e. children selected for the C/U model have lower than average scores on EGRA in comparison to children in the W/C model.	EGRA assessment	Not applicable
Feasibility of the efficacy trial 5. What does the Pilot tell us about the feasibility of the	5.1 Enough in place to allow the intervention to take place the following year at scale (i.e., have are high, enough participants trained to	Delivery team assessment	The delivery team thinks that modifications to programme materials for W/C Grade 1 as well as successful execution of key programme procedures (e.g. recruitment and training), can be

process components of an Efficacy Trial (e.g., school recruitment, retention, or data collection in both intervention and control groups)?	act as trainers/coaches, school/participant retention rates during intervention the intervention materials and training suitably defined and developed)		finalised in time i.e. before delivery commences in January 2024.
	5.2 ≥ 85% of pupils complete the outcome testing in both intervention and control groups.	EGRA assessment	89% of pupils assigned to the W/C arm or the respective control group completed the EGRA assessment at both baseline and endline
Feasibility of the efficacy trial 6. What does the Pilot tell us about the feasibility of the resources of an Efficacy Trial (e.g. measurement instruments or specific equipment used)?	6.1 The training materials, practice books and measurement instruments are appropriate and meaningful. Any modifications to such tools are identified, based on evidence from the pilot, and possible to implement.	Teacher survey FGD	-Most teachers who responded to the teacher survey (65%) considered the practice book a useful aid in the classroom. -However, some teachers who took part in the FGDs felt that there was too much content to cover and not enough time. Teachers also felt that the 'you do' section was difficult to implement. -Coaches and teachers felt that the material was not appropriate for grade 1. -Coaches who took part in the FGDs felt there were too many words and teachers felt that non-sense words were unhelpful. -Coaches felt that the book does not take into account the different levels of ability.
	6.2 Sufficient numbers of training materials and practice books, even after modifications, can be available by the time required.	Delivery team assessment	Sufficient number of materials could be made available in time for delivery in January 2024.
Feasibility of the efficacy trial 7. What does the Pilot tell us about the feasibility of the management components of an Efficacy Trial (e.g. problems with data collection or variability of collected data)?	7.1. Implementing and evaluation partners have the human resource capacity, time available, funding, and positive working relationship status to successfully implement a large-scale RCT.	Partners own assessment	-The evaluation partners have the knowledge and human resource capacity to design and evaluate efficacy trial of LRF! -Successful implementation of LRF! relies on number of delivery partners and evaluators as demonstrated above. As part of the pilot evaluation we documented the decisions made, the actors involved, setbacks and successes. We think that we have established good working relationship and we can take some important lessons for the future.

Dimension	Criteria	Data source	Key Conclusions
<p>Evidence of promise</p> <p>1. In what ways, and to what extent, does 'Let's Read Fluently!' (LRF!) affect teacher and pupil practice as compared to business-as-usual teaching and learning? What are the participants' views on the perceived impact of LRF!?</p>	<p>1.1 The evidence suggests that the intervention could improve oral reading fluency and specific sub-domains of literacy attainment as measured with the EGRA in either model.</p>	<p>EGRA Assessment</p>	<p>There is no evidence that C/U model impacted Grade 1 pupils' literacy attainment, based on pupil's responses on the EGRA assessment. These findings hold for all sub-domains of the EGRA assessment. Given the small sample sizes the analysis is not able to detect relevant differences.</p>
	<p>1.2 Teachers, coaches, parents, and pupils perceive that there is a positive value in this intervention compared to/in addition to business-as-usual.</p>	<p>Teacher survey</p> <p>FGDs</p>	<p>-Most teachers (62.5%) who responded to the survey disagreed that pupils in C/U Grade 1 improved their literacy attainment due to LRF!. Half of teachers (50%) disagreed that pupils' reading comprehension improved.</p> <p>-There were mixed responses from teachers in the FGD, with some seeing no improvements and some noticing some improvement in reading speed, correct pronunciation and writing abilities.</p> <p>-Coaches did not think that C/U Grade 1 pupils gained much in terms of literacy in comparison to Grade 2 and 3 pupils.</p> <p>-Parents and teachers felt that the LRF! practice book was too difficult and advanced for C/U Grade 1 pupils.</p> <p>-Some parents felt that LRF! had positive impact on C/U Grade 1 performance, particularly on formulating words</p> <p>-Pupils in C/U Grade 1 enjoyed the choral reading as part of the 'We do' section, but found certain parts too difficult to grasp.</p>

	1.3 ≥ 50% of Pupils articulate their personal development (ability, confidence) as a result of being part of this intervention.	FGDs with pupils	-C/U pupils in Grade 1 thought LRF! was useful and it made them enjoy reading with peers. -However, pupils also thought that it was challenging to stay seated, and they thought that some part of the practice book were too difficult.
	1.4 ≥ 70% of teachers agree that the book was a useful aid in the classroom.	Teacher survey	Majority (62.5%) of teachers who responded to the survey agreed that the practice book is useful aid in the classroom.
	1.5 ≥ 70% of teachers agree that children are more confident in reading due to Let's Read Fluently.	Teacher survey	Majority of teachers (62.5%) who took part in the survey disagreed that pupils in C/U Grade 1 are more confident in reading due to LRF!.
	1.6 Coaches agree that LRF! had a positive impact on pupils' learning.	FGDs with	Coaches did not think that C/U Grade 1 pupils gained much in terms of literacy in comparison to Grade 2 and 3 pupils.
	1.7 Evidence suggests that unintended negative consequences as a result of the implementation of LRF! is zero or minimal.	EGRA assessment; FGD	-The descriptive analysis based on the EGRA assessment suggested that there are no negative consequences on literacy as result of LRF!. -Coaches felt that some pupils in C/U Grade 1 felt pressured and uncomfortable being in small group. They felt that the psychological impact was greater than the academic achievement for C/U Grade 1 pupils. -Coaches also thought that pupils in C/U Grade 1 were missing regular Maths, English and Religion classes due to LRF!. -Teachers felt that LRF! doesn't align with the national Arabic curriculum, and that teaching LRF! in addition to the national curriculum could be too much, and therefore there is a need for more coordination between the two.
Evidence of promise	2.1 Teachers believe that they have the skills and confidence to successfully deliver the intervention approach following the training and coaching.	Teacher survey	All resource room teachers (100%) who responded to the survey felt that the training equipped them with the skills and knowledge to deliver LRF!.

<p>2. To what extent do teachers develop sufficient skills and confidence through the training and coaching?</p>	<p>2.2 Coaches think that teachers have developed the required skills and confidence through the training and coaching to successfully deliver the intervention approach.</p>	<p>FGD</p>	<p>-Coaches expressed concerns about the 'buy-in' and enthusiasm for LRF! from resource room teachers, who had high workloads.</p> <p>-Coaches also reported that not all resource room teachers were of high calibre, and that they require more support</p>
--	---	------------	---

Dimension	Criteria	Data source	Key Conclusions
Feasibility of intervention 3. Was LRF! delivered as intended in terms of <u>dosage</u> , nature and quality? What modifications were made, with what implications?	3.1 At least 2 out of the 3 sessions per week are delivered as intended (i.e. sessions are 30 min long, with tasks being completed in the recommended length),	Teacher survey	All teachers who responded to the survey delivered 3 sessions per week.
	3.2 At minimum, 50% of the lessons were covered from the practice book; and teachers adopt all pedagogical approaches in the session (“I do”, “we do” “you do”).	Observations	Based on the observation almost all lessons were based on the practice book and they delivered all pedagogical approaches.
	3.3 \geq %70 of teachers attending training sessions.	Teacher survey	All resource room teachers who responded to the survey attended the training.
	3.4 \geq 70% of teachers agree that the training equipped them with the skills and knowledge required to deliver Let’s Read Fluently.	Teacher survey	All resource room teachers who responded to the survey agreed that the training equipped them with the skills and knowledge required to deliver LRF!.
	3.5 \geq %70 of pupils feel happy to undertake reading or working from the practice book at schools.	Pupil survey	Majority of C/U Grade 1 pupils who had copy of the practice book (69%) felt happy to undertake reading or working from the same at school. However, some of the pupils (19%) felt negative about the practice book.
Feasibility of intervention 4. What is the learning about teacher’s use of the Coarse-Grained Diagnostic (RAMP) tool? How successful is it, in use, at identifying the most appropriate pupils for the Literacy Catch-Up implementation model (C/U model only)?	4.1 \geq 50% of teachers (and coaches) in the catch up model agree that the diagnostic tool was successful in identifying the most appropriate pupils for the Literacy Catch-Up implementation model (maybe with few changes).	FGD	<ul style="list-style-type: none"> -Some teachers felt that the diagnostic tool was unnecessary as they could assess pupils reading ability without the tool -Teachers in larger classes felt that the tool was not helpful or accurate as it was easier for pupils to copy answers from their peers -Coaches felt that the diagnostic tool was not able to capture well ability among first graders
	4.2. The evidence suggests that the results of the diagnostic tool were consistent with the baseline (pre) EGRA assessment tool i.e. children selected for the C/U model have lower than average scores on EGRA in comparison to children in the W/C model.	EGRA assessment	The results from the baseline EGRA assessment suggest that pupils who were selected for the C/U model based on the diagnostic tool had significantly lower EGRA scores in comparison to children in W/C. This suggests that the diagnostic tool was successful in selecting children with lower literacy skills.

Feasibility of the efficacy trial 5. What does the Pilot tell us about the feasibility of the process components of an Efficacy Trial (e.g., school recruitment, retention, or data collection in both intervention and control groups)?	5.1 Enough in place to allow the intervention to take place the following year at scale (i.e., have enough participants trained to act as trainers/coaches, school/participant retention rates during intervention are high, the intervention materials and training suitably defined and developed)	Delivery team assessment	Modifications to programme materials for C/U Grade 1 as well as successful execution of key programme procedures (e.g. recruitment and training), can be finalised in time i.e. before delivery commences in January 2024.
	5.2 ≥ 85% of pupils complete the outcome testing in both intervention and control groups.	EGRA assessment	-86% of pupils assigned to the C/U Grade 1 or the respective control group completed the EGRA assessment at both baseline and endline.
Feasibility of the efficacy trial 6. What does the Pilot tell us about the feasibility of the resources of an Efficacy Trial (e.g. measurement instruments or specific equipment used)?	6.1 The training materials, practice books and measurement instruments are appropriate and meaningful. Any modifications to such tools are identified, based on evidence from the pilot, and possible to implement.	Teacher survey FGD	-Most teachers who responded to the teacher survey (65%) considered the practice book a useful aid in the classroom. -However, some teachers who took part in the FGDs felt that there was too much content to cover and not enough time. Teachers also felt that the ‘you do’ section was difficult to implement. -Coaches and teachers felt that the material was not appropriate for grade 1. -Coaches who took part in the FGDs felt there were too many words and teachers felt that non-sense words were unhelpful. -Coaches felt that the book does not take into account the different levels of ability.
	6.2 Sufficient numbers of training materials and practice books, even after modifications, can be available by the time required.	Delivery team assessment	Sufficient number of materials could be made available in time for delivery in September 2023.
Feasibility of the efficacy trial 7. What does the Pilot tell us about the feasibility of the management components of an Efficacy Trial (e.g. problems with data collection or variability of collected data)?	7.1 Implementing and evaluation partners have the human resource capacity, time available, funding, and positive working relationship status to successfully implement a large-scale RCT.	Implementors and evaluators own assessment.	-The evaluation partners have the knowledge and human resource capacity to design and evaluate efficacy trial of LRF! -Successful implementation of LRF! relies on number of delivery partners and evaluators as demonstrated above. As part of the pilot evaluation we documented the decisions made, the actors involved, setbacks and successes. We think that we have established good working relationship and we can take some important lessons for the future.

Dimension	Criteria	Data source	Key Conclusions
<p>Evidence of promise</p> <p>1. In what ways, and to what extent, does 'Let's Read Fluently!' (LRF!) affect teacher and pupil practice as compared to business-as-usual teaching and learning? What are the participants' views on the perceived impact of LRF!?</p>	<p>1.1 The evidence suggests that the intervention could improve oral reading fluency and specific sub-domains of literacy attainment as measured with the EGRA in either model.</p>	<p>EGRA Assessment</p>	<p>There is no evidence that C/U model impacted Grade 2 and 3 pupils' literacy attainment, based on pupil's responses on the EGRA assessment. These findings hold for all sub-domains of the EGRA assessment. Given the small sample sizes the analysis is not able to detect relevant differences.</p>
	<p>1.2 Teachers, coaches, parents, and pupils perceive that there is a positive value in this intervention compared to/in addition to business-as-usual.</p>	<p>Teacher survey FGDs</p>	<p>-Most teachers (87.5%) who responded to the survey agreed that pupils in C/U Grade 2 and 3 improved their literacy due to LRF!. Most teachers (75%) agreed that pupils' reading comprehension improved.</p> <p>-Teachers in the FGDs confirmed that they observed improvements amongst pupils in the C/U classes, particularly amongst 3rd grade pupils.</p> <p>-Coaches thought that C/U Grade 2 and 3 pupils gained more in terms of development in literacy in comparison to Grade 1.</p> <p>-Parents of pupils in C/U Grade 2 and 3 felt that their children's confidence in reading has improved, and some of them observed improvements in children's reading ability.</p> <p>-Pupils in C/U Grade 1 enjoyed the choral reading as part of the 'We do' section, but found certain parts too difficult to grasp.</p>
	<p>1.3 \geq 50% of Pupils articulate their personal development (ability, confidence) as a result of being part of this intervention.</p>	<p>FGDs with pupils</p>	<p>-C/U pupils in Grade 2 and 3 thought LRF! was useful and it made them enjoy reading with peers.</p> <p>-Pupils in Grade 2 and 3 also expressed that they became in reading because they had received gifts.</p>

			-However, pupils also thought that the 'I do' section was challenging.
	1.4 ≥ 70% of teachers agree that the book was a useful aid in the classroom.	Teacher survey	All teachers who responded to the survey agreed that the practice book is useful aid in the classroom.
	1.5 ≥ 70% of teachers agree that children are more confident in reading due to Let's Read Fluently.	Teacher survey	Majority of teachers (62.5%) who took part in the survey agreed that pupils in C/U Grade 2 and 3 are more confident in reading due to LRF!.
	1.6 Coaches agree that LRF! had a positive impact on pupils' learning.	FGDs with	Coaches thought that LRF! gave pupils in Grade 2 and 3 a routine and responsibility for individual work. Coaches also thought that LRF! had positive impact on pupils' confidence in reading.
	1.7 Evidence suggests that unintended negative consequences as a result of the implementation of LRF! is zero or minimal.	EGRA assessment; FGD	-The descriptive analysis based on the EGRA assessment suggested that there are no negative consequences on literacy as result of LRF!. -Coaches did not discuss any negative consequences for C/U pupils in Grade 2 and 3. -Teachers felt that LRF! doesn't align with the national Arabic curriculum, and that teaching LRF! in addition to the national curriculum could be too much, and therefore there is a need for more coordination between the two. -Teachers also felt that pupils got frustrated during the more difficult 'I do' section.
Evidence of promise 2. To what extent do teachers develop sufficient skills and confidence through the training and coaching?	2.1 Teachers believe that they have the skills and confidence to successfully deliver the intervention approach following the training and coaching.	Teacher survey	Most resource room teachers (87.5%) who responded to the survey felt that the training equipped them with the skills and knowledge to deliver LRF!.
	2.2 Coaches think that teachers have developed the required skills and confidence through the training and coaching to successfully deliver the intervention approach.	FGD	-Coaches reported high absenteeism among resource room teachers for Grade 2 and 3

			<p>-Coaches also expressed concerns about the 'buy-in' and enthusiasm for LRF! from resource room teachers, who had high workloads.</p> <p>-Coaches also felt that some teachers did not fully understand the training, and as a result some of them adapted implementation and deviated from the programme.</p> <p>-Coaches also reported that resource room teachers may require more support.</p>
--	--	--	--

Dimension	Criteria	Data source	Key Conclusions
Feasibility of intervention 3. Was LRF! delivered as intended in terms of <u>dosage</u> , nature and quality? What modifications were made, with what implications?	3.1 At least 2 out of the 3 sessions per week are delivered as intended (i.e. sessions are 30 min long, with tasks being completed in the recommended length),	Teacher survey	No data available
	3.2 At minimum, 50% of the lessons were covered from the practice book; and teachers adopt all pedagogical approaches in the session ("I do", "we do" "you do").	Observations	Based on the observation almost all of the lessons were based on the practice book and all pedagogical approaches were delivered during the session.

	3.3 ≥ %70 of teachers attending training sessions.	Teacher survey	Almost all resource room teachers (87.5%) who responded to the survey attended the training.
	3.4 ≥ 70% of teachers agree that the training equipped them with the skills and knowledge required to deliver Let's Read Fluently.	Teacher survey	Almost all resource room teachers (87.5%) who responded to the survey agreed that the training equipped them with the skills and knowledge required to deliver LRF!.
	3.5 ≥ %70 of pupils feel happy to undertake reading or working from the practice book at schools.	Pupil survey	Majority of C/U Grade 2 and 3 pupils who had copy of the practice book (89%) felt happy to undertake reading or working from the same at school. Only one pupil in Grade 2 and 3 felt negative about the practice book.
Feasibility of intervention 4. What is the learning about teacher's use of the Coarse-Grained Diagnostic (RAMP) tool? How successful is it, in use, at identifying the most appropriate pupils for the Literacy Catch-Up implementation model (C/U model only)?	4.1 ≥ 50% of teachers (and coaches) in the catch up model agree that the diagnostic tool was successful in identifying the most appropriate pupils for the Literacy Catch-Up implementation model (maybe with few changes).	FGD	-Some teachers felt that the diagnostic tool was unnecessary as they could assess pupils reading ability without the tool -Teachers in larger classes felt that the tool was not helpful or accurate as it was easier for pupils to copy answers from their peers -Coaches felt that the diagnostic tool was not able to account for learning difficulties, and these pupils needed to be opted out of the selection process
	4.2. The evidence suggests that the results of the diagnostic tool were consistent with the baseline (pre) EGRA assessment tool i.e. children selected for the C/U model have lower than average scores on EGRA in comparison to children in the W/C model.	EGRA assessment	Not applicable
Feasibility of the efficacy trial 5. What does the Pilot tell us about the feasibility of the process components of an Efficacy Trial (e.g., school recruitment, retention, or data collection in both intervention and control groups)?	5.1 Enough in place to allow the intervention to take place the following year at scale (i.e., have enough participants trained to act as trainers/coaches, school/participant retention rates during intervention are high, the intervention materials and training suitably defined and developed)	Delivery team assessment	Modifications to programme materials for C/U as well as successful execution of key programme procedures (e.g. recruitment and training), can be finalised in time i.e. before delivery commences in January 2024.
	5.2 ≥ 85% of pupils complete the outcome testing in both intervention and control groups.	EGRA assessment	-91% of pupils in the C/U Grade 2 and 3 control group completed the EGRA testing at baseline and endline.

			-69% of pupils in treatment group completed the EGRA testing at both baseline and endline
Feasibility of the efficacy trial 6. What does the Pilot tell us about the feasibility of the resources of an Efficacy Trial (e.g. measurement instruments or specific equipment used)?	6.1 The training materials, practice books and measurement instruments are appropriate and meaningful. Any modifications to such tools are identified, based on evidence from the pilot, and possible to implement.	Teacher survey FGD	-All teachers who responded to the teacher survey considered the practice book a useful aid in the classroom. -Teachers confirmed that they would continue to use the practice book in the future. -Coaches and teachers felt that the material was more appropriate for Grade 2 and 3 pupils. -However, coaches also made number of suggestions for improvements e.g. there are too many words, text should be bigger, the lines used to separate the sentences are confusing, vowels are introduced too later. -Coaches felt that the book does not take into account the different levels of ability.
	6.2 Sufficient numbers of training materials and practice books, even after modifications, can be available by the time required.	Delivery team assessment	Sufficient number of materials could be made available in time for delivery in September 2023.
Feasibility of the efficacy trial 7. What does the Pilot tell us about the feasibility of the management components of an Efficacy Trial (e.g. problems with data collection or variability of collected data)?	7.1 Implementing and evaluation partners have the human resource capacity, time available, funding, and positive working relationship status to successfully implement a large-scale RCT.	Implementors and evaluators own assessment.	-The evaluation partners have the knowledge and human resource capacity to design and evaluate efficacy trial of LRF! -Successful implementation of LRF! relies on number of delivery partners and evaluators as demonstrated above. As part of the pilot evaluation we documented the decisions made, the actors involved, setbacks and successes. We think that we have established good working relationship and we can take some important lessons for the future.