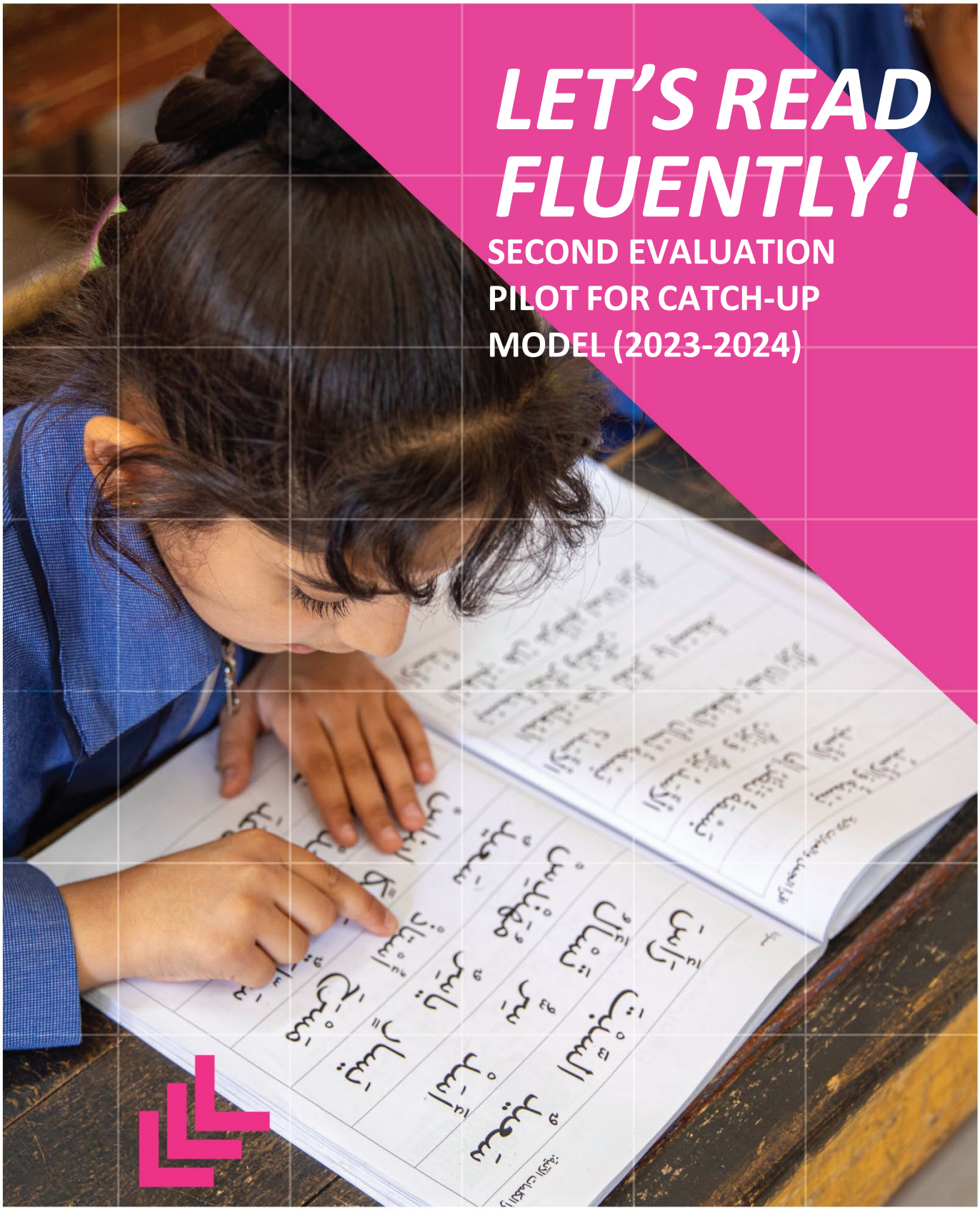
مؤسسة الملكة رانيا
QUEEN RANIA FOUNDATION

Education
Endowment
Foundation

BHP | Foundation



LET'S READ FLUENTLY!

SECOND EVALUATION PILOT FOR CATCH-UP MODEL (2023-2024)

TABLE OF CONTENTS

About the evaluator

The evaluation team includes the following staff from NatCen, as well as Integrated and the Queen Rania Foundation:

Enes Duysak, Hayley Leonard, Abbi Rennick, Charlotte Bessant, Anjhana Damodaran, Rebecca Parker, Gayle Munro, Andi Fugard, Nedjma Koval, Samah Goussous, Marwa Alsamneh, Taimaa Khalaf, Rasha Al-Khateeb.

The lead evaluator was Enes Duysak.

Acknowledgements

We would like to thank all the teachers, parents, and students 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 (Rawan Awwad, Rami Al Assad, Haneen Al Abed) 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 Ignacia Valenzuela 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 co-funded by the Queen Rania Foundation in addition to the Education Endowment Foundation (the latter in partnership with the BHP Foundation, as part of the "Building a global evidence ecosystem for teaching" project).

Acronyms

Modern Standard Arabic	MSA
Let's Read Fluently	LRF!
Whole Class model	W/C model
Catch-up model	C/U model
United Arab Emirates	UAE
Impact Evaluation	IE
Implementation and Process Evaluation	IPE
Queen Rania Foundation	QRF
Education Endowment Foundation	EEF
Queen Rania Teaching Academy	QRTA
Professional Learning Community	PLC
Princess Taghrid Institute for Development and Training	PTI
Early Grade Reading Assessment	EGRA
Program for International Student Assessment	PISA
Organization for Economic Cooperation and Development	OECD
Progress in International Reading Literacy Study	PIRLS
Special Educational Needs and Disabilities	SEND
Early Grade Reading and Mathematics Project	RAMP
Oral Reading Fluency	ORF
Ministry of Education	MoE
File Transfer Protocol	FTP
Research Ethics Committee	REC
UK Economic and Social Research Council	ESRC
UK Government Social Research	GSR
Randomised Controlled Trial	RCT
Arabic Literacy Attainment	ALA
Letter Sound Identification	LSI
Syllable Identification	SI
Word Recoding	WR
Reading Comprehension	RC
Listening Comprehension	LC
Intention-To-Treat	ITT
Confidence Intervals	CI
Minimum Relevant Mean Difference	MRMD
Minimum Relevant Effect Size	MRES
Intracluster Correlation Coefficients	ICCs
Area Under the Receiver Operating Characteristic	AUROC curve
Feasibility of Intervention	FI
Feasibility of Trial	FT
Research Question	RQ

EXECUTIVE SUMMARY

Evaluation Team

The evaluation team includes the following staff from NatCen and Integrated International: Enes Duysak, Hayley Leonard, Abbi Rennick, Charlotte Bessant, Rebecca Parker, Anjhana Damodaran, Andi Fugard, Gayle Munro, Nedjma Koval, Samah Goussous, Marwa Alsamneh, Rasha Al-Khateeb, and Taimaa Khalaf.

The principal investigator was Enes Duysak.

Acknowledgements

We sincerely appreciate the dedication and efforts of all participating **teachers, parents, and students**, whose commitment to improving literacy in Jordanian schools has been invaluable to this study.

We extend thanks to the Queen Rania Foundation team (especially Rawan Awwad, Rami Al Assad, and Haneen Al Abed) for their support. Additionally, we acknowledge the Education Endowment Foundation team led by Ignacia Valenzuela and contributions from Dr. Helen Abadzi and Julie Helson.

Special thanks to the Queen Rania Teacher Academy (QRTA) team (Lubna Al Drainy, Mohammad Salameh, and Sana Al Syouf) for candidly implementing the program and their expertise in teacher training and professional development. We also appreciate the Princess Taghrid Institute (PTI) team for their role in the diagnostic assessments and student selection process.

This research was co-funded by the Queen Rania Foundation and the Education Endowment Foundation, the latter in partnership with the BHP Foundation, as part of the “Building a Global Evidence Ecosystem for Teaching” project.

Overview

The Let's Read Fluently! (LRF!) Catch-Up (C/U) Model is a targeted intervention aimed at supporting students who struggle with reading in Modern Standard Arabic (MSA). This pilot evaluation, commissioned by the Queen Rania Foundation (QRF) and supported by the Education Endowment Foundation (EEF) and BHP Foundation, was conducted by NatCen in collaboration with Integrated International. The study focused on assessing the feasibility of a refined delivery model and its readiness for an efficacy trial.

LRF! is grounded in Dr. Helen Abadzi's cognitive science framework, which emphasizes the development of automaticity in letter recognition and decoding as a precursor to reading fluency. The model uses structured, repetitive practice to support perceptual learning and improve working memory efficiency, addressing key challenges of Arabic script complexity.

Previous Pilot Evaluation

Previous pilot evaluation of Let's Read Fluently! have tested both Catch-Up (C/U) and Whole-Classroom (W/C) models:

- The W/C model was previously piloted and demonstrated promising literacy gains for a broader student population.
- The first C/U pilot revealed challenges in adapting the intervention to resource room settings, prompting modifications in training, coaching, and student selection.

The current pilot aimed to refine the C/U model by enhancing teacher support mechanisms and refining the multi-staged screening process.

Intervention Implementation

The intervention was implemented over a 12-14 week period, with three 30-minute sessions per week delivered in small groups of five or six students by the resource room teacher. The sessions focused on structured reading exercises, decoding strategies, and guided fluency practice, aligned with the LRF framework to reinforce automaticity in reading.

Screening of students occurred through a multi-staged process:

1. Coarse-Grain Screening from the Reading and Mathematics Program (RAMP Tool): Classroom teachers identified the lowest-performing 20% of students.
2. PTI Diagnostic Assessment: Specialists conducted assessments to exclude students with broader learning difficulties, who would not benefit from the intervention.
3. Early Grade Reading Assessment (EGRA) Testing: Students scoring ≤29 correct words per minute for oral reading fluency (ORF) were identified as needing intervention support.

Evaluation Methods

The study employed a Randomised Control Trial (RCT) design, complemented by an Implementation and Process Evaluation (IPE):

Impact Evaluation (IE):

- Sixteen schools were randomly assigned to intervention or control arms. The study was implemented in schools located in Amman,Karak, Madaba, Balqa, Jarash, and Ajloun governorates.
- Primary Outcome: Oral reading fluency.
- Secondary Outcomes: Letter Sound Identification, Syllable Identification, Oral Reading Fluency, Reading Comprehension, Listening Comprehension, Word decoding.
- The instrument used to measure primary and secondary outcomes was the EGRA Grade 2 assessment with the addition of a set of pre-literacy items developed and piloted for the previous pilot evaluation, for both baseline and endline.
- The primary outcome analysis followed an Intention-To-Treat (ITT) approach, with 161 students included in the analyses, taken from a random selection of classrooms in grades 2 and 3 from the 16 schools.
- Evidence of Promise Criteria: Effect sizes and confidence intervals were assessed against Minimum Relevant Effect Size (MRES) thresholds.

Implementation and Process Evaluation (IPE):

Data collection included Focus Group Discussions (FGDs) and structured surveys with classroom and resource room teachers in intervention and teaching-as-usual schools, as well as a FGD with coaches from QRTA. The following IPE domains were explored:

- Fidelity: Measured adherence to the intervention.
- Reach: Measured rate and scope of participation in the intervention.
- Responsiveness: Evaluated the degree to which participants engaged with the intervention.
- Perceived impact: Evaluated whether teachers and coaches perceived the intervention had achieved its intended outcomes.
- Usual practice: Identified how the intervention differed from usual literacy teaching.

The IPE findings were used to inform conclusions on the feasibility of the intervention and to identify any barriers to large scale implementation, including logistical and resource challenges.

Key Findings from current pilot

1. Evidence of Promise:

- The pilot demonstrated indicative evidence of improvement in students' oral reading fluency, the primary outcome investigated.
- Among secondary outcomes: Syllable identification and word decoding showed positive results, while letter sound identification, reading comprehension and listening comprehension were not conclusive.
- Enhanced training contributed to improved instructional delivery.

2. Feasibility of Implementation:

- Resource room teachers benefited from structured training and coaching, yet challenges remain in scaling the model.
- The multi-staged screening process (RAMP, PTI assessment, EGRA) remains inconclusive, with continued concerns about accurately identifying students most in need.

3. Feasibility of the Efficacy Trial

- School Recruitment: Recruitment was largely effective, but over 60% of schools approached were unreachable due to incorrect contact details. Future efficacy trials must account for this by expanding the pool of schools and improving outreach strategies.
- Retention Rates: Both school and student retention were high, an improvement over the previous pilot evaluation, suggesting feasibility in maintaining engagement over the trial period.
- Resource Feasibility: The LRF! training materials, practice book, and intervention resources were generally well-received. However, suggested improvements—such as simplifying the practice book and including visuals—may enhance usability for teachers and students.
- Screening Tools: The feasibility of the coarse-grain screening and PTI tools remains unclear. A move toward standardised literacy and cognitive development assessments may improve accuracy in student selection.

4.. Readiness for Efficacy Trial:

- While the intervention showed promise, scalability is limited due to teacher workload and logistical constraints.

Conclusions and Recommendations

- There is indicative evidence for an effect of the LRF! C/U model on oral reading fluency, as well as syllable identification and word decoding.
- The student screening process remains uncertain, requiring further refinement.
- Teacher training and coaching mechanisms were effective, but scaling is a concern due to operational constraints.
- A full-scale efficacy trial is not yet recommended based on these two pilots for the C/U model. A future pilot evaluation should focus on refining student selection, improving implementation, and assessing scalability.

Given the challenges in scaling the C/U model, findings suggest that the W/C is a more viable candidate for an efficacy trial. The W/C model demonstrated feasibility in reaching a larger student population while maintaining instructional effectiveness. The second C/U pilot provides insights into shared instructional elements that could enhance the W/C model's delivery.

The LRF! intervention aligns with Jordan's national literacy priorities and has the potential to improve early-grade reading outcomes, but further refinements are necessary before scaling the C/U model program.

INTRODUCTION

Context: The importance of early literacy

A strong foundation in literacy is crucial for children's later academic development (Zakaria et al., 2021). Findings from Brombacher et al. (2012) suggest that early literacy difficulties can persist, limiting children's ability to go on to achieve their potential. In some languages, such as Arabic, there are particular challenges with formal literacy learning because it differs from the colloquial form of language used at home (Abadzi, 2017). Research conducted in Jordan in 2017 and 2018 reported that the proportion of Grade 2 and Grade 3 students reaching oral reading fluency benchmarks of 46 words per minute was low, at 13.2% in 2017 and 19.1% in 2018 (RTI International, 2018). This is worrying, particularly when interpreted in line with findings from The World Bank (2019), who report that it is unlikely that students will make up for learning loss during the next stages of their education. Early intervention to support Arabic literacy development is therefore of great importance.

Although students in Jordan are expected to know Modern Standard Arabic (MSA) by the time they enroll in school, the reality is that they are tasked with absorbing it alongside developing their more colloquial literacy at home. This diglossia[1] can result in students not being able to make sense of MSA's visual complexities and grammatical framework due to the differences in verb and noun endings, expressions, vocabulary, and pronunciation (Abazdi, 2017). The visual complexity of the Arabic script specifically limits its accessibility and retention for reading comprehension (Abadzi, 2017). The script consists of 29 letters and eight diacritics[2], with some letters having up to four different shapes depending on their position in a word. Adding to this, letters can look very different depending on the font in which they are written (Eckert et al., 2020). Considering the linguistic challenges that readers in Arabic face, it is important to identify approaches that will help students with literacy attainment (Huri, 2012). One of these approaches is Let's Read Fluently! (LRF!).

Background to this pilot

Let's Read Fluently! (LRF!) is an intervention that aims to support students to overcome obstacles in learning to read Arabic and successfully develop foundational literacy skills. There are two models of LRF! delivery: a Whole-Class Teaching and Learning approach (W/C), and a Catch-Up (C/U) model. The W/C approach supports all students in a class, while the C/U model provides targeted support for selected students who are identified as falling behind their classmates in MSA Arabic literacy.

Similar approaches to the LRF! model in other countries, such as Cambodia, the Gambia and Egypt, have shown evidence of promise (Abazdi, 2013). The similarities between interventions include the use of a textbook with a simple functional design, the gradual introduction of a new letter/concept, independent reading, and feedback from the teacher, all of which are used in the LRF! model. There are also early results from a small-scale pilot conducted in the United Arab Emirates (UAE), which demonstrated how students' reading ability increased following an LRF!-style classroom intervention, namely, being able to read more letters and making fewer errors than their peers (Eckert et al., 2020). This suggests the LRF! model as a whole may have a positive impact on Arabic reading fluency among early grade students.

[1] Diglossia is when there are multiple varieties of the same language throughout a community.

[2] Diacritics are marks placed above or below (or sometimes next to) a letter in a word to indicate how letters are pronounced, and to distinguish between words of similar spelling.

A previous pilot study of LRF! in Jordan was carried out in 2021-22 by the National Centre for Social Research (NatCen), in collaboration with Integrated International (Integrated), School-to-School International and Oxford MeasurEd. The pilot found evidence that the W/C approach could be effective but did not find evidence of promise for the C/U approach (Dimova et al., 2023). In particular, it was reported that the students participating in the programme were not always able to engage appropriately with the intervention (for example, the content was found to be too difficult for those in Grade 1), and that resource room teachers who implemented the C/U model outside of usual class time would have benefited from additional training and support.

The current report describes the evaluation of a recommissioned pilot of the LRF! C/U model. It aimed to understand whether a revised delivery approach, addressing the limitations of the previous pilot, could show evidence of promise. The new pilot assessed the feasibility of the LRF! C/U model, feasibility of an efficacy trial, and readiness for an efficacy trial. It includes an Impact Evaluation (IE) component and an Implementation and Process Evaluation (IPE) component.

This pilot was funded by the Queen Rania Foundation (QRF) and supported by the Education Endowment Foundation (EEF) and the BHP Foundation. The evaluation team was led by NatCen working together with Integrated, a research organisation based in Jordan.

Rationale for the revised pilot

Some of the relevant key findings for the LRF! C/U model from the previous pilot (Dimova, et al. 2023) were used to inform revisions to the programme and the logic model. These are listed in Table 1.

Table 1: Previous pilot findings and subsequent changes to programme

Key findings from previous pilot	Revision to the current pilot's programme and logic model
Resource room teachers reported difficulty in being able to adapt the pace of learning to students' needs while still completing the syllabus.	The maximum capacity of students in each resource room session was capped at six by Jordan's Ministry of Education. The training for resource room teachers was extended to two days instead of one. Classroom teachers would no longer be trained to deliver the LRF! C/U model. The training was for resource room teachers only. Beyond the three coaching sessions that were already part of the intervention, resource room teachers also received additional support through fortnightly online meetings. Queen Rania Teacher Academy (QRTA) established a Professional Learning Community (PLC), which resource room teachers could draw ongoing support from during the intervention via online meetings.
Some resource room teachers found delivering the LRF! C/U model challenging, due to working with students facing difficulties with their learning.	
Resource room teachers typically had less experience than classroom teachers.	

Some resource room teachers felt that the practice book had a 'one-size-fits-all' approach and suggested that having books for different abilities could be a solution.	The reading practice book was modified based on the feedback given by Dr. Helen Abadzi and teachers from the previous pilot phase. This feedback included that the page content was too dense, words were too similar, and the presence of typos and unclear text. Changes therefore included replacing some syllables, words and phrases, changing fonts, font size and spacing. A book for different abilities was not implemented for this pilot.
LRF! C/U was too difficult for students in Grade 1.	The C/U model was implemented with students in Grades 2 and 3 only. Grade 1 was no longer included.
The pace of content was too fast for struggling students, leaving resource room teachers unable to finish more than a page of the practice book during some of their sessions.	
The 'You do' stage of LRF! was challenging for C/U students, leaving some feeling frustrated.	
The coarse-grain screening tool alone did not identify the students most suited for the C/U model. The tool identified a very high proportion of students struggling with literacy. The evidence suggested that in some cases teachers responded to this by selecting students for the intervention based on their own judgement, which they felt was more accurate than the diagnostic tool. In the end, C/U students in the previous pilot were found to have varying needs, which led to some being removed from the intervention during delivery.	The approach to identifying eligible students for inclusion in the C/U model was updated. Rather than relying on scores from the coarse-grain tool alone, an additional diagnostic assessment tool developed by Princess Taghrid Institute for Development and Training (PTI), along with Early Grade Reading Assessment (EGRA) scores, were used to help screen out those for whom LRF! C/U support was not considered suitable.

This pilot evaluation focused on the aspects of the C/U model that had changed, rather than re-examining unchanged aspects of the programme where the pilot evidence had already been gathered in 2021-22 (Dimova et al., 2023).

Intervention description

Why

It is estimated that early readers in Arabic need a level of automaticity[3] in oral reading fluency[4] of 45-60 words per minute (RTI, 2012). This fluency allows working memory to be freed up for

[3] Automaticity is defined as being able to complete a task with no conscious effort.

[4] 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.

comprehension. Data from RTI (2018) using the Early Grade Reading Assessment (EGRA) suggests that only around 19% of Grade 2 and Grade 3 students meet, or exceed, the lowest levels of this benchmark. Alongside that, a significant number of students in Jordan (16.6%) scored zero in oral reading fluency.

Indicators of early literacy, such as oral reading fluency, are important for the later development of reading (EEF, 2023). The 2018 Program for International Student Assessment (PISA) found that for 15-year-old Jordanian students, attainment levels were behind the Organization for Economic Cooperation and Development (OECD) average by an equivalent of more than one grade in reading. Only one in five students performed at or around the average OECD reading score and two in every five performed below the minimum proficiency level in reading (QRF, 2020). The 2021 Progress in International Reading Literacy Study (PIRLS), which reports on reading levels in 57 countries, found that 10-year-old Jordanian students had significantly lower average reading achievement in comparison to 55 of the other countries (IEA, 2021).

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 (World Bank Group, 2023).

When diacritics are used, Arabic is a transparent language – that is, there is a reliable relationship between letters and sounds. These low EGRA scores therefore likely reflect a gap in phonics skills, or the understanding of the relationships between a written letter (or symbol) and its sound (EEF, 2021). The importance of phonics for reading attainment is reflected in the EEF's Teaching and Learning Toolkit and other literature (Seidenberg, 2017; Castles et al., 2018).

It is also important to note that learning losses from the consequences of COVID-19, including sustained school closures, could still be having an impact on children's educational attainment (Cortés-Albornoz et al., 2023). This context reinforces the need for interventions to support literacy acquisition and strong evidence to understand what works (UNICEF, 2023).

The LRF! C/U approach is designed for the lowest achieving 20% of students in terms of Arabic literacy. This approach is considered appropriate for those students who are less likely to progress in a larger classroom size and need tailored support to advance their learning. The practice book is designed to enable those who find it harder to learn new letters to be supported through the use of gradual introduction of new letters, as well as consistent repetition of symbol-sound relationships.

Who

Both resource room teachers and students can be considered recipients of the LRF! C/U model.

Resource room teachers received a two-day training course from the Queen Rania Teacher Academy (QRTA). They then delivered the LRF! C/U model to students in the resource room via small group tuition. Concurrently, they engaged in three coaching sessions and had access to fortnightly online meetings from QRTA, to provide ongoing support in carrying out the intervention.

The LRF! C/U model is targeted at students in Grades 2 and 3 who are among the lowest achieving 20% of students in terms of Arabic literacy. It has not been designed to provide support for students with Special Educational Needs and Disabilities (SEND).[5]

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

Students were identified for the LRF! C/U model support based on their performance in three assessments: the start-of-term coarse-grain screening tool, an additional diagnostic assessment administered by the PTI, and the EGRA assessment.

Selection of students for LRF! C/U model support

Students were identified for inclusion in the intervention through the following process:

- At the start of the school year in September 2023, classroom teachers carried out the **coarse-grain screening tool**[6] developed by the Early Grade Reading and Mathematics Project (RAMP)[7] with all students in their class. This is a part of the usual practice in Jordanian schools at the start of the school year.
- Where there were more than three classes per grade, three classes were randomly selected for the evaluation.
- Students who scored in the lowest 20% of their class in Grades 2 and 3 in the coarse-grain tool undertook an additional **diagnostic assessment** carried out by PTI[8]. The purpose of this additional screening was to identify students with characteristics that may be associated with SEND, for whom the LRF! C/U model may not be the most appropriate form of support.
- The **Oral Reading Fluency (ORF) sub-scale of the EGRA**[9] was used as the final screening criterion to identify students with reading proficiencies that made them eligible for the programme. Students who scored 29 correct words per minute or less for the ORF sub-scale of EGRA were identified as eligible for the programme. This is in line with the four categories of reading proficiency for Jordanian students in Grades 2 and 3 published by RTI International (2023); those scoring 29 correct words per minute or less are categorised into the lowest reading proficiency categories of 'non-readers' (ORF = 0) and 'beginning readers' (1 <= ORF <= 29).

Following this screening process, students in Grades 2 and 3 who scored in the lowest 20% of their class by the coarse-grain tool, and were additionally found to be eligible after the PTI diagnostic assessment and EGRA assessment, were selected for inclusion in the LRF! C/U sessions. The student screening process is described in greater detail in the *Recruitment* section below.

[6] You can find the copy of the coarse-grain reading assessment from this webpage: Jordan Remedial Study: Reading Diagnostic Assessment Tool and Stimulus Sheet | SharEd (rti.org)

[7] The Early Grade Reading and Mathematics Project (RAMP) aims to improve learning outcomes for reading in Arabic and math in grades K2-G3 for all public schools in Jordan. This involves improving early grade reading and math learning materials, better preparing teachers and administrators to provide effective reading and math instruction through in-service induction and pre-service training, mentoring and supervision, engaging communities for participation in the education of all children and holding schools accountable for results and supporting the Government of Jordan's efforts to institutionalise early grade reading and math policies, standards and assessments. For more information on RAMP see: https://pdf.usaid.gov/pdf_docs/PA00THHW.pdf .

[8] The steps taken during the PTI assessment process are explained in detail in Appendix 7.

[9] More information can be found here: https://earlygradereadingbarometer.org/

What

The LRF! C/U model involves a practice-focused pedagogy and student practice book developed by cognitive psychologist Dr Helen Abadzi and the Al Qasimi Foundation[10] 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 (Abadzi, 2020). It has been developed to help students build 'low level' neurological functions: rapidly distinguishing letter shapes, chunking, and decoding sounds and words. Altogether, the programme comprised the following elements:

Awareness raising sessions

School principals and directorate supervisors for resource room teachers attended a three-hour awareness session conducted by QRTA to receive information about LRF!.

Teacher training and coaching

QRTA carried out a two-day training session for resource room teachers on how to use the LRF! C/U method and practice book. The training aimed to provide participants with an understanding of:

- the rationale for the LRF! C/U model,
- the role of resource room teachers in delivering the LRF! C/U model,
- the learning experience teachers are being asked to facilitate, and
- how to support the involvement of parents/carers – for example, encouraging students' use of the LRF! practice book at home, and supporting parents in this through WhatsApp messages.

The training included opportunities to practise the new teaching and learning techniques, and to explore potential barriers and how they can be overcome. Following the training sessions, resource room teachers received three coaching visits and access to fortnightly online meetings. The coaching visits were delivered by QRTA staff, using a coaching model that had been designed to enhance implementation effectiveness.

Delivery of LRF! C/U sessions

The LRF! C/U model support is a form of small group tuition that was delivered by resource room teachers in the resource room, using the practice book. Eligible students received the intervention for 12 to 14 weeks, with three 30-minute sessions per week. The LRF! C/U model support was intended to be delivered to groups of five to six students per session, who have similar literacy learning needs.

LRF! C/U sessions were provided during the normal school day. The scheduling of sessions was agreed between classroom teachers and resource room teachers in each school. The Ministry of

[10] https://www.alqasimifoundation.com/

Education (MoE) gave permission to use one 'free activity period'[11] and two 30-minute sessions from two Arabic periods allocated on different days to implement this intervention.

The LRF! C/U model adopts an *'I do, we do, you do'* pedagogical approach using the practice book. First, using large versions of the textbook, resource room teachers introduce the letter-sound, or letter combinations, and model how to 'read' it ("*I do*"). This is followed by an opportunity for students to practise 'reading' using the echo reading method[12] ("*we do*"). These two steps are completed in the first 10 to 15 minutes of the session. Following this, resource room teachers ask students to independently work through the student 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, the teacher's role is to encourage engagement with the task and provide one-to-one feedback (reinforcing the fluency of reading or the actual improvement, correcting reading mistakes and writing down notes to follow up and keep track of each student). This stage of independent practice with teacher feedback lasts for around 15-20 minutes. This is a key feature of the LRF! C/U model, as research indicates that individuals need to independently and repeatedly practise decoding to develop the automaticity needed for fluent reading (Abadzi & Martelli, 2014).

Delivery of materials to schools

All students selected for the programme received a copy of the LRF! C/U model practice book, which had been updated based on results from the previous pilot. Students were encouraged to take the practice book home for extra practice with their parents/carers.

Parental engagement

Schools and resource room teachers also engaged with parents and carers to involve them in the LRF! C/U model and encouraged them to support students at home.

Schools held an awareness-raising meeting for parents at the start of the programme. Schools and resource room teachers also used different communication channels (for example, WhatsApp messages to parents) about the support needed for practice at home following the lessons.

How

In the programme, students are taught to process written text more quickly by first repeating individual letters and words to the point of automation. This is intended to enable them to decode words 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 (Alzubi & Attiat, 2021).

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 (Rudnicky & Kolers, 1984), so the practice book uses large font sizes and spacing.

[11] There is a set curriculum for the free activity period that teachers implement with students; this primarily focuses on enhancing personal skills and values.

[12] Echo reading is when students will repeat out loud what the teacher has read.

- The Arabic script is dense and visually complex, with students often identifying words and understanding their meaning slowly (Abadzi, 2012).[13] The practice book and the LRF! model recognise this, and place importance on repetition and teacher feedback, in order to encode Arabic script into memory.
- New letter shapes are introduced slowly, one by one.
- A phonics-based approach is followed in which students gradually decode words using their phonics knowledge, rather than using other clues or seeking help.
- 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).
- Students need to see meaning in text, so real words and sentences are introduced as soon as possible.
- The use of pictures in the text is minimised to ensure students learn letter sounds, rather than guessing.
- The practice book includes audio recordings accessible via a QR code, so students can listen at home.

The student practice book stresses repetition of patterns, alongside lots of practice in recognising them. See the examples below in Figure 1.

Figure 1: Excerpts from the student practice book

[13] A student reading Arabic script has to undertake several stages to comprehension. They must decipher the text, predict the vowels but keep multiple alternative words in their working memory to test the meaning, and then make linguistic sense (Abadzi, 2012).

الْمالُ	رِمالُ	رُموزُ	لَذيذُ
مَدَدُ	مِزْمارُ	الْوادي	لامَ
رَمْزُ	رَمْلُ	مالَ	مَوْلِدُ
رامي	مَوْرِدُ	مُرادُ	مارِدُ
مَلاذُ	مَدى	لَزِمَ	مَوْزُ

اقْرَأُ الْجُمَلَ وَالْعِباراتِ الْآتِيَةَ:

الْمَوْزُ لَذيذٌ.

49

لَزِمَ	مَدَدُ	مَوْزُ	رَمْزُ
رَمْلُ	مَوْرِدُ	مَوْلِدُ	زَمْزَمُ
مِزْمارُ	الْمالُ	لَذيذُ	الْوادي
الْوالي	دَليلُ	الْوَزيرُ	ماذا
مَدارُ	الْوالي	رَمى	زَمْزَمُ
مالي	دَليلُ	مَلَلُ	رامَ
مُرورُ	الْوَزيرُ	مَرامُ	ميلادُ

48

اقْرَأُ الْجُمَلَ وَالْعِبارَاتِ الآتِيَةَ:

يَعْدو الْمُهْرُ وَراءَ الْأَرانِبِ.

أَعْطاني والِدي هَدِيَّةً يَوْمَ مَوْلِدي.

مَتى أَزْهَرَتِ الْوُرودُ؟

أَزْهَرَتِ الْوُرودُ يَوْمَ الْأَرْبِعاءِ.

مَتى رَنَّ الْمُنَبِّهُ؟

رَنَّ الْمُنَبِّهُ بَعْدَ الْغُروبِ بِثَلاثينَ ثانِيَةً.

أَوَدُّ أَنْ أَنامَ بِهُدوءٍ تامٍّ.

لِماذا رَمى وائِلٌ ثِيابَهُ؟

رَمى وائِلٌ ثِيابَهُ لِأَنَّها مُهْتَرِئَةٌ وَرَثَّةٌ.

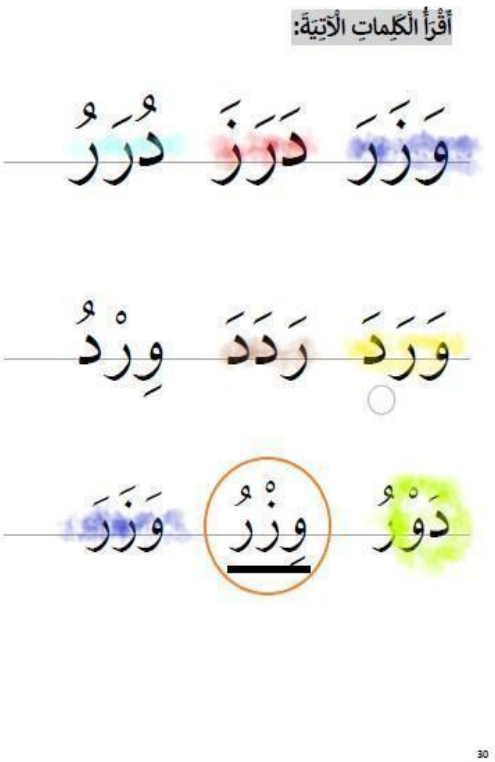
هَلْ تَوَدُّ أَنْ تَتَناوَلَ الْغَداءَ؟

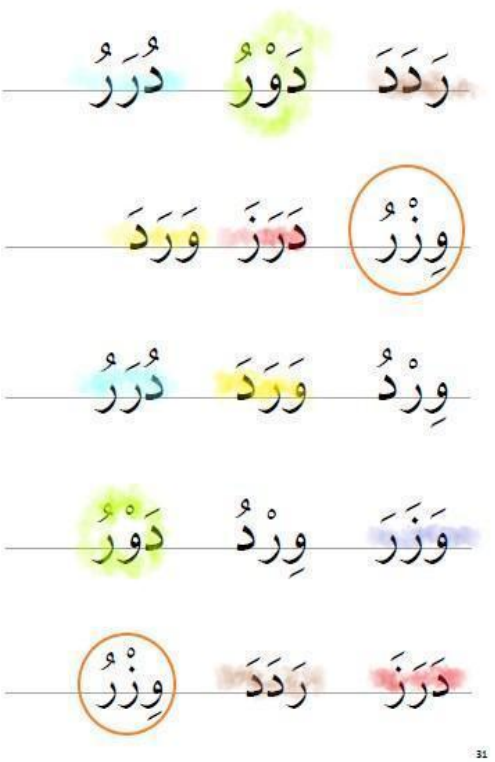
نَعَمْ، أَوَدُّ أَنْ أَتَناوَلَ الْغَداءَ.

ماذا تَناوَلْتَ عَلى الْغَداءِ؟

تَناوَلْتُ الْأَرُزَّ الْمُبَهَّرَ مَعَ اللَّبَنِ.

لا أَتَناوَلُ الْمِياهَ الْغازِيَّةَ أَبَدًا.


أَقْرَأُ الْكَلِماتِ الْآتِيَةَ:
وَزَرَ دَرَزَ دُرَرُ
وَرَدَ رَدَدَ وِرْدُ
دَوْرُ وِزْرُ وَزَرَ
30


رَدَدَ دَوْرُ دُرَرُ
وِزْرُ دَرَزَ وَرَدَ
وِرْدُ وَرَدَ دُرَرُ
وَزَرَ وِرْدُ دَوْرُ
دَرَزَ رَدَدَ وِزْرُ
31

When

The LRF! C/U model was delivered across at least 12 of 14 weeks of the school semester, with at least two out of three of the 30-minute LRF! C/U sessions delivered per week. A detailed timeline is included in the *Timeline* section below.

Tailoring

While the content of each session was set, resource room teachers did have some flexibility over how they could facilitate it. They were encouraged to draw on their professional judgement to tailor instructions according to students' needs. They were also expected to ensure that they progressed through the content of the practice book, while at the same time ensuring students were able to adequately master each lesson as they did so.

Updated logic model

The LRF! C/U logic model was updated in response to findings from the previous pilot and is shown in Figure 2 below. The updating process included a logic model workshop in August 2023 attended by key members of QRF and the evaluation team. Moving from left to right, its components are:

Inputs describe the resources required to develop and deliver the C/U model.

Activities describe the work required to implement the C/U model. This is divided into the organisational activities carried out by QRF and QRTA, and the activities carried out in schools to deliver the programme.

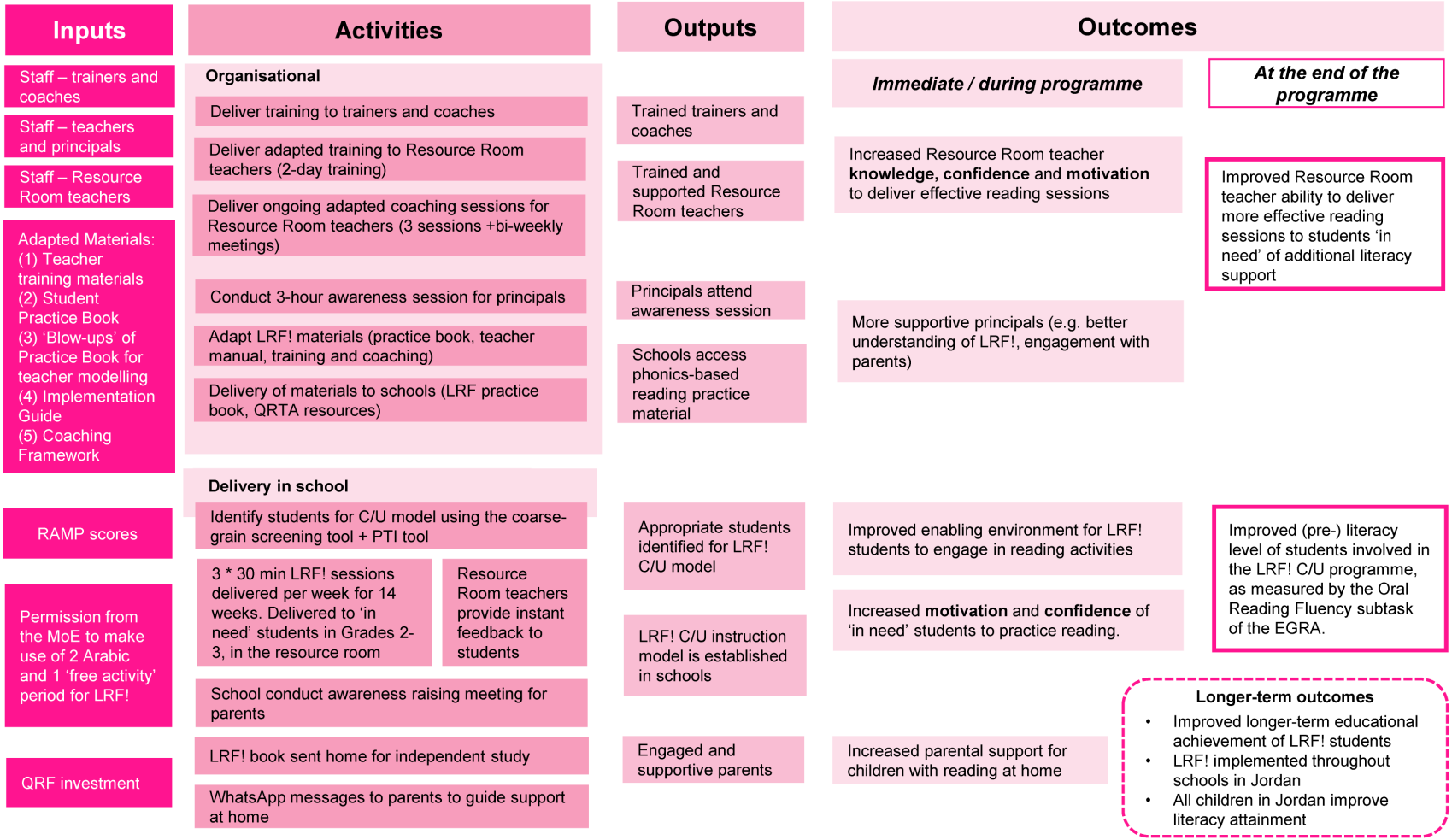
Outputs describe the product of the activities; these are divided in line with the above.

Outcomes describe the results that will be achieved if outputs are delivered. We distinguish outcomes that will arise during implementation, and those that will arise at the end of the programme.

Longer-term outcomes are shown in a box in the bottom right-hand corner of the logic model. These are the results that are expected to emerge in the future. This captures the longer-term vision of the LRF! C/U model.

The boxes of the revised logic model that are **outlined in black** show the key components of the revised C/U approach that are the focus of this pilot evaluation. Due to changes in QRTA's delivery approach for this pilot, the activities, outputs and outcomes in the logic model that relate to classroom teachers were removed from the logic model.

Figure 2: Updated LRF! C/U model logic mode



Notes:
a) Activities, outputs and outcomes relating to class teachers have been removed from the logic model to reflect QRTA's adapted approach to C/U model training
b) Boxes outlined in black are key elements of the revised pilot implementation and evaluation; other boxes will not be the focus of the revised evaluation

RESEARCH QUESTIONS

The research questions addressed by the current pilot are listed below. We have detailed whether each question is addressed by IE or IPE methods, or a combination of both. As this is not a full-scale pilot and is instead looking at key aspects of the revised LRF! C/U model based on findings from the previous pilot, the majority of research questions were developed to be investigated by IPE activities. Some questions concerning the feasibility of an efficacy trial have already been well explored in the previous pilot and did not need to be explored again.

Evidence of promise

1. In what ways, and to what extent, does the LRF! C/U model affect school, teacher, and student practice as compared to usual practice teaching and learning? (IPE)
2. How do teachers perceive the intervention and any changes that it has delivered? (IPE)
3. Is there evidence to support the revised logic model? (IE and IPE)
4. Is there any evidence of unintended consequences (negative or positive) as a result of the implementation of the LRF! C/U model? (IPE)

Feasibility of intervention

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

Assessing feasibility 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 usual practice groups? (IE and IPE)
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 (including the PTI diagnostic tool)? (IE and IPE)

Assessing readiness for trial

13. What changes, if any, are needed to the logic models? (IPE)

14. What changes to the intervention, implementation models, support, or materials need to be made? (IPE)

15. What can we learn from the pilot about minimal detectable effect size estimates, intra-cluster correlations, pre-and-post correlations, and sample sizes? (IE)

Project teams

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 and led EGRA testing, IPE data collection and contributed to evaluation design, analysis and reporting. Members of the evaluation and delivery team are outlined in Table 2 and Table 3, respectively.

Table 2: Evaluation team

Name	Project role	Organisational role
Enes Duysak	Principal Investigator and strategic lead	Research Director, Evaluation, NatCen
Hayley Leonard	Implementation and Process Evaluation lead	Research Director, Children and Families, NatCen
Abbi Rennick	Day-to-day project manager	Senior Researcher, Children and Families, NatCen
Charlotte Bessant	Project management and implementation and process evaluation support	Researcher, Children and Families NatCen
Rebecca Parker	Project management and impact evaluation support	Researcher, Evaluation NatCen
Anjhana Damodaran	Impact evaluation support	Senior Researcher, Evaluation, NatCen
Andi Fugard	Impact evaluation oversight and QA	Co-Director of Evaluation, NatCen
Gayle Munro	Implementation and process evaluation oversight and QA	Director of Children and Families, NatCen
Nedjma Koval	CEO	Integrated
Samah Goussous	Senior Project Manager	Integrated
Marwa Alsamneh	Logistics Officer	Integrated
Rasha Al-Khateeb	MEL Officer	Integrated
Taimaa Khalaf	MEL Officer	Integrated

Table 3: Delivery team

Name	Department/Project role	Role and team
Rola Said	Academic Affairs (Advisor)	Director of Education - QRTA
Dr. Ibrahim Al Awadi	Academic Affairs (Advisor	CPD Manager
Lubna Al Drainy	Academic Affairs (Technical project lead and coach)	Senior Education Specialist - QRTA
Mohammad Salameh	Academic Affairs (Coach)Reviews of research design, plans, data, and tools	Senior Teacher Educator - QRTA
Sana Al Syouf	Project Management (manage the project plan and arrange logistics)	Project Manager - QRTA
Fadi Al Mobaidein	Support Services (training and material procurement)	Procurement & Logistics Manager - QRTA
Ala'a Saymeh	Finance (keep track of project financial expenses and reporting)	Finance Manager - QRTA
Robert Palmer	Oversight and executive guidance	Executive Director, Research and Program Development - QRF
Rami Al-Assad	Project Management Office - Project Management	PMO Associate Manager - QRF
Haneen Al Abed	Review of training and intervention materials to reflect changes from previous trial; liaison with coaches	Arabic Literacy Technical Associate Manager - QRF
Rawan Awwad	Reviews of research design, plans, data, tools, and reports	Monitoring, Evaluation and Learning Manager - QRF

DATA PROTECTION, ETHICS, AND TRIAL REGISTRATION

Data Protection

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 was undertaken by Integrated, with students assigned a unique identifier. Test results were submitted to EGRA's Tangerine tool[14] and a pseudonymised dataset was transferred to NatCen via NatCen's Secure File Transfer Protocol (FTP) server. 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 can only be accessed by authorised personnel. At Integrated, 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.

[14] Tangerine is a tool used for data recordings for tests like EGRA.

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 carry out regular testing to ensure this process is 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.

For this evaluation, NatCen is a data controller and Integrated is a data processor. Furthermore, the only potential personal data to be transferred by NatCen to QRF for this evaluation could be the name, position and email address of key project contacts in relation to their contribution to the evaluation and to provide these individuals with further details of the evaluation. For this evaluation, QRF and NatCen each are a data controller in respect of their own processing of this specific data.

NatCen agreed a data retention period of seven years from the project inception with QRF. Once this period has expired, all relevant parties 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. Ethical approval was granted in November 2023. 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 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 participants are not identifiable in the outputs

Trial registration

The trial was registered on the Open Science Foundation in April 2024.

(https://doi.org/10.17605/OSF.IO/J35WZ)

METHODS

Pilot design

The pilot evaluation was designed as a two-arm cluster randomised controlled trial (RCT), with schools as the unit of randomisation and students as the unit of analysis. Eight schools were randomly allocated to receive the LRF! C/U model and eight schools were randomly allocated to usual practice. Within recruited schools, three classes per grade were randomly selected for the evaluation (if there were more than three classes per grade), and students were screened for their eligibility for the LRF! C/U model. Eligible students took part in baseline and endline EGRA testing and were included in the IE analysis for this pilot.

Due to this being a pilot evaluation, the study was not intended to provide a robust estimate of the causal impact of the LRF! C/U programme. Consequently, while this pilot evaluation involves comparing literacy outcomes between those receiving the LRF! C/U model and usual practice, this evidence should be treated as indicative of the potential promise of this intervention, and not as a robust causal estimate.

Table 4 presents the pilot design, including unit of randomisation, stratification, variables and measures of outcomes.

Table 4: Study design for the revised LRF! C/U model pilot

Trial design, including number of arms		**Pilot cluster RCT**
Unit of randomisation		School level
Stratification variables (if applicable)		Regions and urban/rural classification
Primary outcome	**Variable**	Oral reading fluency
	Measure (instrument, scale, source)	**Source:** EGRA Grade 2 assessment with the addition of a set of pre-literacy items **Instrument**: EGRA + pre-literacy tool **Subdomain**: Oral reading fluency
Secondary outcome(s)	**Variable(s)**	Arabic literacy attainment and specific sub-domains of Arabic literacy attainment
	Measure(s) (instrument, scale, source)	**Source:** EGRA Grade 2 assessment with the addition of a set of pre-literacy items **Instrument:** EGRA + pre-literacy tool **Sub-domains:** Arabic literacy attainment, letter sound identification, syllable identification, reading comprehension, word decoding and listening comprehension
	Variable	Oral reading fluency

Baseline for primary outcome	**Measure (instrument, scale, source)**	**Source:** EGRA Grade 2 assessment with the addition of a set of pre literacy items **Instrument:** EGRA + pre-literacy tool **Sub-domain**: Oral reading fluency
Baseline for secondary outcome	**Variable(s)**	Arabic literacy attainment and 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:** Arabic literacy attainment, letter sound identification, syllable identification, reading comprehension, word decoding and listening comprehension

School recruitment

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

- They were not part of the previous pilot.
- They had students in Grades 2 and 3.
- They were a single shift school.[15]
- They were not taking part in any other literacy interventions, other than the RAMP that has been carried out in Grades 1-3 in all MoE schools in Jordan since 2015.
- They had a resource room teacher.
- They were not delivering blended teaching.[16]
- They were not in a Syrian refugee camp.[17]
- They were from the middle, near north and near south[18] regions.

The recruitment process involved the following steps:

- QRF shared an updated list of schools in Jordan, including school characteristics and contact information.
- NatCen researchers cleaned the data and excluded schools that were not eligible for the trial according to the criteria above. NatCen then randomly ordered the list of eligible schools.
- NatCen shared the randomly ordered list of schools with Integrated.
- Integrated followed the list to approach schools for their consent to participate.[19]

School recruitment took place in November 2023. QRF supported the recruitment process by leading on all communications with the MoE. The MoE had already provided written permission for the intervention and evaluation to take place; they also approved the EGRA and PTI tools for use in schools. NatCen and QRF drafted school and parent information sheets, covering the details of the programme and evaluation. The

[15] In Jordan, schools may operate on one shift or two shifts. Schools operating on two shifts (morning and afternoon shift) have different groups of students in the morning and afternoon shift, while schools operating on one shift have one group of students during the whole day.

[16] In September 2021, the MoE decided to implement blended teaching in some schools due to the COVID-19 pandemic. Schools that are required to implement blended teaching have 2-3 days face-to-face teaching and 2-3 days pre-recorded online teaching in a week. This form of blended teaching would have a negative effect on the implementation of LRF! C/U and would not reflect usual practice in schools. Therefore, schools that use blended teaching will not be eligible for the pilot.

[17] When schools in the Syrian refugee camps and single shift schools were excluded from the list of eligible schools, we would be excluding all schools in the refugee camps and the Syrian evening schools from the list of eligible schools. Our final list of eligible schools included schools that had either Syrian refugee students who are integrated to the Jordanian Educational System or no Syrian refugee students. The term “integrated schools” is used by the MoE and indicates schools where Syrian refugee children and Jordanian children are taught together in the same classrooms.

[18] The North and South regions were narrowed to near north and near south for logistical reasons.

[19] In two regions, deviations from the established recruitment process occurred. Specifically, instead of contacting schools according to the randomly ordered list, Integrated asked contacted schools to refer other schools for participation. This led to instances where schools not originally in the immediate sequence were contacted earlier than scheduled. No schools outside the established recruitment list were contacted.

teachers from the participating primary schools were then asked to distribute parent information sheet with parents/carers. Following this, parents/carers opted their children out if this was their preference.

The whole recruitment process is depicted in the *Participant flow including losses and exclusions* section, below. In total, 16 schools were recruited for the pilot evaluation from Amman, Ajloun, Jerash, Karak, Balqa and Madaba governorates.

Randomisation

Schools recruited for the pilot were randomly allocated to receive the LRF! C/U model or usual practice.

We adopted the same stratified randomisation approach that was used in our previous LRF! pilot. 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)[20]. Amman, the country's capital, and the biggest city in Jordan, is located in the middle region. To equally represent schools that are in the middle region but not in Amman, we divided the middle region into two geographical regions: Amman, and the remainder of the middle region. In addition, we formed near north and near south regions for logistical reasons. The near north region covers Jerash and Ajloun governorates and the near south region covers Karak governorate only. Therefore, we had four geographical regions altogether: (1) Amman, (2) middle (excluding Amman:), (3) near south and (4) near north. Given that there were four geographical regions, and a school could be either in a rural or urban area, we had eight strata altogether.

Randomisation was carried out by an analyst at NatCen in December 2023. Randomisation was undertaken in Stata 17 and both the 'do' and 'log' files were saved as a record of the randomisation process.

Participant selection

As discussed earlier ('Selection of students for LRF! C/U model support'), students were identified for the intervention using a combination of routine coarse-grain assessment data, the PTI diagnostic tool, and the EGRA assessment. This process is described in more detail in this section.

All classes in Grades 2 and 3 were eligible for the LRF! C/U model. However, due to the resource room teacher's limited capacity, three classes per grade were randomly selected (if there were more than three classes per grade). At the start of the school term in September 2023, all classroom teachers administered the coarse-grain screening tool developed by the RAMP with their students in Grades 2 and 3. The coarse-grain screening tool is part of usual practice in Jordanian schools at the start of the school year. QRTA compiled the scores from the coarse-grain reading assessment and identified the lowest performing 20% of students in each class based on these scores. This information was shared with QRF, PTI and NatCen.

For the lowest performing 20% of students only, practitioners from PTI carried out an additional screening tool between December 2023 and February 2024, which was made up of multiple processes to judge inclusion/exclusion in the intervention. This involved first having a telephone call with parents or carers to ask about the students' developmental history, lasting between 15 and 45 minutes. Then PTI carried out a language development test[21] and learning difficulty test[22] with students. These tests were carried out by

[20] In Jordan: North region consists of Irbid, Ajloun, Jerash, and Mafraq governorates. Middle region consists of Balqa, Amman, Zarqa, and Madaba governorates. South region consists of Karak, Tafilah, Ma'an, and Aqaba governorates.

[21] The language development test assesses speech organs, evaluates linguistic, receptive and expressive skills. It also involves examining visual and auditory discrimination and assessing visual and auditory perception and memory.

[22] The learning difficulties test includes tests for auditory discrimination, comprehension of same-meaning sentences, following instructions, understanding others' speech and understanding pronouns. The test, in general, evaluates difficulties in understanding and comprehension, speech capabilities and reading and writing skills.

PTI's specialist team and lasted about 30-45 minutes per student, with breaks as required by the student. If a student exhibited indicators of delay in language skills using the language development and/or the learning difficulties tests, they were identified by PTI to have delays in language development and were not recommended for the LRF! C/U model. Additionally, if the PTI therapist noticed indicators of global developmental delay using these tests and/or the student was reported to have indicators of global developmental delay by parents/carers during the screening phone call, these students were also not recommended for the LRF! C/U model.[23] If students were identified with learning difficulties that were not indicative of global developmental delay or language delay (such as reading or writing difficulties), they were recommended for the intervention, as the programme is suitable with these students. A comprehensive report about the student was then created, which explained their linguistic abilities and described any learning difficulties with reading and writing.

After students were screened using the PTI tool and those with potential SEND were excluded from further testing, the remaining students took part in EGRA testing, described in the following section, in January and February 2024. Since LRF! aims to improve students' reading fluency, ORF is crucial for identifying those eligible for the C/U model. Students categorised as non-readers or beginning readers – those scoring 29 correct words per minute or less – were deemed eligible for the programme. The reading proficiency categories used in this assessment are based on RTI International (2023).[24]

Impact Evaluation methods

Outcome measures

The previous pilot evaluation combined a set of pre-literacy items and the Jordanian EGRA to create a single learning metric for reading. This tool is referred to as EGRA+pre-lit. The process of developing the pre-literacy items and rationale for doing so are described in detail in the previous pilot report (Dimova et al., 2023). During the set-up meetings, it was agreed to use the same tool that was used in the previous pilot.

As the LRF! C/U model mainly focuses on improving students' reading fluency, NatCen and Integrated agreed with QRF and EEF that the primary outcome measure would be ORF. This was obtained by administering the EGRA+pre-lit. As part of EGRA+pre-lit testing, students 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. This ensures differentiation between students if students read the passage in less than 60 seconds.

Acquisition of reading skills is regarded as a developmental process (Chall, 1996). Reading fluency and comprehension are not standalone skills, but they are built on other skills, such as letter sound identification and decoding. Therefore, it was important to assess the effect of the programme on reading skills that could be predictive of future successful reading (RTI International, 2015).[25] Specific sub-domains of Arabic Literacy Attainment (ALA) as measured by the EGRA+pre-lit were included as secondary outcome measures with separate metrics produced for each of these sub-domains from the EGRA+pre-lit. For the timed sub-domains of the EGRA, the per-minute score was calculated as explained by the EGRA toolkit (RTI International, 2015). This scoring system allows differentiating students who have got the same number of

[23] The steps taken during the PTI assessment process are explained in detail in Appendix 7.

[24] RTI International (2023) provides the categories of reading proficiency identified for Jordanian Grade 2 and 3 students. Students were categorised into four categories based on their Oral Reading Fluency (ORF) score: non-readers, beginning readers, progressing readers, and proficient readers.

[25] Please note that the word decoding and listening comprehension sub-domains of EGRA are additionally administered and assessed in the current pilot evaluation compared to the previous pilot evaluation.

correct items but finished in different time periods. The per-minute scores were calculated using the following formula:

$$Per - minute\ score = \frac{Subtask\ score}{(Time\ given - Time\ remain)} * 60$$

The first sub-domain of the EGRA is Letter Sound Identification (LSI). It assesses a student's ability to associate sounds with letters. In this task, students were given 100 letters and asked to read the letter sounds within one minute. A per-minute score LSI was created to assess students' ability on LSI.

Students also completed a sub-domain of the EGRA which is used to assess their ability to identify syllables. In this sub-domain, students 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 the student's ability to identify syllables.

Students also completed a sub-domain of the EGRA which is used to assess students' decoding ability (i.e., the sub lexical route of word processing (RTI International, 2015). This sub-domain of the EGRA assessment included a list of 50 one- and two-syllable nonwords to read within one minute. The number of correct nonwords was used to create a per-minute score for decoding nonwords.

Furthermore, once students completed reading the short story for the ORF measure, they were asked five questions related to the short story to assess their Reading Comprehension (RC). They were given a maximum of 15 seconds for each question. The RC score is the number of correct answers out of a maximum possible score of five.

The last sub-task of EGRA implemented in this pilot trial was Listening Comprehension (LC). This sub-task of EGRA included a passage read by the assessor and was followed by oral comprehension questions answered by the students. During one minute, students were asked five questions about the passage. The LC score is the number of correct answers out of a maximum possible score of five.

As the primary outcome measure changed from ALA in the previous pilot evaluation to ORF in the present evaluation, the ALA was included as part of the secondary outcome measures to maintain consistency with the previous pilot evaluation. The ALA was a single metric derived from the sub-domains of the EGRA, each of which has a different scale due to the varying number of questions. Standardising each sub-scale is essential; without standardisation, some sub-domains would disproportionately influence the average score. To standardise the sub-domains, each was adjusted to have a mean of zero and a standard deviation of one. We then calculated the average of these standardised scores to create the ALA metric. Additionally, we calculated Cronbach's alpha to assess the internal consistency of the measure. EGRA+pre-lit was administered at both baseline (February 2024) and endline (May 2024). The primary and secondary measures at baseline were the same as those at endline. All eligible students from both C/U and usual practice schools were included in baseline and endline assessments.

Impact evaluation analysis

Primary outcome analysis

In line with the EEF analysis guidance[26], the primary outcome analysis followed an Intention-To-Treat (ITT) approach[27]. The analysis used a three-level model to account for the clustering of students (level 1) in classes (level 2) within schools (level 3). Random assignment to conditions was at the school level (level 3). This model included school and class-level random intercepts and accounted for the baseline ORF and the stratification variables. The basic form of the model is:

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

where students (*i*) are clustered in classes (*j*) within schools (*k*). The intervention effect was estimated by β_2, β_3 represented a vector of strata fixed effects for the schools (i.e., their geographical location), β_0 was the intercept, β_1 was the slope for baseline scores, v_k a school-level random effect, u_{jk} a classroom-level random effect, and e_{ijk} the residual term. In line with the EEF analysis guidance[28], other additional covariates were not considered. This was to ensure comparability, transparency, and reproducibility in scientific research. The primary analysis also included detailed descriptive analysis: histograms, means, and standard deviations, for all measures, groups, and time points. The analysis was carried out using Stata 17. Both the syntax used, and outputs of analysis were saved as a record of the process.

Where a sample size is small, the standard statistical tests may not perform well on non-normally distributed measures. As this is the case in our primary outcome measure (please see findings section for more information on this), we used a non-parametric bootstrap method to deal with the non-normality (Sainani, 2012). The estimated effect of the intervention is presented with 95% confidence intervals to capture the associated uncertainty. Following Campbell (2008), we calculated bootstrapped standard errors and constructed bootstrapped confidence intervals with 1000 replications.

As this is a pilot evaluation of the LRF! C/U model which is not powered to detect meaningful differences in outcomes, the focus of the IE is not on attributing causality for changes in outcomes to the LRF! C/U model. Rather it assesses evidence of promise supporting the revised logic model. Due to the small sample size of the trial, estimates of effect sizes will be highly imprecise. Furthermore, studies with lower statistical power have higher likelihood of Type I error, meaning that an evaluation concludes that a programme has had an impact when in reality it has had no impact. Similarly, studies with lower statistical power have higher likelihood of Type II error, meaning that an evaluation concludes that a programme has had no impact when in fact it has had impact (Christley, 2009). We therefore caution against drawing conclusions around the impacts of the intervention on any outcome measure.

Given the challenges associated with pilot evaluations and small sample size, it is recommended to focus on estimation and confidence intervals rather than formal hypothesis testing (Lee et al., 2014). For the IE, we specifically examined Confidence Intervals (CI) to inform our assessment of the evidence of promise. We interpreted the CIs in relation to the Minimum Relevant Mean Difference (MRMD), which indicates the smallest mean difference that is deemed significant for policy decisions (Lee et al., 2014). While there are no similar studies in Jordan that would enable us to infer this for our pilot study, we consider a small effect

[26] Evaluation design | EEF (educationendowmentfoundation.org.uk)

[27] The intention to treat (ITT) approach compares C/U model students with students receiving the usual practice irrespective of whether the intervention students actually receive the intervention. For more information please see: Gupta S. K. (2011). Intention-to-treat concept: A review. *Perspectives in clinical research*, *2*(3), 109–112. https://doi.org/10.4103/2229-3485.83221

[28] Evaluation design | EEF (educationendowmentfoundation.org.uk)

size - a Minimum Relevant Effect Size (MRES) of 0.2 standard deviations - as relevant for educational programmes.[29]

For each outcome measure, we transformed a MRES of 0.2 standard deviations into a MRMD using the relevant pooled standard deviation.[30] We then followed the steps below to assess the evidence of promise:

- **Check if the CI for an estimate crosses zero:** If it does, this suggests that there could be no difference between the C/U model and usual practice group.
- **Determine if MRMD falls within the CI:** If it does, this indicates a future efficacy trial could likely observe the MRMD at the appropriate level of confidence.

If our assessment shows that the CI for an estimate excludes zero and the MRMD falls within the CI, this will indicate evidence of the programme's impact on the relevant outcome measure.

For consistency and comparability with other evaluations, the difference in the means between the intervention and usual practice groups at endline was also expressed as a standardised effect size using Hedges' g, accompanied by 95% confidence intervals. Following EEF guidelines, the numerator was the unstandardised effect estimate given by β_2 in the multilevel model specified above, which adjusted for the baseline score and the stratification variable. The denominator was the unconditional pooled standard deviation of the primary outcome at endline. The formula used 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 students in both groups and $s_1{}^2$ and $s_2{}^2$ were the within-group variances in the primary outcome at endline.

We also reported school and class-level Intracluster Correlation Coefficients (ICCs) alongside 95% confidence intervals in analyses. We used the following multilevel model with a random effect by school, v_k a classroom-level random effect, u_{jk}, and the residual term, e_{ijk}. The model does not include 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 analyses

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

For all defined secondary outcomes (e.g., letter sound identification, syllable identification, and reading comprehension), we followed an ITT approach similar to that of the primary outcome analysis. This statistical model took the form of a multilevel model, where students were clustered in classes within schools, and took account of the stratification variables. The basic form of the model is:

[29] Cohen (1992) categorised 0.2 standard deviation as a small effect size, 0.5 standard deviation a medium effect size and 0.8 a large effect size.

[30] For example, for our primary outcome measure, the pooled standard deviation was 11.23. After multiplying this number with 0.2 standard deviations, we find the MRMD of 2.24.

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

where students (*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), v_k a school random effect, u_{jk} a classroom-level random effect, and e_{ijk} the residual term.

Similar to the primary outcome analysis, we used a non-parametric bootstrap method to deal with the violation of normality in secondary outcome measures. Estimates of the intervention effects were presented alongside uncertainty estimates (i.e. 95% confidence intervals). Following Campbell (2008), we calculated bootstrapped standard errors and constructed bootstrapped confidence intervals with 1000 replications.

For all secondary outcome measures, we also reported the effect size with confidence intervals at the 95% level using Hedges' g as previously described.

Additional analysis

As part of our further investigation, we conducted a descriptive analysis of the LRF! pre-literacy items administered to students at both baseline and endline assessments. The pre-literacy assessment tool includes the following items:

1. Oral vocabulary: Task 1
2. Oral vocabulary: Task 2
3. Recognise Letter Names
4. Recognise High Frequency words

The mean pre-literacy scores at baseline and endline for the C/U model group and usual practice group were calculated. For each pre-literacy item, we descriptively assessed whether students improved from baseline to endline and if there was a statistically significant difference between the C/U model and usual practice group at endline. The results are presented in Appendix 1.

While the protocol did not set out a coarse-grain screening tool assessment, as part of our additional investigations, we also evaluated the coarse-grain screening tool's effectiveness. The screening tool was intended to identify the lowest performing 20% of students for further diagnostic assessment by PTI. We used the Area Under the Receiver Operating Characteristic (AUROC) curve and Youden statistic to assess the effectiveness of the screening tool in identifying eligible students. The results are presented in Appendix 2.

Implementation and Process Evaluation methods

An Implementation and Process Evaluation (IPE) was carried out to address the evaluation domains set out in Table 5 below. The IPE domains of interest were informed by the EEF's IPE guidance (EEF, 2022) and IPE Handbook (Humphrey et al., 2019). The domains we assessed were fidelity (the extent that those implementing the LRF! C/U adhered to the model), including dosage (how much of the intended C/U model of LRF! had been delivered), quality (how well different components of the LRF! C/U model had been delivered), and adaptation (any changes made to the LRF! C/U model). In addition, we investigated reach (the rate and scope of participation in the LRF! C/U model), responsiveness (the degree to which participants engaged with the LRF! C/U model), perceived impact (the perception of teachers and coaches as to whether the LRF! programme had achieved its intended outcomes), and usual practice (what usual literacy teaching looked like in the absence of the LRF! C/U programme). Each of the research methods is mapped on to the IPE dimensions, research questions, and analysis methods in Table 5, and is described in more detail below.

Table 5: Mapping of IPE dimensions, RQs, data collection methods and analysis

IPE dimension	Research questions addressed	Research and data collection methods	Data analysis methods
Fidelity: Dosage	RQ3, RQ5, RQ13	Resource room teacher focus group discussion	Qualitative; framework analysis
		Resource room teacher end of programme survey	Quantitative; descriptive analysis
		Classroom teacher end of programme survey	
Fidelity: Quality	RQ5, RQ8, RQ9, RQ10	Resource room teacher focus group discussion	Qualitative; framework analysis
		Coaches focus group discussion	
		Resource room teacher post-training survey	Quantitative; descriptive analysis
		Resource room teacher end of programme survey	
	RQ5, RQ8	Classroom teacher end of programme survey	
Fidelity: Adaptation	RQ3, RQ5, RQ13, RQ14	Resource room teacher focus group discussion	Qualitative; framework analysis
		Coaches focus group discussion	
		Resource room teacher end of programme survey	Quantitative; descriptive analysis
		Classroom teacher end of programme survey	
Reach	RQ3, RQ6, RQ12, RQ13	Resource room teacher focus group discussion	Qualitative; framework analysis
		Coaches focus group discussion	
	RQ7	Resource room teacher post-training survey	Quantitative; descriptive analysis

	RQ3, RQ6, RQ12, RQ13	Resource room teacher end of programme survey	
		Classroom teacher end of programme survey	
	RQ1	Attendance data	
Responsiveness	RQ3, RQ7, RQ7, RQ10, RQ11, RQ13, RQ14	Resource room teacher focus group discussion	Qualitative; framework analysis
		Coaches focus group discussion	
	RQ7, RQ10, RQ11, RQ14	Resource room teacher post-training survey	Quantitative; descriptive analysis
	RQ2, RQ3, RQ7, RQ9, RQ10, RQ11, RQ13, RQ14	Resource room teacher end of programme survey	
	RQ2, RQ3, RQ10, RQ11, RQ13, RQ14	Classroom teacher end of programme survey	
Perceived impact	RQ2, RQ3, RQ3, RQ13	Resource room teacher focus group discussions	Qualitative; framework analysis
		Coaches focus group discussion	
		Resource room teacher end of programme survey	Quantitative; descriptive analysis
		Classroom teacher end of programme survey	
Usual practice	RQ1	Resource room teacher focus group discussion	Qualitative; framework analysis
	RQ1	Usual practice classroom/ resource room teacher focus group discussions	

Methods were chosen to address the research questions and focus on the changes to the pilot represented in the revised logic model (i.e., focusing on changes to training and support for resource room teachers, and on the screening process for eligible students). Data collection therefore focused on teachers (resource room and classroom), and coaches. To better understand the impact of the revised training model, surveys were conducted with resource room teachers directly after they received the initial training, and again at the end of the intervention. Broader issues relating to screening and perceived impacts on students and teachers were covered in the endline survey and were discussed in more depth in focus groups with resource room teachers and coaches, and in a survey with intervention classroom teachers.

To improve our understanding of usual practice when teaching reading (which may have changed since the last pilot), we conducted focus groups with mixed groups of resource room and classroom teachers from the teaching-as-usual group, as well as asking about usual practice amongst intervention teachers.

Data were triangulated across methods to address each research question, with quantitative data from surveys being incorporated into our thematic framework (see IPE Analysis section below). For RQs 3, 11, and 12, data from the IPE were used to complement IE data and to support the conclusions drawn from the analyses.
All data collection was administered by Integrated. Integrated has a team of trained enumerators who are experienced in collecting data within Jordanian schools. NatCen delivered an online training session to Integrated in March 2024 on qualitative data collection, in Arabic.

Training attendance data

Attendance data for each of the LRF! training and coaching sessions were gathered by QRTA and shared with NatCen by QRF. The data recorded the number of teachers who attended the two-day training, the three coaching sessions, and the bi-weekly coaching visits. This information was shared with NatCen in Excel and analysed to produce numerical counts of the number of attendees.

Surveys

All surveys were completed in Arabic, with Integrated enumerators asking participants the questions on a telephone call and inputting their responses using Alchemy survey software. At the beginning of the telephone call, verbal consent was given by participants, and they were reminded of their right to withdraw from the evaluation.

Surveys collected background data e.g., teacher and school information, as well as information about the number and types of sessions delivered. The remaining questions (topics listed below) asked teachers to select all options that applied to them, or to indicate their level of agreement with a number of statements on the following scale: Strongly agree, Agree, Neither agree nor disagree, Disagree, Strongly disagree (with an additional 'Don't Know' option). Enumerators read out statements or options one at a time and recorded the responses before presenting the next survey item. Results were shared with NatCen in Excel for analysis.

Resource room teacher post-training survey

The post-training survey consisted of multiple questions assessing resource room teachers' opinions of the training content and materials, the trainers, the location of training, and the training's objectives. All resource room teachers who attended the training completed the survey.

Resource room teacher post-intervention survey

For resource room teachers, the post-intervention survey covered opinions on the coaching visits and bi-weekly online meetings. Questions also asked about the coaches, the training objectives, the process of identifying students for the programme, implementation, student engagement, and perceived outcomes.

Classroom teacher post-intervention survey

Classroom teachers who had students in the C/U model were asked about implementation from their perspective, the process of identifying students for the programme, student engagement with LRF!, parental engagement, and perceived outcomes.

Focus group discussions

Focus groups took place with the following groups: QRTA coaches; resource room teachers (C/U model only); classroom and resource room teachers (usual practice group only). Participants from all schools in the intervention and usual practice groups were invited to focus groups, ensuring a spread of geographical regions and a mix of urban and rural settings were represented.

At the beginning of the discussions, verbal consent was given by participants, including consent to audio record the session. Attendees were also reminded of their right to withdraw from the evaluation.

Integrated adopted a flexible approach for the facilitation of focus groups, providing options to conduct them online or face-to-face. Online was reported to yield the best attendance, meaning all groups were therefore conducted in this way. In addition to being audio recorded, notes were taken during all focus groups which summarised the discussion. These were translated into English and shared with NatCen as supplementary information, prior to the audio files being translated.

Coaches

The focus group with coaches explored their perception of the programme, their observations of implementation to date, including enablers and barriers to delivery, and perceived impacts to date, and was designed to be 75 minutes long.

Usual practice group – classroom and resource room teachers

For the usual practice group, the aim was to understand different aspects of teaching-as-usual, including how students, including those who have difficulty reading, are typically taught to read, reflections on what enables and hinders teaching to read, and usual engagement with homework. These focus groups were designed to be 35-40 minutes long.

LRF! C/U model – resource room teachers

Lastly, the focus group with resource room teachers delivering the intervention aimed to explore their understanding and perceptions of the programme, their experiences of implementation to date, and their insights into any perceived impacts. This group was designed to be 60-75 minutes long.

The number of participants who completed each of the above methods are detailed below in Table 6.

Table 6: IPE methods and participant numbers

Method used	Number of participants
Post-training survey: Resource room teachers	8
Post-intervention survey: Resource room teachers	8
Post-intervention survey: Classroom teachers	37[31]
Focus group: Coaches	2
Focus group: Resource room teachers (intervention)	7
Focus group: Resource room and classroom teachers (usual practice)	26

IPE Analyses

Once the audio files were translated and transcribed, they were analysed using NatCen's Framework approach (Ritchie et al., 2003), adapted for bilingual working. The framework approach is a type of thematic analysis which evidences the relationship between themes and anonymised cases.

Using themes covered in discussion guides and any other themes identified in the data, we assembled a matrix in which each row represented a focus group and each column a theme/sub-themes. We then summarised the qualitative data in the matrix, including illustrative verbatim quotes where appropriate. Once all data had been coded in this matrix, we then moved onto analysis. This involved a phase of 'detection', including studying what participants said about a particular phenomenon, listing these and then sorting them thematically. Once we identified different themes in the data, we created higher-level categories that worked as meaningful conceptual groupings for participants' views and experiences within and across schools.

Summaries were produced in English and shared with the wider team for review. Core team members then came together for an analytical planning meeting, where key themes, patterns and issues identified in the qualitative data were discussed, and within team verification of findings established. Any findings where team members had alternative points of view were discussed and a consensus was reached.

[31] Integrated approached all 38 classroom teachers who were planned to have LRF! students. 37 completed the survey.

Descriptive statistics for attendance and survey data were produced using Excel. Relevant data from these sources were then incorporated into the analysis framework, providing additional detail to the themes identified through the qualitative analyses.

Findings were reviewed in detail and mapped against the pilot RQs and logic model in preparation for reporting.

Readiness for trial criteria

One of the main objectives of this pilot evaluation was to assess whether the LRF! C/U model is ready for an efficacy trial. In line with the previous trial, a list of criteria was developed specific to this 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 fell into the following categories: evidence of promise, feasibility of intervention, feasibility of the efficacy trial, and readiness for 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 engagement with LRF!, the practice book is used, the training is effective), and considered the quality and appropriateness of the intervention materials, training and the revised diagnostic tool. The criteria around the feasibility of the efficacy trial and the readiness for 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, and what changes may need to be made to allow this to happen. The success criteria are presented in Table 7, alongside their associated research questions. A full mapping of the success criteria to data collection methods and times is provided in Appendix 4.

Table 7: Success criteria and associated research questions for the revised LRF! C/U model pilot

Dimension	Research Question	Criteria (linked to logic model)
Evidence of Promise	EP1. In what ways, and to what extent, does the LRF! C/U model affect school, teacher, and pupil practice as compared to usual practice and learning?	**EP1a.** Resource room teachers use the updated LRF! practice book in more than half[32] of their literacy lessons and follow the pedagogical approaches outlined through the teacher manual, training and coaching. **EP1b.** Attendance data show that resource room teachers attended two-day training and three coaching sessions, and the majority of additional bi-weekly meetings with coaches. **EP1c.** Students use the practice book in more than half of lessons and show motivation to take it home.

[32] There was a total of 42 lessons (3 lessons per week for 14 weeks), meaning half would be 21.

	EP2. How do teachers perceive the intervention and any changes that it has delivered?	**EP2.** More than half of classroom and resource room teachers perceive that there is a positive value in this intervention compared to/in addition to usual practice: - Resource room teachers report positive outcomes of the new training and support system put in place by QRTA on their delivery of literacy lessons compared to usual practice. - Resource room teachers report increased knowledge, confidence and motivation to deliver effective reading sessions. - Resource room / Classroom teachers believe that students involved in the CU programme have increased confidence and motivation to practise reading. - Resource room / Class teachers agree that students involved in the LRF! C/U programme have improved (pre-)-literacy level.
	EP3. Is there evidence to support the revised logic model?	**EP3a.** Coaches and resource room intervention teachers report improved resource room teacher ability to deliver more effective reading sessions to students 'in need' of additional literacy support. **EP3b.** Coaches and intervention classroom and resource room teachers agree that students involved in the LRF! C/U programme have improved (pre-) literacy level. **EP3c.** Results from the EGRA tests suggest that the LRF! C/U programme could improve oral reading fluency and specific sub-domains of literacy attainment.[33]

[33] We need to be cautious about interpreting the results of the end line assessment given the small sample size.

	EP4. Is there any evidence of unintended consequences (negative or positive) as a result of the implementation of the LRF! C/U programme?	**EP4.** Teachers and coaches report minimal or no negative consequences as a result of the implementation of the LRF! C/U programme on: a) students participating in the programme b) other students in the class c) resource room or class teachers d) parents
Feasibility of Intervention (FI)	FI5. Was the LRF! C/U model delivered as intended in terms of dosage, nature and quality? What modifications were made, with what implications?	**FI5.** Resource room teachers and coaches report that the intervention was delivered as intended in terms of a) Dosage: at least two out of three of the 30-minute LRF! C/U sessions are delivered per week for nearly all of the programme. b) Nature: there is evidence that the practice book is being used regularly as part of the lesson, and teachers adopt all pedagogical approaches in each session ("*I do, we do, you do*"). c) Quality: LRF! C/U sessions were effective and engaging for students 'in need' of additional literacy support.
	FI6. What is the learning about the use of the PTI diagnostic tool? How successful is it, in use, at identifying the most appropriate students for the C/U model?	**FI6a.** Teachers and coaches agree that the coarse-grain screening tool + PTI tool effectively identified appropriate students for the CU programme.[34] **FI6b.** Teachers and coaches report minimal or no negative consequences as a result of the implementation of the PTI tool with students (on students or parents).
	FI8. To what extent do resource room teachers develop sufficient	**FI8.** Resource room teachers and coaches perceive that the resource room teachers have the

[34] The success criteria were agreed prior to the EGRA being confirmed for use for additional screening.

	skills and confidence through the training and coaching?	skills and confidence to effectively deliver the CU programme following the training and coaching sessions.
Feasibility of Trial (FT)	FT11. 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 usual practice groups?	**FT11a.** Evidence that there is enough in place to allow the intervention to take place the following year at scale: - school/participant retention rates during intervention and evaluation are high - the intervention materials and training is suitably defined and developed **FT11b.** More than half of eligible students[35] complete the outcome testing in both intervention and usual practice groups.
	FT12. What does the pilot tell us about the feasibility of the resources of an efficacy trial, e.g. measurement instruments or specific equipment used? (including use of the PTI tool)	**FT12a.** The training materials, practice books and measurement instruments are appropriate and meaningful, i.e. the EGRA and PTI tests provide relevant data about literacy attainment and developmental difficulties, respectively, and are age- and context-appropriate.

[35]Eligible students are defined as those who were included in the sample and received the intervention after applying inclusion/exclusion criteria from the coarse-grain screening tool, PTI, and EGRA assessments.

TIMELINE

Table 8: Timeline of the evaluation activities

Activity	Date
Logic model workshop	Aug 23
Logic model agreed	Nov 23
Coarse-grain screening tool carried out by teachers in Jordan, as per usual practice	Sep 23
Ethics approval complete for pilot evaluation	Nov 23
MoE approvals for training secured	Nov 23
Schools recruited	Nov 23
School randomisation	Dec 23
Pupil selection (from PTI tool)	Dec 23 – Feb 24
Pilot LRF! Catch up Model implementation - Semester 2	
Resource room teacher training and principal awareness sessions delivered	Jan 24
EGRA baseline testing	Feb 24
LRF! resources delivered to schools	Feb 24
Coaching delivered	Feb – May 24
Post-training surveys with resource room teachers (C/U model)	Feb 24
Intervention delivered (12 weeks)	11 Feb – 16 May 24
Focus group discussions with coaches	6-13 May 24
Focus group discussions with resource room and class teachers (usual practice group)	6-13 May 24
End of programme surveys with resource room and classroom teachers (C/U model)	6-13 May 24
Focus group discussions with resource room teachers (C/U model)	12-16 May 24
EGRA endline testing	19-23 May 24

FINDINGS

Participant flow including losses and exclusions

The participant flow diagram is presented in Figure 3 below. The diagram shows that out of the 4,070 schools in Jordan, 874 were identified as eligible for the trial. Of these 874 eligible schools, 125 were approached by Integrated for the pilot evaluation, resulting in 16 schools agreeing to participate. As part of the recruitment process, Integrated maintained a recruitment log which provided reasons for schools not participating.[36] Table 9 below shows that the main reason schools did not take part in the trial was Integrated being unable to reach them due to incorrect or disconnected phone numbers. This was the case for 64 out of the 109 schools that did not participate. Other reasons for non-participation are presented in Table 9.

Table 9 also shows the number of schools that agreed to take part in the evaluation and the number of schools that Integrated tried to reach by geographical region. The table illustrates clearly that it was more challenging to recruit schools in the near north, while it was easier to recruit schools in the near south. This was similar to the previous pilot, where school recruitment was easier in the near south in comparison to other geographical regions. Future recruitment processes should consider these regional differences.

[36] The recruitment process was designed to follow a randomly ordered list of schools in each region. Recruitment was to continue until two schools from each region agreed to participate, at which point recruitment would stop. However, in two regions, a deviation from this protocol occurred. After successfully recruiting two schools, the recruitment process continued, and an additional school in each region was contacted and agreed to participate in the evaluation. These additional schools, referred to as "Backup schools" were not included in this pilot evaluation since the required two schools had already been recruited. The "Backup schools" are not included in the counts in Table 9 and Figure 3.

Figure 3: Participant flow diagram

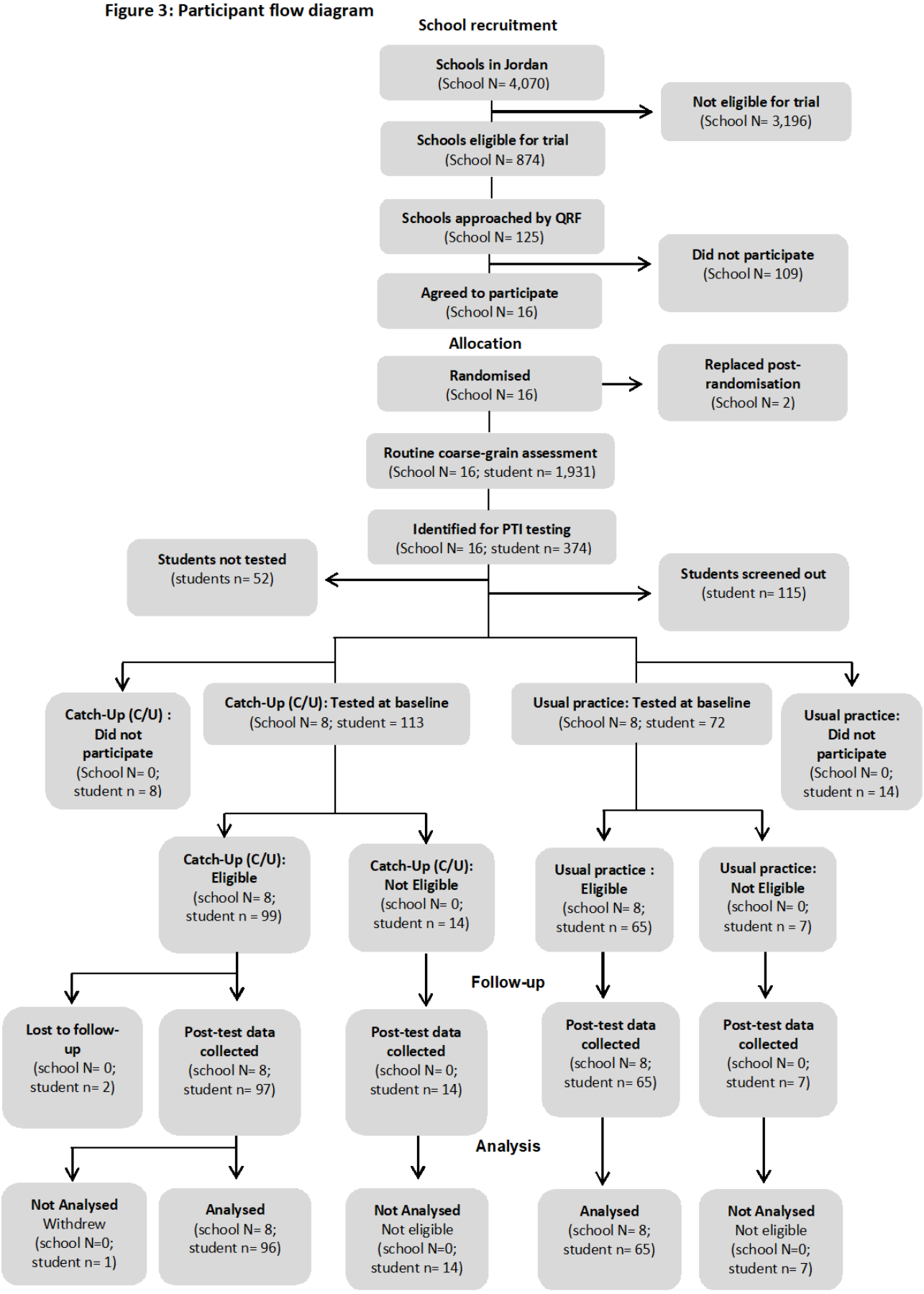


Table 9: Common reasons why schools across the different regions did not take part in the evaluation

	Number of schools by geographical regions				
Reasons for not participating	**Amman**	**Middle excluding Amman**	**Near North**	**Near South**	**Total**
Prefer not to participate	0	1	0	2	3
Unable to reach school/wrong number/disconnected	15	17	27	5	64
Taking part in another literacy intervention	0	0	1	0	1
Double shift school	0	0	1	0	1
No official resource room teacher	0	0	1	0	1
No resource room	4	5	25	3	37
Unknown reason	0	0	0	2	2
Number of schools invited to participate, but did not participate for the different reasons above	**19**	**23**	**55**	**12**	**109**
Number of schools agreed to participate	**4**	**4**	**4**	**4**	**16**
Total number of schools	**23**	**27**	**59**	**16**	**125**

The recruited 16 schools remained in the trial for randomisation, with eight schools assigned to the C/U model group and eight schools assigned to the usual practice group. However, following the recruitment and prior to baseline data collection, two schools were deemed ineligible for the trial: one school from the C/U model group and one school from the usual practice group. This was due to both schools already delivering a reading initiative and thus, not meeting the eligibility criteria[37]. Integrated followed the recruitment process explained in the School Recruitment section above and recruited two replacement schools from the random list of schools provided to them by NatCen during the recruitment process.[38]

Following the routine coarse-grain assessment at the start of the school term, 374 students in the 16 schools that agreed to take part in the trial were identified as the lowest performing 20% of students in their class. These students were invited to take part in an additional diagnostic assessment carried out by PTI. However, 52 students could not be tested due to factors such as absences and unavailable parental contact details. A further 115 students took part in the assessment and were screened out, as the assessment indicated that the C/U model may not be the most appropriate form of support for them.

[37] This information was not shared with Integrated during recruitment. These schools were identified by QRF, who were involved of the delivery of the other initiatives.

[38] While we acknowledge that this adjustment would not typically be undertaken in an efficacy or effectiveness trial, where strict adherence randomization is essential. However, as a pilot study, the primary focus was on assessing the feasibility of intervention and trial, and refining trial procedures, making this modification appropriate in this context.

Figure 3 shows that 207 students were invited to take part in baseline EGRA testing; 22 students (8 C/U model; 14 usual practice) did not participate.

Reasons for not participating in baseline EGRA testing were the following:

- 11 students were absent from school (50% of students that did not participate)
- two students did not participate due to parent objection (9.09% of students that did not participate)
- five students refused to participate (22.73% of students that did not participate)
- four students did not take part for an unknown reason (18.18% of students that did not participate)[39]

All students were also given the opportunity to opt out of the EGRA testing if they wished. This was firstly done through verbal agreement, and then more formally through students clicking a button at the start of the online EGRA test. No students opted out of the EGRA testing via verbal agreement at this stage.

In total, 113 students from the eight C/U model schools took part in baseline EGRA testing, with 99 students identified as eligible for the programme, due to scoring 29 correct words per minute or less for the ORF sub-scale of EGRA. In the eight usual practice schools, 72 students took part in baseline EGRA testing and 65 students were identified as eligible for the programme.[40]

At endline, data was collected for 97 eligible students in the eight C/U model schools and 65 eligible students in the eight usual practice schools.[41] Eligible students in the C/U model group (n= 2) did not participate in the follow-up because they were absent during the testing period. These students are listed under the 'lost to follow-up' category in Figure 3. All eligible students in the usual practice group participated in EGRA testing at endline.

Following EGRA baseline testing, one eligible student in the C/U model group withdrew from the trial. The data that had already been collected from them was deleted. They are listed under the 'Not analysed' boxes in Figure 3.

Furthermore, Figure 3 shows that 14 students in the eight C/U model schools and seven students in the eight usual practice schools were deemed ineligible for the programme, due to scoring 29 correct words per minute or more for the ORF sub-scale of EGRA. Despite being ineligible for the programme, data was collected for these students at follow-up. These students were excluded from the analysis and are listed under the 'Not analysed' boxes in Figure 3.

[39] Three students did not take part in baseline EGRA testing for unknown reasons but did participate in follow-up testing. However, due to the absence of baseline ORF scores, their eligibility for the programme could not be assessed and they were therefore excluded from the analysis.

[40] As this is a pilot study and has a small number of schools, there was a higher likelihood of having different group sizes between treatment and control groups. Hence, the schools that ended up in the C/U model and usual practice groups have different populations of students in Grades 2 and 3. In an efficacy trial with a sufficient number of schools, we would expect not to observe significant differences in groups sizes.

[41] The data collection team mobilised a "clean-up" crew to test students that missed EGRA follow-up testing and increase uptake.

At the analysis stage, data was available for 96 students in the eight C/U model schools and 65 students in the eight usual practice schools.

Attrition

Table 10 shows student level attrition from the trial. Baseline data was available for 99 eligible students in eight C/U model schools and 65 eligible students in eight usual practice schools. At the analysis stage, data was available for 96 eligible students in eight C/U model schools and 65 eligible students in eight usual practice schools.

The overall ratio between students analysed and students identified as eligible for the programme is 161 to 164. The retention rate is 98.17% and the attrition rate is 1.83%. In the C/U model group, the retention rate is 96.97% (attrition of 3.03%) with 96 students in the analysis from 99 students identified as eligible for the programme. In the usual practice group, the retention rate is 100.00% with no loss of students from those identified as eligible for the programme to those analysed.

Table 10: Student level attrition from the trial (primary and secondary outcomes)

		C/U model	Usual practice	Total
Number of students	Eligible for programme	99	65	164
	Analysed	96	65	161
Student attrition (eligible for programme to analysis)	Number	3	0	3
	Percentage	3.03%	0.00%	1.83%

Evidence of Promise

RQ1: In what ways, and to what extent, does the LRF! C/U model affect school, teacher, and student practice as compared to usual practice teaching and learning? (IPE)

This section presents findings which illustrate usual practice in terms of: teacher training and support, how students are identified as needing additional support with literacy, the teaching of literacy, and homework. The ways in which the LRF! C/U model differs from these existing practices are then presented.

Usual practice

Training and support

The usual amount of training and support received by both classroom and resource room teachers for teaching literacy varied. In the usual practice focus group, some classroom teachers suggested that they had only received training after initially becoming a teacher, while others shared more recent experiences of training. For classroom teachers who mentioned their previous training, they highlighted this was effective in helping them learn a variety of teaching strategies and understanding differences in students' reading ability.

> ***"We have been exposed to many training sessions, for example, on reading strategies, including some that I might have missed online, which were honestly very important."*** Usual practice focus group participant

Some resource room teachers spoke in focus groups about receiving a day of training on a yearly basis, but this was about fluency in general, and not specific to students who had difficulty reading. Other resource room teachers suggested they had not received training but would have found it useful.

Across focus groups with resource room and classroom teachers, participants highlighted the lack of previous guidance for teaching those struggling to read. This meant that teachers had to be proactive in researching on both the internet and from external texts to find information that supported them and their students. Resource room teachers also mentioned the benefits of exchanging advice with one another.

> ***"I haven't taken any courses at all over the span of nearly 20 years – none related to reading or, honestly, to resource rooms either...To be honest, there was nothing except visiting each other...[and sharing experiences between the teachers]"*** Usual practice focus group participant

Identification of students requiring additional literacy support

Across the usual practice focus groups, teachers spoke of a variety of ways that students in need of additional literacy support were identified. In these groups, there was agreement that informal tests were given to students in second and third grade at the beginning of the year in order to determine their academic level. Additional tests were sometimes administered in the resource room for students who struggle with reading. There was agreement within the usual practice focus groups that tests were supplementary to teachers' knowledge about students' reading ability, and often, teachers found it easy to diagnose those who were struggling. This was evident when the student could not link the sounds of the letters, which would have helped them to read continuously. Often, more value was placed on teacher opinion than the administered tests.

Finally, some classroom teachers described starting each lesson with basic literacy to gauge the reading ability of the class. Others would review student ability each week and increase the difficulty as time went on.

Teaching literacy

Within the usual practice focus group, teachers explained their varied approaches to teaching literacy. Some approaches involved reading exercises which had a particular focus on word segmentation, which helped students to pronounce letters and gradually increase their reading speed. Other teachers spoke about the importance of focusing on comprehension to help students to read independently. Some achieved this by creating stories for students based on their reading ability, with varying sentence difficulty and frequency of pictures, while others would display the words being learnt around the school building, to increase familiarity and engagement.

> ***"There are situations where a student reads but you feel they don't have comprehension. They read without proper expression, without intonation, just focusing on finishing the words. We've now moved to reading with understanding, with proper intonation, and at the right speed."*** Usual practice focus group participant

In addition, a few usual practice focus group participants provided students with higher literacy skills with external texts from subjects such as science, maths and religion. This was to increase their awareness of words in different contexts.

> ***"I don't ask them to dictate just for the Arabic language, but I ask them to dictate in maths, religion, social studies and science, why?… I need to make the child aware and alert so that every***

> ***word they come across, wherever they encounter it, they try to read it, master it, and memorise its form.*** *"* Usual practice focus group participant." Usual practice focus group participant

Lastly, the importance of tasks being achievable for students was discussed in the usual practice focus groups. Teachers in this group highlighted how tailoring activities to different abilities encouraged students to remain motivated in learning literacy.

> ***"When I assign reading activities, I don't give the same level to everyone…For the struggling student, since they are already struggling, I don't want them to feel their weakness; I want them to feel they are achieving something... I give them skills that match their level."*** Usual practice focus group participant

Homework

Some resource room teachers in the C/U model focus group described how, in usual practice teaching, they would record themselves speaking and send this to students to enable practice at home. This was to assist with completion of homework, for instance when parents were not cooperative or supportive of their child. In general, it was agreed across intervention and usual practice focus groups that videos were a widely used method of helping to improve engagement with students and parents alike.

Changes in practice

Training and support

All LRF! resource room teachers received training and coaching. All eight of the teachers who completed the post-intervention survey reported attending the two-day LRF! training run by QRTA and receiving the maximum number of three coaching sessions. This was supported by attendance data supplied by QRTA. The teachers also reported attending the maximum number of bi-weekly online coaching meetings, between five and six. The QRTA team reported there being four main sessions, with an additional two; one which introduced the programme and the other which discussed the mechanisms for selecting the sample of students. This differed from QRTA's attendance data, which stated from the eight teachers involved, five attended three of the bi-weekly online coaching meetings, while the other three teachers attended just one.

Teaching literacy

During the delivery of LRF!, the practice book and pedagogical approaches (namely the *'I do, we do, you do'* approach) were used consistently by the majority of resource room teachers.

All resource room teachers who completed the post-intervention survey reported using the practice book in 100% of lessons.[42] Discussions within the resource room teacher focus group suggested this was because the book was useful in supporting with teaching reading.

> ***"What makes it easier for me is that the words are already prepared, the texts are ready, and they have proven their effectiveness."*** C/U model resource room teacher focus group participant

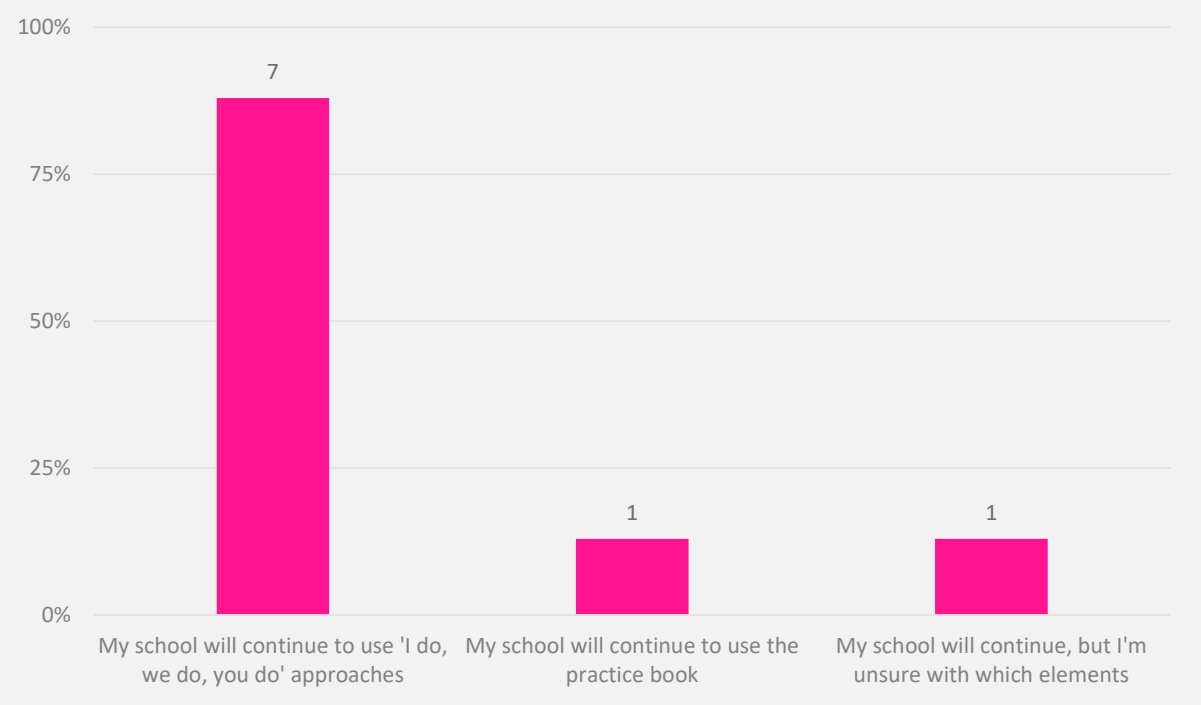
In terms of the pedagogical approaches, the focus group discussions with coaches and resource room teachers verified that the *'I do, we do, you do'* approach was being followed correctly, and that even

[42] The survey answer options indicated that 36 lessons equated to 100%, whereas the success criteria outlined 100% being 42 lessons. This discrepancy means we cannot be certain whether teachers meant 36 lessons or 100% when they responded.

teachers who did not take part in the intervention began to use the technique due to its perceived usefulness. Similarly, one resource room teacher thought the techniques were so useful that they delivered LRF! sessions to students who were not struggling, or those who were not enrolled on the initiative.

Speaking to its long-term effectiveness, almost all (88%; 7/8) resource room teachers who completed the post-intervention survey said that the *'I do, we do, you do'* approach would continue to be used in their school after the programme had ended (Figure 4), while over half of classroom teachers (59%; 22/37) agreed they would use the approach post-intervention.

Figure 4: Use of LRF! skills and approaches after the programme has ended



Source: Resource room teacher post-intervention survey, n=8, multiple choice allowed

In contrast, the continued use of the practice book was much less certain. Only one resource room teacher selected in the post-intervention survey that their school would continue to use the practice book post-LRF! (Figure 4), and coaches reported during their focus group that teachers had reservations about the book, because of the length of sentences and the lack of pictures. Suggestions around changes that could be made to the practice book are explored under RQ12.

Use of the practice book

Resource room teachers delivering LRF! and classroom teachers observed high levels of motivation among students to take the practice book home (88%; 7/8 resource room teachers; 89%; 33/37 classroom teachers agreed in the post-intervention survey).

All resource room teachers who responded to the post-intervention survey said that students took the practice book home a few times a week, and that they assigned homework in the practice book in all lessons. Likewise, respondents (75%; 6/8) felt that students were confident using their practice book in the resource room and all felt they used it for more than half of lessons.

In line with reports from the usual practice focus group, coaches highlighted that students were more motivated to use the practice book at home when their parents were engaged, but that parental engagement was not a given. An alternative was that parents would do the homework themselves.

> ***"The difficulties were in implementing it [the practice book] at home. That's the only challenge. The problem was that not all parents were following up with it at home."*** Coaches focus group participant

Resource room teachers shared a similar sentiment, that students struggled to complete work at home without parental help.

> ***"But, of course, the parents are not cooperative. They only play the audio recording for their children and leave them on their own."*** C/U model resource room teacher focus group participant

In order to bolster engagement, teachers drew on techniques that they used during usual practice. For example, in some instances, teachers would set homework and ask parents to send videos of their child practising back to them.

> ***"I launched a thing called 'The Home Video' for the mothers' group. It was a home video competition - send me a video of the students' reading, and the best one would receive a prize and recognition and so on…I truly saw results from this…mothers began competing to show whose daughter was better."*** C/U model resource room teacher focus group participant

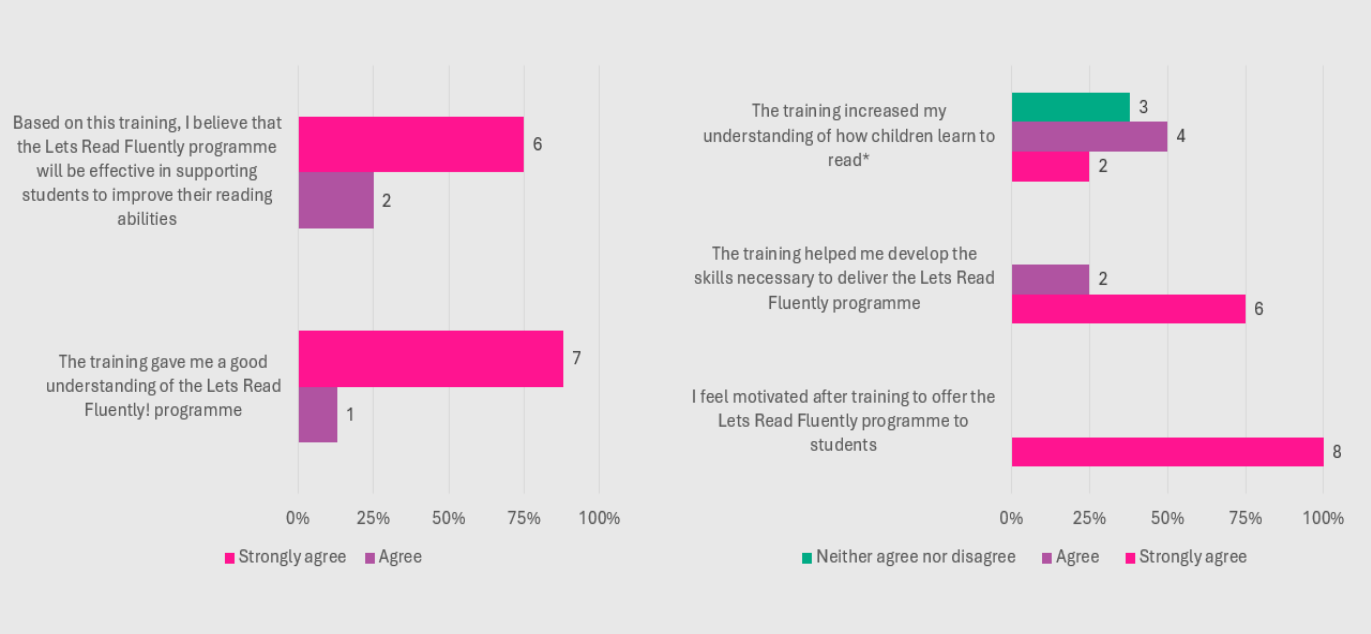
RQ2: How do teachers perceive the intervention and any changes that it has delivered? (IPE)

This section explores outcomes of the LRF! C/U model programme and training received on teachers and students, notably the impact on their understanding, motivation, and confidence in comparison to usual practice.

Perceived outcomes of the training and support

As portrayed in Figure 5 below, resource room teachers reported positive outcomes of the new training system. All resource room teachers agreed that the two-day training provided them with a good understanding of the LRF! C/U model. All also reported that it helped them to develop necessary skills, including an enhanced understanding of how children learn to read, and left them feeling motivated to deliver the programme. As a result of the training, all teachers expressed confidence in the potential effectiveness of the programme prior to delivering the intervention. Very similar findings were replicated when resource room teachers gave feedback on the coaching sessions, in the post-intervention survey

Figure 5: Resource room teacher perspectives of the training outcomes



Note. *1 respondent selected both 'Agree' and 'Neither agree nor disagree'.

Source: Resource room teacher post-training survey, n=8

As can be seen in Figure 5 the resource room teachers had more mixed views on whether the training increased their understanding of how children learn to read. This could be a reflection of the training, or the resource room teachers already feeling like they had a good level of understanding.

In contrast to the survey findings, it was reported during the coaches focus group that levels of resource room teacher motivation varied. Some appeared to enjoy using elements such as the practice book, whereas others found delivery of lessons boring and tiresome. This was particularly related to the repetitive element of the programme. In addition, some teachers had to repeat the same lesson multiple times a week when the number of students eligible for LRF! C/U model was too large for one group and therefore split into smaller groups, again contributing to their boredom with delivery. The theme of teacher (and student) boredom is explored later in this section (RQ2), in RQ5, and in RQ14.

During the coaches focus group, participants reflected on changes made to the training since the previous pilot, particularly its extension to two days. This additional time was seen as more suitable, specifically due to the second day teaching the application of LRF! C/U model and how it worked in practice, cementing the resource room teacher's knowledge. This is evidenced further in RQ8.

> ***"You guarantee that he will return to his school truly understanding the methodology correctly."***
> Coaches focus group participant

Perceived impact of the LRF! C/U model on students' confidence and motivation to practise reading

Both teachers and coaches reported improvements in students' confidence and motivation as a result of the intervention.

The majority of resource room teachers who completed the post-intervention survey (88%; 7/8) felt that students who had received the LRF! C/U model had increased confidence and motivation to practise reading. In the C/U model resource room teacher focus group, one view was that confidence increased alongside students' enjoyment of lessons. Interestingly, fewer classroom teachers (49%; 18/37) who answered the post-intervention survey reported increased motivation of students to practice reading. During the coaches focus group, it was suggested that the repetitive methods used in the LRF! C/U model were easier for students to understand than usual approaches, which in turn increased their levels of motivation to practice reading.

> ***"Their confidence in themselves has greatly increased. There was a shy student who didn't speak, but through implementing the methodology, we noticed that the student's confidence improved."*** Coaches focus group participant

Almost all classroom teachers (95%; 35/37) who answered the post-intervention survey felt that students were motivated to attend LRF! C/U model sessions. However, focus group discussions with coaches and C/U model resource room teachers portrayed a range of experiences for students. Some participants highlighted how students found lessons enjoyable and enjoyed practices such as chanting and reading stories. In contrast, the experience of other students seemed similar to that of some teachers, as they appeared bored during each lesson because of the repetitive reading and listening activities. This led to lower student motivation.

> ***"At the beginning, the students were enthusiastic, [...] but starting from the middle of the term onwards, I began to notice when I called them to the class, that their expressions showed they weren't interested anymore because they were simply bored. It was the same method over and over again."*** C/U model resource room teacher focus group participant

Perceived impact of LRF! C/U model on students (pre-) literacy level

Results from the post-intervention survey showed that most classroom teachers (86%; 32/37) believed students had improved their literacy levels. This positive view was not unanimously shared; findings from the focus groups with coaches and resource room teachers delivering the intervention suggested that effects on students' pre-literacy level varied. Coaches highlighted in focus groups that some students could read letters and words, but others had seen little change and still required continued support.

> ***"There has been improvement in some students... student X has improved their reading compared to when they started, while student Y is still only reading letters and words. The issue of improvement is different from one student to another".*** Coaches focus group participant

On one hand, some resource room teachers suggested during their focus group that the intervention was successful because students could now distinguish letters and vowels of words. There was agreement that the intervention improved pre-literacy levels for those who had previous knowledge of letters, and that these students could now read for the entire lesson. Some resource room teachers attributed this success to the repetitive approaches used within the LRF! C/U programme, and that these approaches helped even the lowest performing students learn how to read.

> ***"With repetition and practice, and through persistence, I was surprised to see that even the weaker students reached the stage where they could distinguish the vowel marks."*** C/U model resource room teacher focus group participant

In contrast, the post-intervention survey showed that under half of resource room teachers (3/8; 38%) believed that students who had taken part saw an improvement in their pre-literacy level. Focus groups with coaches and resource room teachers also highlighted that the intervention was not effective in improving reading ability for some students. Resource room teachers voiced skepticism around the programmes impact due to the exclusion of writing exercises and the absence of exams, which were seen to be a motivator of student learning.

> ***"The program also lacks assessment…there was only an evaluation at the end of the week. The student, naturally, if there is no exam, doesn't study."*** C/U model resource room teacher focus group participant

In addition, participants in the coaches and C/U model focus groups suggested that unless parents followed up with homework and supported their children, pre-literacy level would not change.

> ***"For the ones [students] whose parents weren't following up with them, they remained the same".*** C/U model resource room teacher focus group participant

RQ3: Is there evidence to support the revised logic model? (IE and IPE)

The two key outcomes identified in the revised logic model are: improved (pre-) literacy level of students involved in the LRF! C/U programme, and improved ability of resource room teachers to deliver effective reading sessions. Evidence for each of these outcomes is explored in turn below, using evidence from both the IPE and IE.

Improving (pre-) literacy level of students involved in the LRF! C/U programme

As outlined in the previous section, coaches and teachers had mixed perceptions of the impact of the LRF! C/U programme on students' pre-literacy level. This section presents the results from the sub-tests of the EGRA test to explore the impact of LRF! on students selected for the programme from baseline to endline testing.

Given its small sample size, this pilot evaluation has limited statistical power, resulting in a greater likelihood of Type I and II errors, and leading to highly imprecise estimates.[43] Consequently, the findings from the impact evaluation should be approached with great caution and no conclusions regarding the intervention's impact on any outcome measure should be drawn at this stage.

Primary outcome analysis

The primary outcome measure of this evaluation is ORF which is measured as the number of correct words per minute. The primary outcome analysis included 161 students. For students in the analysis, the baseline ORF score[44] has a mean of 3.09 and standard deviation of 5.83. The distribution of the baseline ORF score is illustrated in Figure 6. The endline ORF score has a mean of 8.24 and a standard deviation of 11.63. The distribution of the endline ORF score is illustrated in Figure 7.

[43] This is explained in detail in the *Impact Evaluation Analysis* section above.

[44] As part of EGRA testing, students 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.

Figure 6 shows that ORF scores are not normally distributed at baseline, with most students obtaining a total score of 0 on this sub-task. Therefore, there is a large floor effect for this measure.[45] Both Figures 6 and 7 show a right skewed distribution indicating that many students have persistently low scores (between 0 - 5). The figures show a slight improvement at endline, as there is a reduction in low scores and an increase in moderate or higher scores, as indicated by the increase in the mean (from 3.09 at baseline to 8.24 at endline) and the broader spread of scores (ranging 0 - 25 at baseline and 0 - 50 at endline). The endline distribution shows a more gradual decline in frequency after initial peak, indicating a small reduction in the floor effect observed at baseline.

Figure 6: Baseline ORF Score

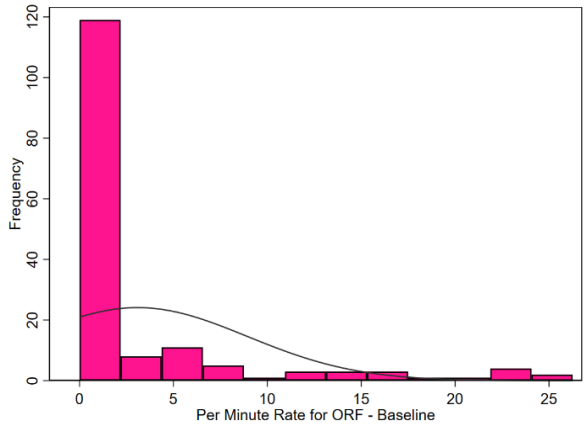
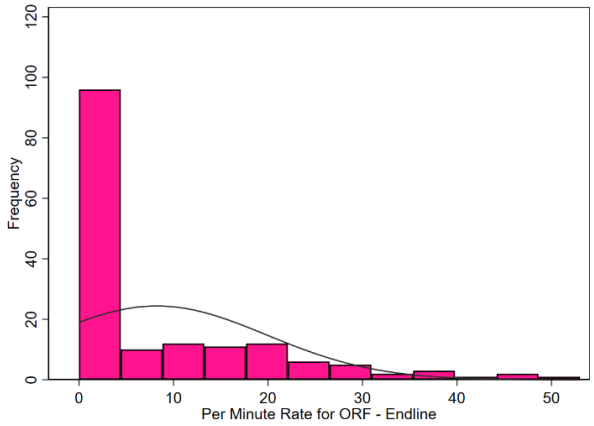


Figure 7: Endline ORF Score



The endline ORF score is correlated to the baseline ORF score with r = 0.69, suggesting a moderate-to-strong positive correlation. The unadjusted ORF mean score difference between the C/U model group and the usual practice group is 2.98. In the multilevel model that accounted for the baseline ORF, the stratification variable and the clustering of students, the adjusted difference in ORF means between the intervention and the usual practice group is 2.82 with a 95% bootstrapped CI of 0.55 and 5.60 (Table 11). The 95% bootstrapped CI excludes zero and includes the MRMD of 2.24, suggesting that there is indicative evidence of the impact of the C/U model on oral reading fluency. The standardised effect size associated with the adjusted difference in means is 0.25 (95% CI: -0.07, 0.56).

The post-intervention ICCs were estimated from endline ORF data from both the C/U and usual practice schools. We used the unadjusted model when calculating the ICCs.[46] Based on ORF data, the ICC within schools is 0.026 with a 95% confidence interval of 0.00093 and 0.44. The ICC in classes within schools is 0.026 with a 95% confidence interval of 0.00093 and 0.44.

Table 11: Primary outcome analyses (ORF)

			C/U model		Usual practice		
Outcome	**Unadjusted diff. in means**	**Adjusted diff. in means (95% CI)**	**n (missing)**	**Variance of outcome**	**n (missing)**	**Variance of outcome**	**Pooled var.**

[45] Floor effect occurs when a significant number of respondents score at the lower end of the measurement scale.

[46] The ICC measures the degree of similarity between units within the same cluster such as students within the same classroom or school. Units within the same cluster may exhibit similarities due to being exposed to similar environmental characteristics. A higher ICC indicates a stronger correlation of responses within a cluster, meaning that individuals within the same cluster are more similar in their responses.

Oral Reading Fluency	3.12	2.82 (0.55, 5.60)	96 (14)	147.23	65 (7)	95.21	126.29

Table 12: Effect size estimation - primary outcome analyses (ORF)

	Unadjusted means				Effect size	
	C/U model		Usual practice			
Outcome	**n (missing)**	**Mean (95% CI)**	**n (missing)**	**Mean (95% CI)**	**Total n**	**Hedges' g (95% CI)**
Oral Reading Fluency	96 (14)	9.50 (7.25, 12.1)	65 (7)	6.38 (4.07, 8.76)	161	0.25 (-0.07, 0.56)

Secondary outcome analysis

The protocol set out six secondary outcome measures: Arabic Literary Attainment (ALA), Letter Sound Identification (LSI), Syllable Identification (SI), Word Recoding (WR), Reading Comprehension (RC) and Listening Comprehension (LC). The results of the secondary outcome analyses are presented below.

Arabic Literacy Attainment (ALA)

The first secondary outcome measure in this trial is ALA. The baseline ALA score has a mean of -0.13 and standard deviation of 0.60. The endline ALA score has a mean of -0.10 and standard deviation of 0.67. The distribution of this measure is provided in Figure 8 and Figure 9 below. This measure is right skewed at both baseline and endline.

The unadjusted mean difference between the C/U model group and the usual practice group is 0.28 (Table 13). The adjusted mean difference between the C/U model group and the usual practice group is 0.15 with a 95% bootstrapped CI of -0.06 and 0.31. The 95% bootstrapped CI includes the MRMD of 0.13, but it also includes zero, suggesting that there may be no difference between the C/U and usual practice groups. Therefore, our analysis suggests no indicative evidence of impact of the C/U model on the ALA measure.

Figure 8: Baseline ALA score

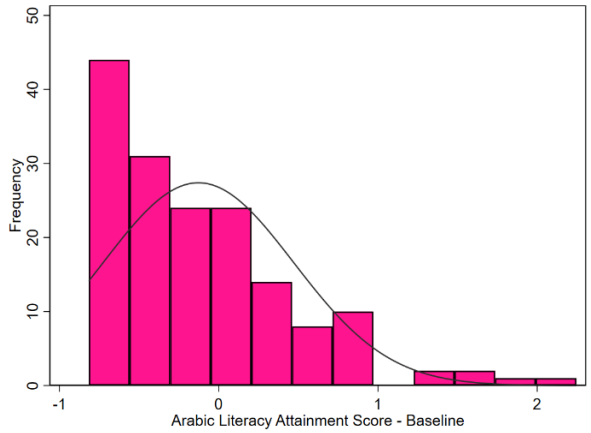


Figure 9: Endline ALA score

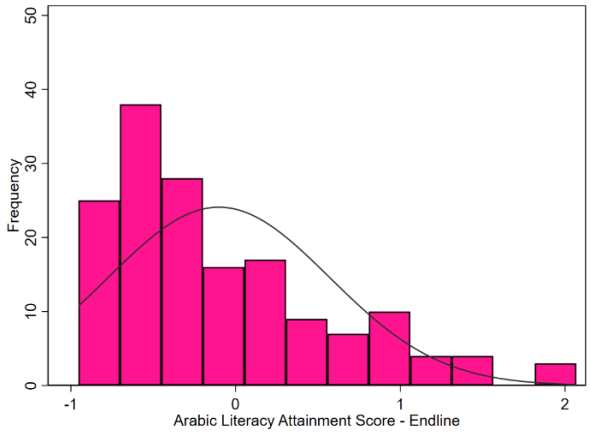


Table 13: Secondary outcome analyses for Catch-Up model (ALA)

			C/U model		Usual practice		
Outcome	**Unadjusted diff. in means**	**Adjusted diff. in means (95% CI)**	**n (missing)**	**Var. of outcome**	**n (missing)**	**Var. of outcome**	**Pooled var.**
Arabic Literacy Attainment (ALA)	0.28	0.15 (-0.06, 0.31)	96 (14)	0.5	65 (7)	0.22	0.39

Table 14: Effect size estimation – secondary outcome analyses for Catch-Up model (ALA)

	Unadjusted means				Effect size	
	C/U model		Usual practice			
Outcome	**n (missing)**	**Mean (95% CI)**	**n (missing)**	**Mean (95% CI)**	**Total n**	**Hedges' g (95% CI)**
Arabic Literacy Attainment (ALA)	96 (14)	0.01 (-0.12, 0.16)	65 (7)	-0.27 (-0.39, -0.16)	161	0.24 (-0.08, 0.55)

Letter Sound Identification (LSI)

The second secondary outcome measure is LSI, which is measured as the number of correct letter sounds read per minute. The baseline LSI score has a mean of 21.50 and standard deviation of 19.90. The endline LSI score has a mean of 30.13 and standard deviation of 26.10. The distribution of this measure is provided in Figure 10 and 11, below. The mean LSI scores increased from baseline to endline, suggesting some improvement in LSI scores. However, this measure is not normally distributed, with around half of the students obtaining scores close to 0. There is, therefore, a small to moderate floor effect for this measure.

The unadjusted mean difference between the C/U model group and the usual practice group is 9.89 (Table 15). The adjusted mean difference between the C/U model group and the usual practice group is 3.52 with a 95% bootstrapped CI of -2.45 and 11.16. The 95% bootstrapped CI includes 0 and the MRMD of 4.88, suggesting no indicative evidence of impact of the C/U model on letter sound identification. The standardized effect size associated with this difference is 0.14 (95% CI: -0.17, 0.46).

Figure 10: Baseline LSI score

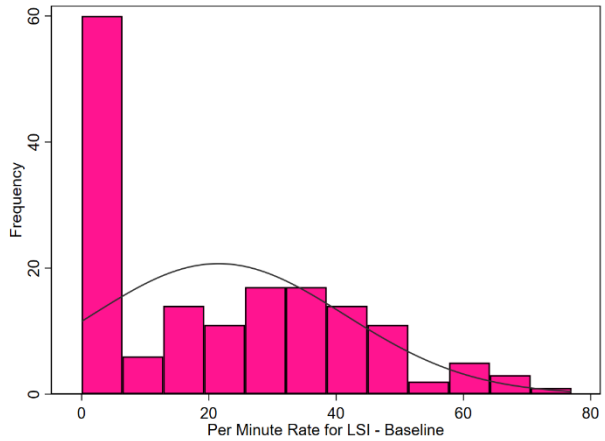


Figure 11: Endline LSI score

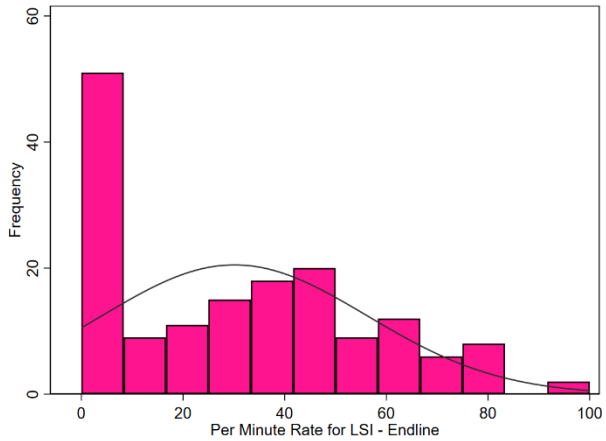


Table 15: Secondary outcome analyses for Catch-Up model (LSI)

			C/U model		Usual practice		
Outcome	**Unadjusted diff. in means**	**Adjusted diff. in means (95% CI)**	**n (missing)**	**Variance of outcome**	**n (missing)**	**Var. of outcome**	**Pooled var.**
Letter Sound Identification (LSI)	9.89	3.52 (-2.45, 11.16)	96 (14)	653.32	65(7)	509.23	595.32

Table 16: Effect size estimation - secondary outcome analyses for Catch-Up model (LSI)

	Unadjusted means				Effect size	
	C/U model		Usual practice			
Outcome	**n (missing)**	**Mean (95% CI)**	**n (missing)**	**Mean (95% CI)**	**Total n**	**Hedges' g (95% CI)**
Letter Sound Identification (LSI)	96 (14)	34.13 (29.45, 39.28)	65 (7)	24.23 (18.58, 29.37)	161	0.14 (-0.17, 0.46)

Syllable Identification (SI)

The third secondary outcome measure is SI, which is measured as the number of correct syllables per minute. The baseline SI score has a mean of 8.14 and standard deviation of 10.69. The endline SI score has a mean of 13.78 and standard deviation of 14.16. The distribution of the baseline and endline SI scores are illustrated in Figure 12 and 13, respectively. This measure is not normally distributed, with the majority of students obtaining a total score of 0 on this sub-task. There is, therefore, a moderate to large floor effect for this measure.

The unadjusted mean difference between the C/U model group and the usual practice group is 8.05 (Table 17). The adjusted mean difference between the C/U model group and the usual practice group is 5.60 with

a 95% bootstrapped CI of 1.95 and 8.96. The 95% bootstrapped CI both excludes 0 and includes the MRMD of 2.60, suggesting that there is indicative evidence of impact of the C/U model on syllable identification. The standardised effect size associated with this difference is 0.42 (95% CI: 0.10, 0.74).

Figure 12: Baseline SI score

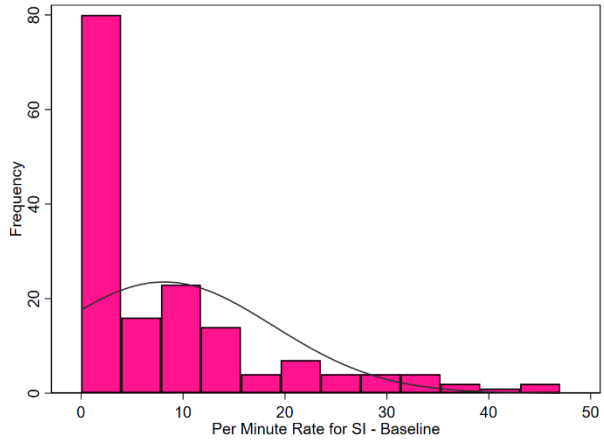


Figure 13: Endline SI score

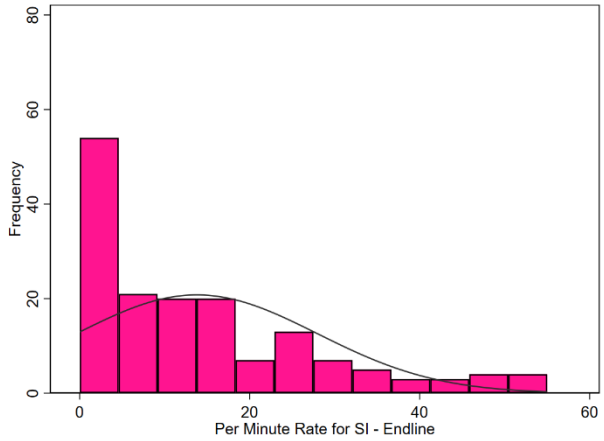


Table 17: Secondary outcome analyses for Catch-Up model (SI)

			C/U model		Usual practice		
Outcome	**Unadjusted diff. in means**	**Adjusted diff. in means (95% CI)**	**n (missing)**	**Var. of outcome**	**n (missing)**	**Var. of outcome**	**Pooled var.**
Syllable Identification (SI)	8.05	5.60 (1.95, 8.96)	96 (14)	217.06	65(7)	110.84	174.31

Table 18: Effect size estimation - secondary outcome analyses for Catch-Up model (SI)

	Unadjusted means				Effect size	
	C/U model		Usual practice			
Outcome	**n (missing)**	**Mean (95% CI)**	**n (missing)**	**Mean (95% CI)**	**Total n**	**Hedges' g (95% CI)**
Syllable Identification (SI)	96 (14)	17.03 (14.32, 20.2)	65 (7)	8.98 (6.4, 11.6)	161	0.42 (0.1, 0.74)

Reading Comprehension (RC)

The fourth secondary outcome measure is RC, which is measured as the number of reading comprehension questions a student answered correctly.[47] The baseline RC score has a mean of 0.50 and a standard deviation of 0.97. The distribution of the baseline RC score is illustrated in Figure 14. The endline RC score has a mean of 0.89 and a standard deviation of 1.37. The distribution of the endline RC is illustrated in

[47] The measure takes value between 0 and 5.

Figure 15. This measure was not normally distributed, with the majority of students obtaining a total score of 0 on this sub-task. Therefore, there was a large floor effect for this measure.

The unadjusted mean difference between the C/U model group and the usual practice group is 0.44 (Table 19). The adjusted mean difference between the C/U model group and the usual practice group is 0.26 with a 95% bootstrapped CI of -0.25 and 0.66. The 95% bootstrapped CI includes 0 and the MRMD of 0.26, suggesting no indicative evidence of the impact of the C/U model on reading comprehension. The standardised effect size associated with this difference is 0.20 (95% CI: -0.11, 0.51).

Figure 14: Baseline RC score

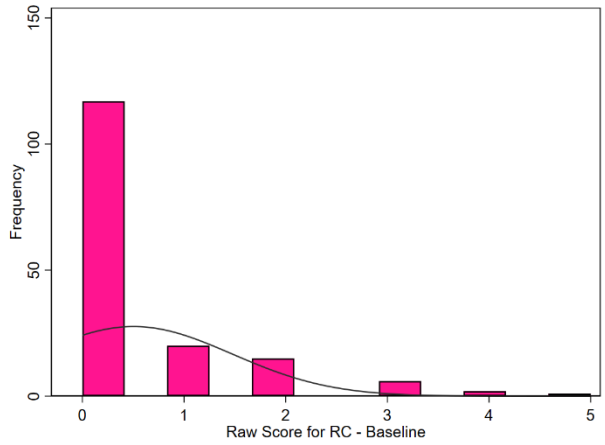


Figure 15: Endline RC score

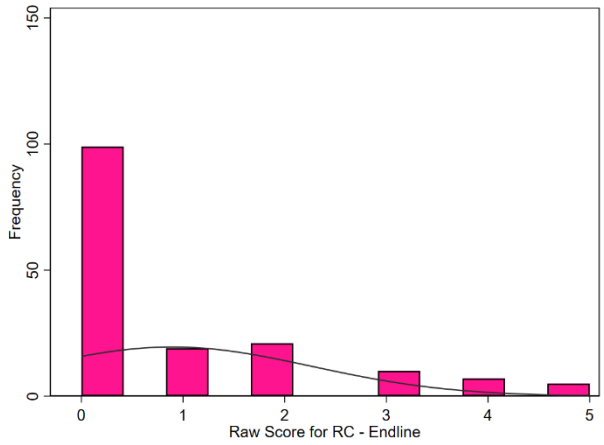


Table 19: Secondary outcome analyses for Catch-Up model (RC)

			C/U model		Usual practice		
Outcome	**Unadjusted diff. in means**	**Adjusted diff. in means (95% CI)**	**n (missing)**	**Var. of outcome**	**N (missing)**	**Var. of outcome**	**Pooled var.**
Reading Comprehension	0.44	0.26 (-0.25, 0.66)	96 (16)	2.3	65 (7)	0.85	1.72

Table 20: Effect size estimation - secondary outcome analyses for Catch-Up model (RC)

	Unadjusted means				Effect size	
	C/U model		Usual practice			
Outcome	**n (missing)**	**Mean (95% CI)**	**n (missing)**	**Mean (95% CI)**	**Total n**	**Hedges' g (95% CI)**
Reading Comprehension	96 (16)	1.07 (0.81, 1.4)	65 (7)	0.63 (0.38, 0.83)	161	0.20 (-0.11, 0.51)

Word Decoding (WD)

The fifth secondary outcome measure is WD, which is measured as the number of one- and two-syllable nonwords students read correctly per minute. The baseline WD score has a mean of 2.87 and a standard deviation of 4.14. The distribution of the baseline WD is illustrated in Figure 16. The endline WD score has a

mean of 4.65 and a standard deviation of 5.88. The distribution of the endline WD is illustrated in Figure 17. This measure is not normally distributed, with a large proportion of students obtaining a total score of 0 on this sub-task. Therefore, there was a large floor effect for this measure.

The unadjusted mean difference between the C/U model group and the usual practice group is 2.31 (Table 21). The adjusted mean difference between the C/U model group and the usual practice group is calculated as 1.76 with a 95% bootstrapped CI of 0.16 and 3.01. The 95% bootstrapped CI both excludes 0 and includes the MRMD of 1.10, suggesting that there is indicative evidence of the impact of the C/U model on word decoding. The standardised effects size associated with this difference is 0.32 (95% CI: 0.0002, 0.63).

Figure 16: Baseline WD score

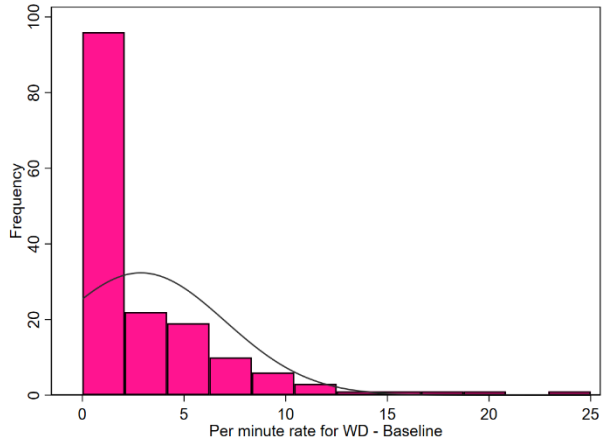


Figure 17: Endline WD score

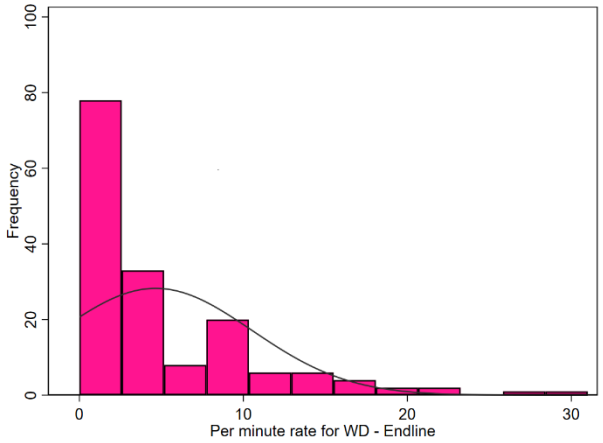


Table 21: Secondary outcome analyses for Catch-Up model (WD)

			C/U model		Usual practice		
Outcome	**Unadjusted diff. in means**	**Adjusted diff. in means (95% CI)**	**n (missing)**	**Var. of outcome**	**n (missing)**	**Var. of outcome**	**Pooled var.**
Word Decoding	2.31	1.76 (0.16, 3.01)	96 (16)	38.52	65 (7)	17.88	30.21

Table 22: Effect size estimation - secondary outcome analyses for Catch-Up model (WD)

	Unadjusted means				Effect size	
	C/U model		Usual practice			
Outcome	**n (missing)**	**Mean (95% CI)**	**n (missing)**	**Mean (95% CI)**	**Total n**	**Hedges' g (95% CI)**
Word Decoding	96 (16)	5.58 (4.44, 6.89)	65 (7)	3.28 (2.31, 4.32)	161	0.32 (0.0002, 0.63)

Listening Comprehension (LC)

The last secondary outcome measure is LC, which is measured as the number of listening comprehension questions a student answered correctly.[48] The baseline LC has a mean of 1.09 and a standard deviation of 1.19. The distribution of the baseline LC is illustrated in Figure 18. The endline LC score has a mean of 1.55 and a standard deviation of 1.21. The distribution of the endline LC is illustrated in Figure 19. This measure is right skewed and thus, is not normally distributed.

The unadjusted mean difference between the C/U model group and the usual practice group is 0.01 (Table 23). The adjusted mean difference between the C/U model group and the usual practice group is calculated as 0.08 with a 95% bootstrapped CI of -0.25 and 0.6. The 95% bootstrapped CI includes both 0 and the MRMD of 0.24, suggesting no indicative evidence of impact of the C/U model on listening comprehension. The standardised effect size associated with this difference is 0.07 (95% CI: -0.25, 0.38).

Figure 18: Baseline LC score

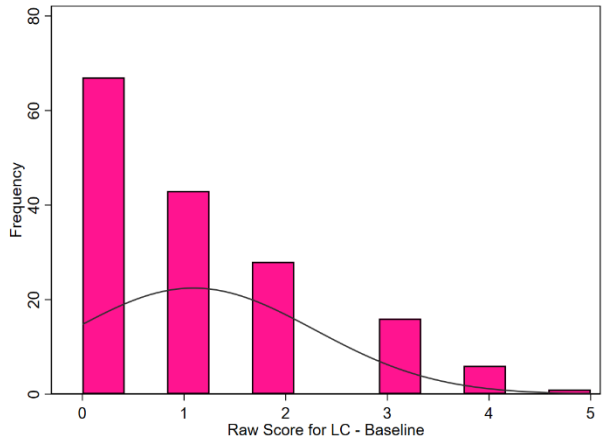


Figure 19: Endline LC score

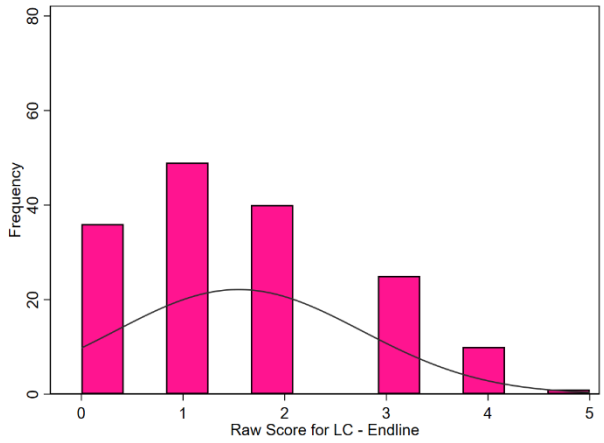


Table 23: Secondary outcome analyses for Catch-up model (LC)

			C/U model		Usual practice		
Outcome	**Unadjusted diff. in means**	**Adjusted diff. in means (95% CI)**	**n (missing)**	**Var. of outcome**	**n (missing)**	**Var. of outcome**	**Pooled var.**
Listening Comprehension	0.01	0.08 (-0.25, 0.6)	96 (16)	1.49	65 (7)	1.35	1.43

Table 24: Effect size estimation - secondary outcome analyses for Catch-Up model (LC)

	Unadjusted means				Effect size	
	C/U model		Usual practice			
Outcome	**n (missing)**	**Mean (95% CI)**	**n (missing)**	**Mean (95% CI)**	**Total n**	**Hedges' g (95% CI)**
Listening Comprehension	96 (16)	1.55	65 (7)	1.54	161	0.07 (-0.25, 0.38)

[48] The measure takes value between 0 and 5.

		(1.29, 1.78)		(1.28, 1.82)		

Summary of IE findings

Due to its sample size, this pilot evaluation has small statistical power, meaning that it has higher likelihood of Type I and II errors and estimates are highly imprecise. Therefore, the findings from the IE should be treated with extreme caution and should not be over-interpreted at this point. Based on the primary outcome analysis, there is indicative evidence for a small effect of the LRF! C/U model on oral reading fluency. Indicative evidence of effects on syllable identification and word decoding are also presented in the secondary outcome analysis, but there is no indicative evidence of effects on Arabic literacy attainment, letter-sound identification, reading comprehension and listening comprehension.

Furthermore, we undertook additional descriptive analysis for the LRF! pre-literacy items. Our findings on the pre-literacy items indicate that students in the C/U model group, on average, name more food items and animals than students in the usual practice group. Similarly, students in the C/U model group, on average, name more letters and read more words at endline than students in the usual practice group. Please see the Appendix 1 for more information on this additional analysis.

Delivering effective reading sessions to students 'in need' of additional literacy support

This section explores perceived changes in resource room teachers' ability to deliver effective reading sessions to those 'in need' of additional support.

Usual practice for delivering reading sessions to students 'in need' of additional support varied hugely. Adjustments could range from teachers creating specific educational plans for students, to support remaining largely unstructured. Some teachers would provide support by offering struggling students more attention than those with a higher reading ability. Based on findings from the previous pilot, the LRF! C/U model programme was updated by increasing the frequency of sessions for these learners and altering the practice book. Coaches reported these changes made the programme better suited to students in need of more support.

> ***"We feel that its application with the struggling students has taken a better turn after modifications in the book, after modifications in the routine [the LRF! programme]."*** Coaches focus group participant

Only a minority (25%; 2/8) of resource room teachers who completed the post-intervention survey reported improved ability to provide appropriate literacy support to students. When classroom teachers were asked the same questions about whether resource room teachers' ability had improved, the results opposed this. The majority (97%; 36/37) felt resource room teachers were better able to deliver effective reading sessions to those 'in need' of additional support. A potential reason for the difference in opinion could be that resource room teachers already rated their ability to provide appropriate literacy support highly, and therefore did not observe an improvement.

Coaches also reported improved ability of the resource room teacher to provide appropriate literacy support to students, noting that the delivery of content by the resource room teacher was *"**perfect**"*. However, they did voice concern over the flexibility needed when supporting students. Opinions voiced during the coaches focus group indicated that often one-to-one support was needed for struggling students. Although the 'You do' stage of the approach allows some one-to-one support (see Delivery of the

LRF! intervention, p.9) it was not perceived by everyone to be possible under the LRF! C/U model methodology.

> ***"Sometimes, with struggling children, you need to provide one-on-one support, […] This is not possible with the 'Let's Read Fluently' strategies, where the teacher provides the prompt, followed by repetitive reading, and then independent reading**."* *Coaches focus group participant*

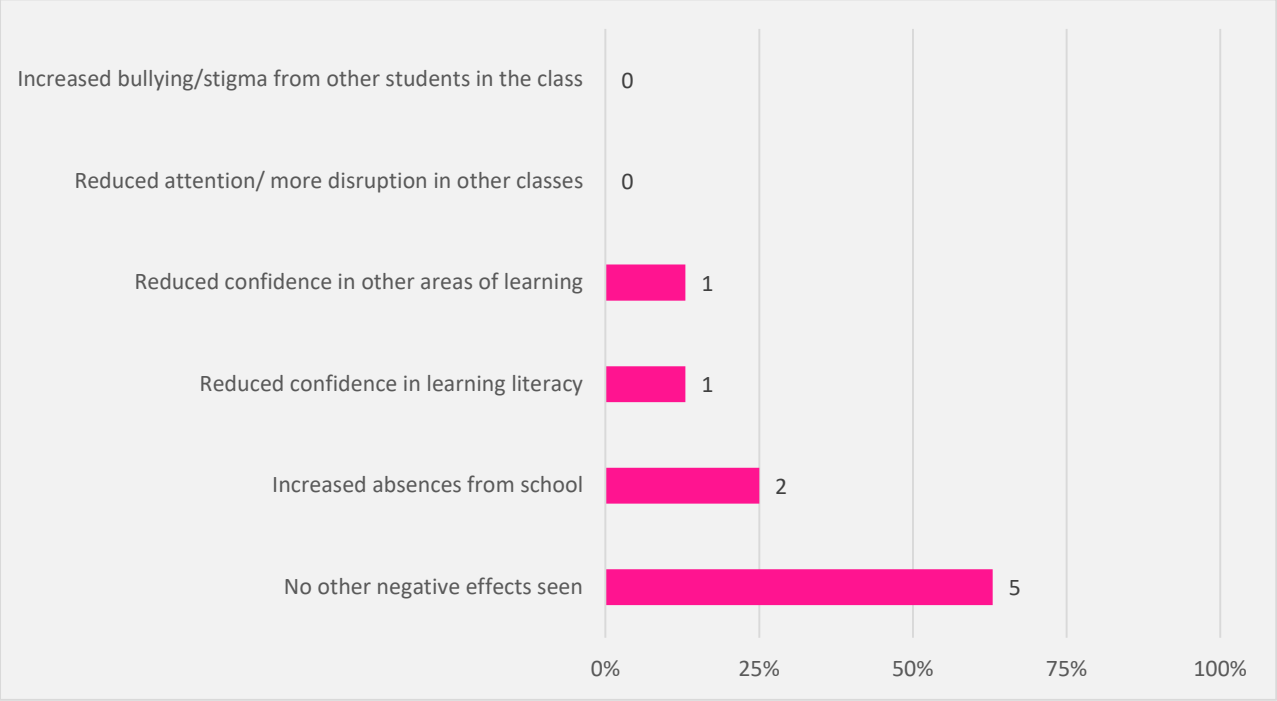
RQ4: Is there any evidence of unintended consequences (negative or positive) as a result of the implementation of the LRF! C/U model? (IPE)

This section presents any perceived unintended consequences as a result of the implementation of the LRF! C/U model on students participating in the programme, other students who were not eligible, teachers, and parents.

Impact on students participating in the programme

Results from the post-intervention survey show that the majority of resource room (63%; 5/8) and classroom teachers (84%; 31/37) felt that there were no negative effects on students who had taken part in the programme, and that typically, unanticipated consequences were positive. As displayed in Figure 20, a very small number of negative effects on students were reported.

Figure 20: Perceived negative effects on students who have taken part in LRF!

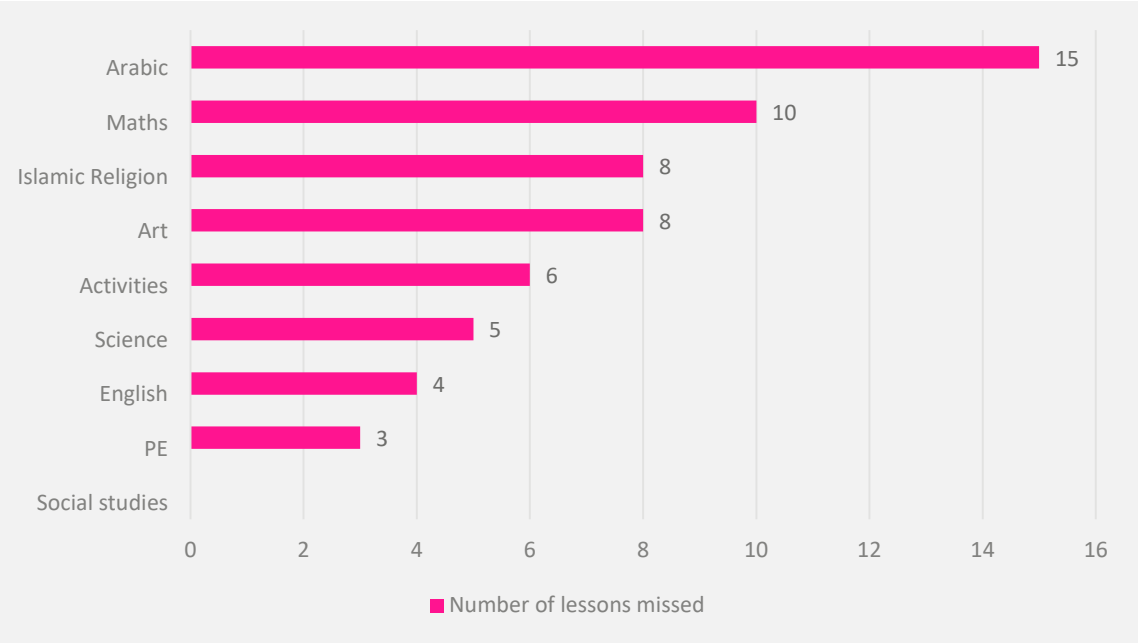


Source: Resource room teacher post-intervention survey, n=8, multiple selections allowed

In the post-intervention survey, all resource room teachers delivering the LRF! C/U model reported observing student development in other academic capabilities as a result of the programme. From a list of other potential positive consequences of LRF!, over half (63%; 5/8) selected increased confidence in other classes, while half (50%; 4/8) also selected improved attention or behaviour in other classes for students participating in LRF!.

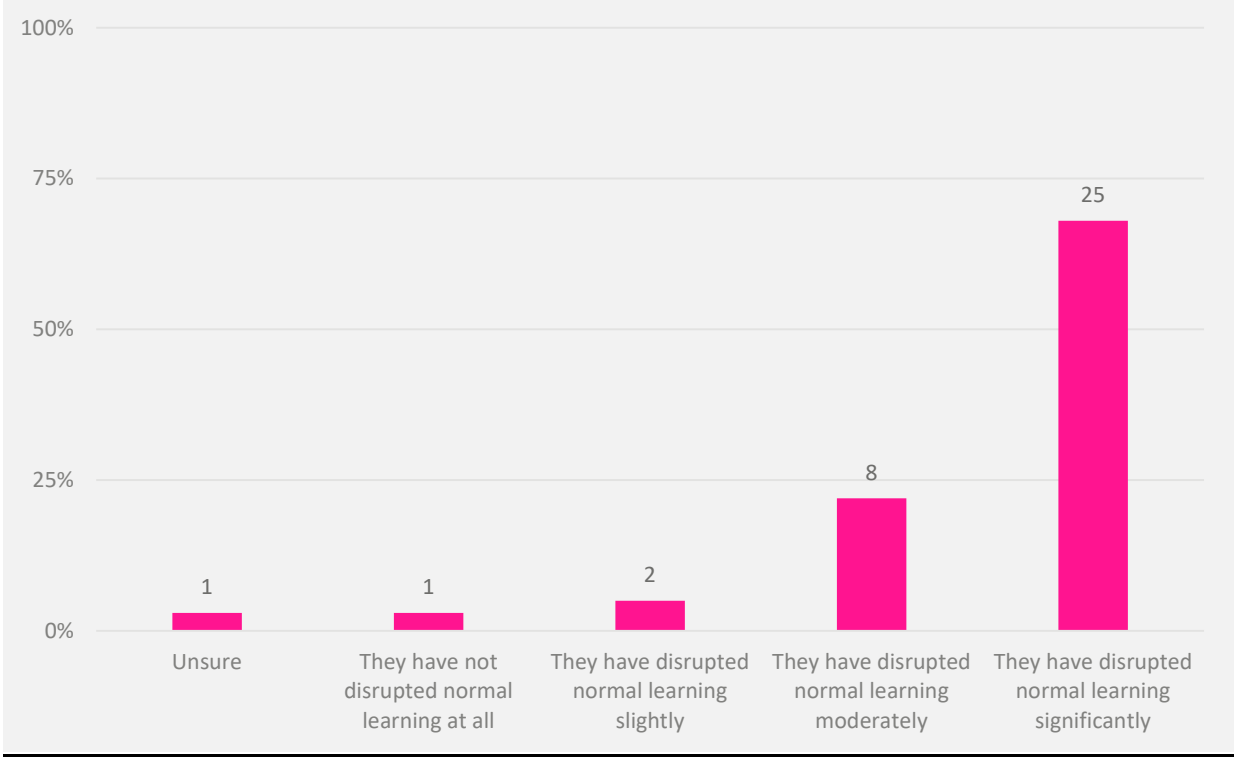
An additional point to consider was that attending LRF! C/U model sessions reduced student attendance in other lessons, which was not intended in the programme design. In the post-intervention survey, classroom teachers were asked which lessons students missed to attend LRF! C/U model sessions. Participants gave a variety of responses, with Arabic and Maths being the most common answers (Figure 21). As the questionnaire did not investigate the number of classes missed in each subject, it is not possible to know if only one or two classes in a subject were missed over the semester or significantly more. However, teachers reported a high level of disruption to normal learning for students participating in LRF! (Figure 22).

Figure 21: Lessons students missed to attend LRF! C/U model sessions



Source: Classroom teacher post-intervention survey, n=37, multiple choice allowed

Figure 22: Extent to which LRF! C/U sessions disrupted normal learning



Source: Classroom teacher post-intervention survey, n=37

Impact of LRF! C/U model on other students in class

In the post-intervention survey, most resource room teachers (88%; 7/8) and classroom teachers (95%; 35/37) reported there being no negative consequences for the students not eligible for the LRF! C/U model programme as a result of implementation.

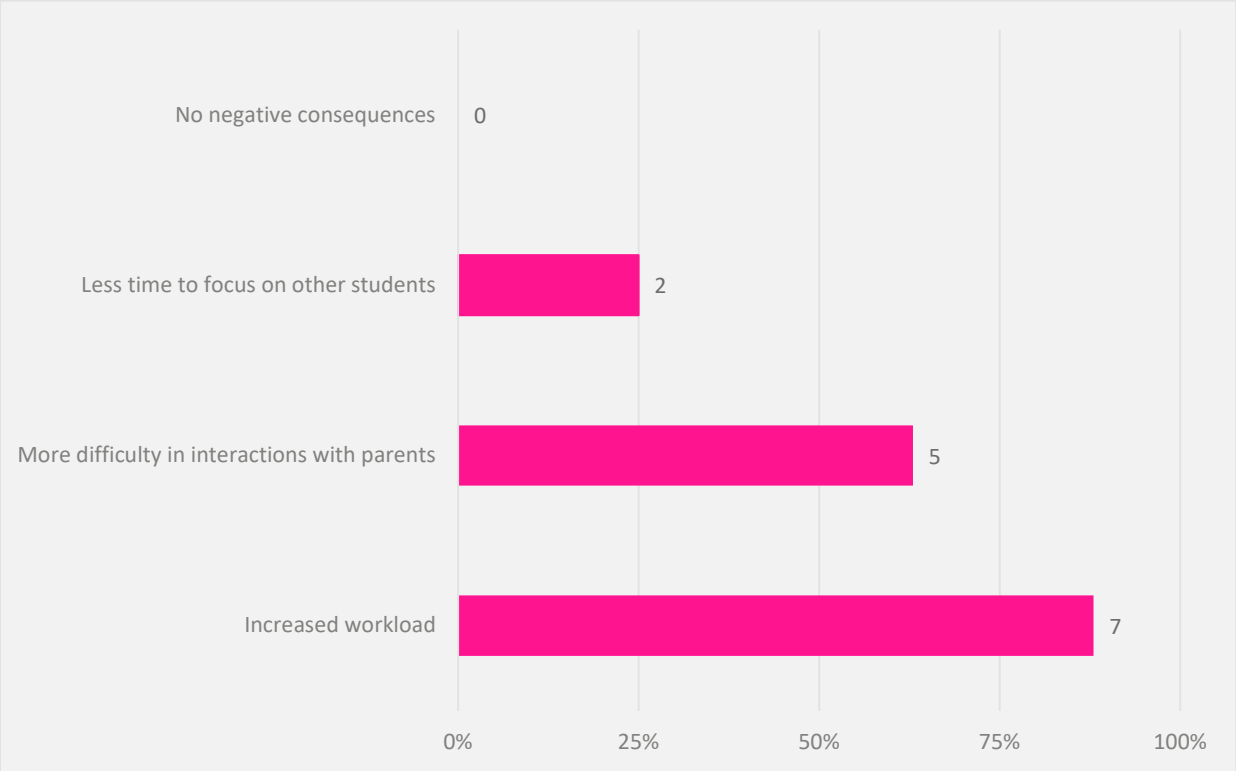
Impact on teachers

In the classroom teacher post-intervention survey, the most commonly selected positive impact was having a better understanding of how to identify students with reading difficulties (75%; 6/8), closely followed by having a better understanding of different approaches to teaching early literacy (63%; 5/8). Similarly, coaches reported during their focus group that resource room teachers were better able to assess students' improvement and keep records of progress as a result of the programme, which was not done previously.

> ***"They are learning how to conduct formative assessment, meaning, she constantly focuses on points of progress."*** Coaches focus group participant

Both classroom and resource room teachers reported experiencing negative consequences since the implementation of LRF! C/U model. For resource room teachers delivering the intervention, the most commonly reported consequence was having an increased workload (88%; 7/8). Full results are displayed in Figure 23.

Figure 23: Negative consequences on resource room teachers as a result of the LRF! C/U model programme



Source: Resource room teacher post-intervention survey, n=8, multiple choice allowed

Increased resource room teacher workload and the subsequent fatigue from this were also mentioned throughout different points in the focus groups with coaches and C/U model resource room teachers. Teachers spoke about the added effort needed in order to engage parents with the programme, and to encourage them to stay motivated.

> ***"Every morning we would say, 'Good morning, mothers!' We would energise them and wake them up by telling them, 'Today, we need to work on this specific story', so the follow-up was driven by motivation from the teacher."*** C/U model resource room teacher focus group participant

Similar to students, some resource room teachers found the lessons boring and tiresome, particularly as a result of the repetition involved in the programme.

> ***"I'm in a conflict. It's exhausting and draining, and it takes a lot of effort - especially the repetition part, which is very challenging. But, in the end I felt that the repetition was what truly brought results…Still, it was tiring".*** C/U model resource room teacher focus group participant

Most classroom teachers reported in the post-intervention survey that there were no negative consequences for them as a result of the programme (84%; 31/37). Those who did identify negative consequences reported having less time to focus on current literacy curriculum (8%; 3/37) and more difficulty in their interactions with parents (8%; 3/37), but in general, most felt there was nothing negative to report.

Impact on parents

Resource room and classroom teachers reported minimal unintended consequences on parents as a result of the LRF! C/U model. In the post-intervention survey, the majority of classroom (70%; 26/37) and resource room (63%; 5/8) teachers who completed their respective surveys felt there were no negative consequences on parents. Additionally, most classroom teachers felt that parents of students receiving

LRF!, and those not receiving LRF!, had a positive perception of the programme (84%; 31/37, and 73%; 27/37, respectively). This contrasts against discussions which took place in the resource room teacher focus group, where teachers explained how some parents associate the resource room with weak students, and did not support their children attending it.

Feasibility of intervention

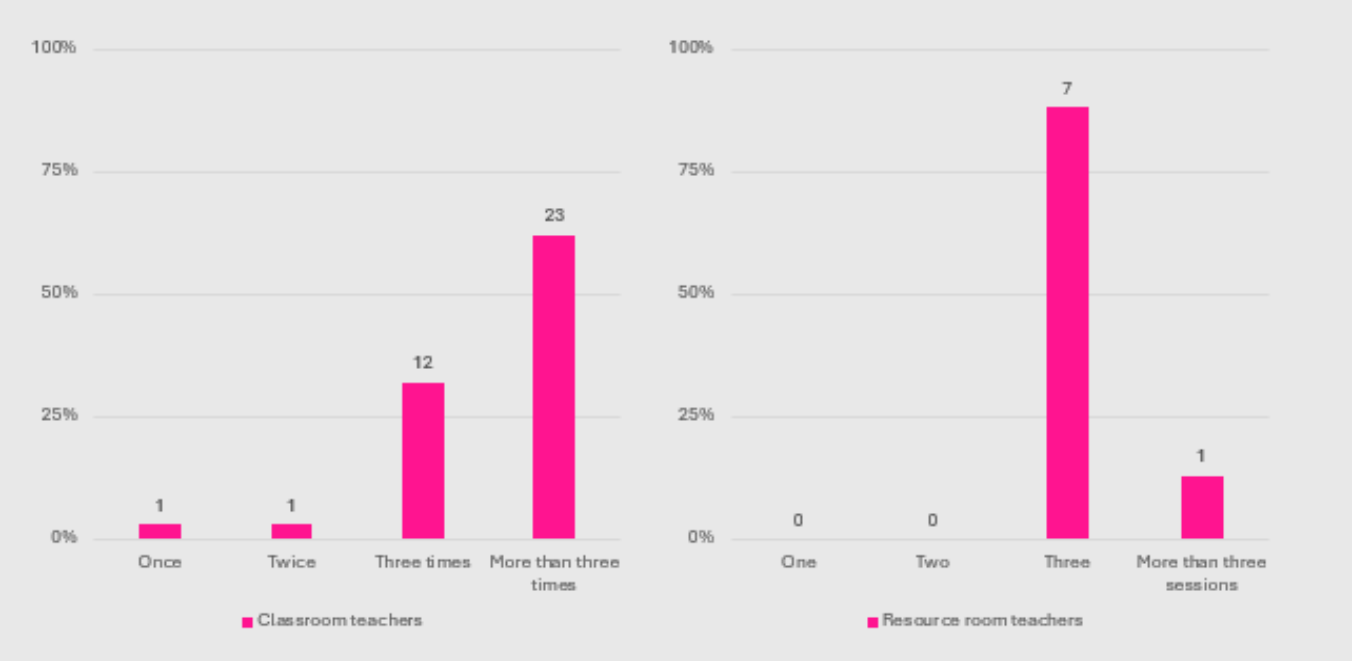
RQ5: Was the LRF! C/U model delivered as intended in terms of dosage, nature and quality? What modifications were made, with what implications? (IPE)

This section considers the dosage, nature and quality of the LRF! C/U model in comparison to usual practice, while identifying the adaptations which were made by teachers on the programme.

Dosage

The reported number of LRF! C/U model sessions delivered per week differed between resource room teachers and classroom teachers. While the majority of resource room teachers (88%; 7/8) reported delivering three sessions of LRF! C/U model per week, per group, as intended, the majority of classroom teachers (62%; 23/37) selected that students were sent to the resource room for their LRF! C/U model sessions more than three times per week. Results are displayed in Figure 24.

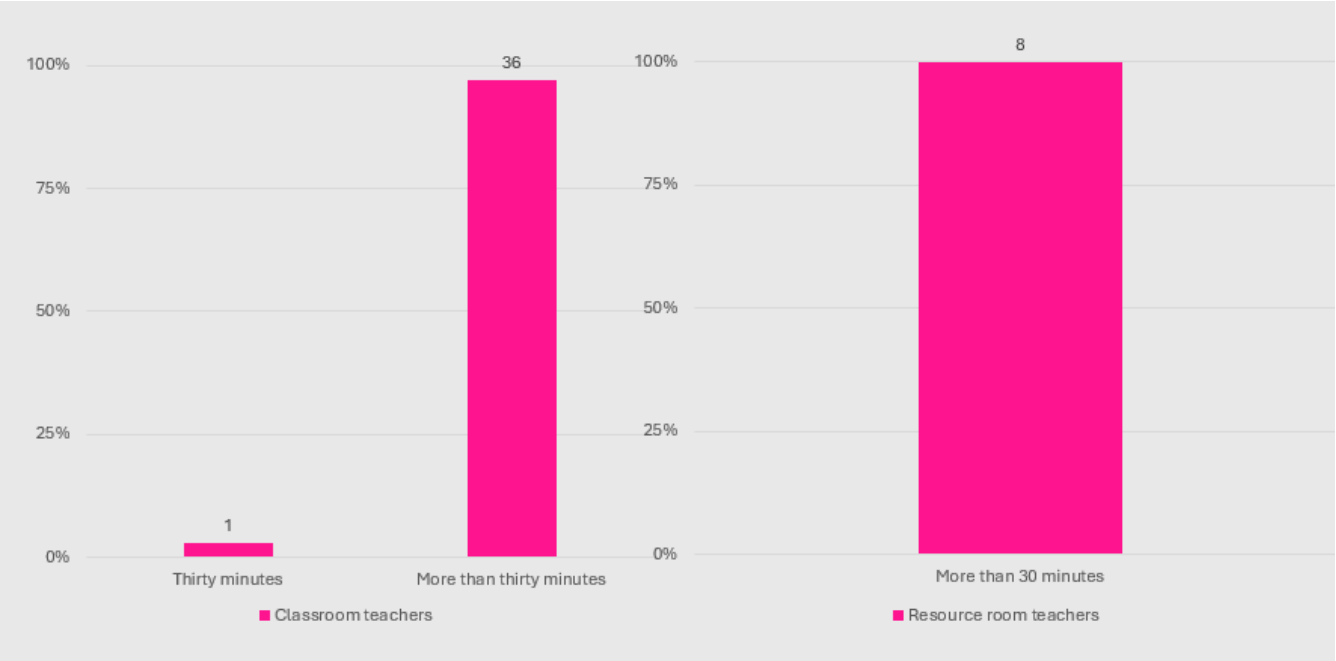
Figure 24: Number of times per week LRF! took place, reported by classroom and resource room teachers



Source: Classroom teacher post-intervention survey, n=37, and resource room teacher post-intervention survey, n=8

In terms of the length of LRF! C/U model sessions, almost all classroom teachers (97%; 36/37) and all resource room teachers reported that they were longer than planned. Results are displayed in Figure 25.

Figure 25: Duration of LRF! sessions, reported by classroom and resource room teachers



Source: Classroom teacher post-intervention survey, n=37, resource room teacher post-intervention survey, n=8

Resource room teachers mentioned during the focus group that the length of lessons was dependent on the students enrolled, and often materials took too long to get through, meaning that the intended activities could not be completed. This may also explain any discrepancy between classroom teacher and resource room teacher reports of the number of sessions delivered; it is possible that the intervention material for one lesson was spread over more sessions so that it could all be covered.

There were slight differences in the reported duration of the intervention; the majority (75%; 6/8) of resource room teachers reported delivering it for 13 weeks, while the majority of classroom teachers who completed the survey most commonly selected that it was delivered for 14 weeks or more (43%; 16/37), followed by 12 weeks (16%; 6/37)

The post-intervention survey results indicate that there was consistent use of the LRF! C/U model practice book and pedagogical approaches. All resource room teachers reported using the practice book in 100% of their lessons, and all used the 'I do, we do, you do' approach in their lessons, either "consistently" (75%; 6/8) or "most of the time" (25%; 2/8). More information on the use of the practice book and pedagogical approaches can be found under RQ1.

Quality

Student engagement in usual practice literacy lessons was facilitated through a combination of techniques, with several teachers mentioning the use of interactive elements in the usual practice focus group. Examples included clapping while reading to identify syllables, or the introduction of competitions, such as timed reading sessions. One teacher suggested that matching students of similar reading ability together facilitated engagement. Importantly, teachers identified that a tailored approach to teaching literacy to each student increased their engagement.

> ***"There is no…specific method for any specific person, every day new things emerge."*** Usual practice focus group participant

Opinions on whether the LRF! C/U model was engaging for students in need of additional support were mixed. The majority of resource room teachers (76%; 6/8) who completed the post-intervention survey agreed that students were engaged. However, coaches noted during their focus group that engagement was sometimes variable and that this may have been dependent on the region the intervention was delivered in and the teaching style used by the teacher.

> ***"In reality, there is a difference between one area and another. Sometimes this is due to the teacher's role, approach, and methods. … But, in general, we can say that the level of discipline in this area is high."*** Coaches focus group participant

As mentioned in RQ2 and RQ14, at times students were observed to find LRF! C/U model sessions boring due to the repetitive nature of the methodology.

> ***"The programme/curriculum was very rigid, boring and exhausting...I enjoyed the regular [literacy] lesson more…To be honest, boredom was evident in most of the students."*** C/U model resource room teacher focus group participant

In order to improve engagement, resource room teachers made modifications to the LRF! C/U model. These included offering material incentives such as sweets to motivate students, which students appreciated.

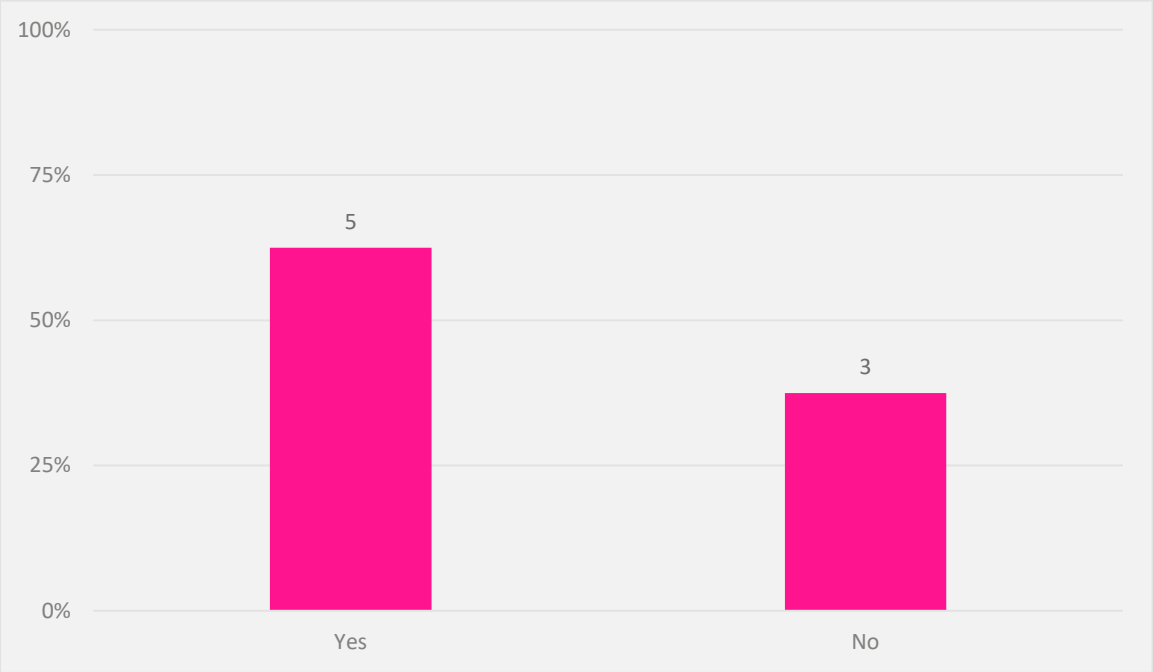
RQ6: What is the learning about the use of the PTI diagnostic tool? How successful is it at identifying the most appropriate students for the LRF! C/U model? (IPE)

In this section, we discuss whether the PTI diagnostic tool identified the appropriate students for the LRF! programme in comparison to previous practices of identifying students in need of additional literacy support. The consequences of the use of the PTI tool on students are also explored.

Level of accuracy of the PTI diagnostic tool

More than half of resource room teachers (63%; 5/8) believed that the correct students were identified for the programme. However, this meant that there were still over a third (38%; 3/8) who reported that the students identified for the programme were incorrect (Figure 26).

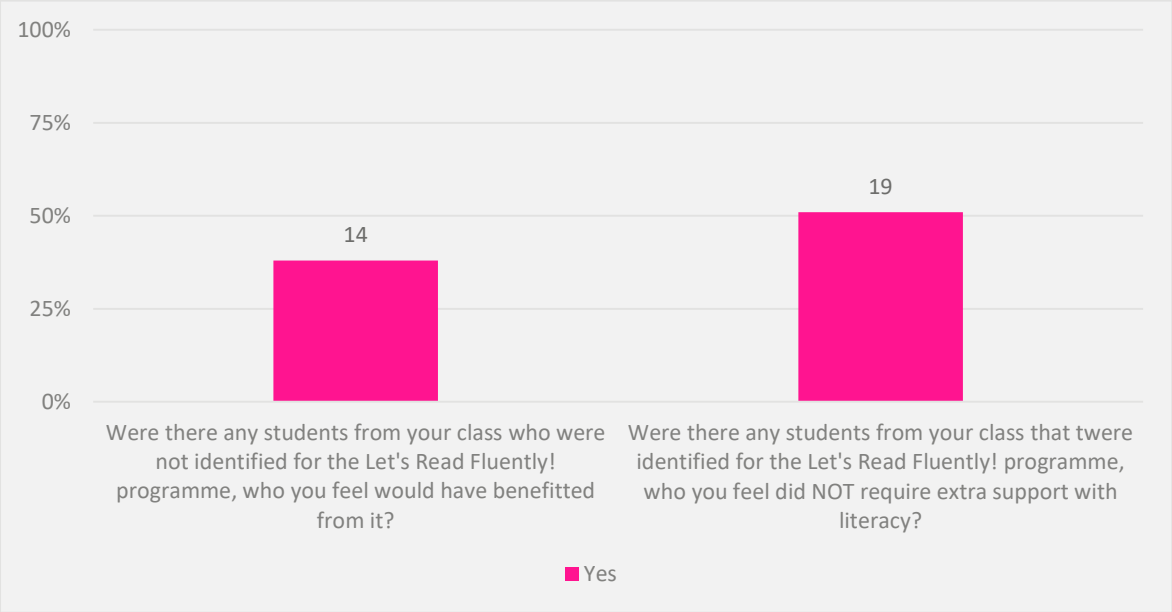
Figure 26: Resource room teacher views on whether the correct students were identified for LRF!



Source: Resource room teacher post-intervention survey, n=8

Over half of classroom teachers (51%; 19/37) reported there being students identified for LRF! who did not require extra support with literacy. These teachers were from a variety of schools, but did overlap with the three resource room teachers who reported the same issue (see Figure 26). In addition, over a third of classroom teachers (38%; 14/37) reported how students in their class would have benefited from the LRF! C/U model but were not identified (Figure 27).

Figure 27: Classroom teacher views on the students selected for LRF!



Source: Classroom teacher post-intervention survey, n=37

Results from the focus group also suggested problems identifying appropriate students, with resource room teachers highlighting the challenge of distinguishing between student skill sets in Grade 2, due to their overall lower ability.

According to the survey, the most common reasons amongst the three resource room teachers who thought the students selected were not suitable were either that they could not focus (66%; 2/3), or that they found the level of literacy required too hard (66%; 2/3). Resource room intervention teachers across the focus group highlighted the ineffectiveness of the intervention for students who could not read letters prior to starting the programme, or for students with learning difficulties. One respondent suggested screening tools were not effective at ruling out those with SEND.

In addition, coaches felt sampling was problematic, particularly as it did not align with the schools' recommendations of which students required support.

> ***"The names on the list that the school provided us with were different from the ones sorted by the diagnostic tool."*** Coaches focus group participant

As a result, coaches reported that some teachers did not use the students identified by the tools and relied on their own judgment, feeling this was more accurate.

Consequences of the PTI tool on students or parents

As discussed under RQ4, students who took part in the LRF! C/U model were required to miss other lessons in order to attend (Figure 21), and learning was viewed as significantly disrupted (Figure 22). Any further impact of this was not explored in the survey. However, in the resource room teacher focus group, teachers explained how in their view, not all of the students identified had difficulty reading, which meant they did not benefit from the programme. This is despite those with high EGRA ORF scores being excluded from the intervention. For those students who teachers viewed as not requiring the LRF! C/U model, but were still missing ordinary lessons, it is possible this impacted their learning in other subjects.
In terms of the consequences related to parents, it was widely reported throughout focus groups that a number of parents were hesitant about the programme because of its association with the resource room,

which was already stigmatised. Some parents seemed not to agree that their child would need support from the resource room.

> ***"Some mothers asked why it was [delivered in] a resource room because they associated the resource room with weak students".*** C/U model resource room teacher focus group participant

RQ7: What were the facilitators and barriers to engagement in the resource room teacher training and coaching sessions? (IPE)

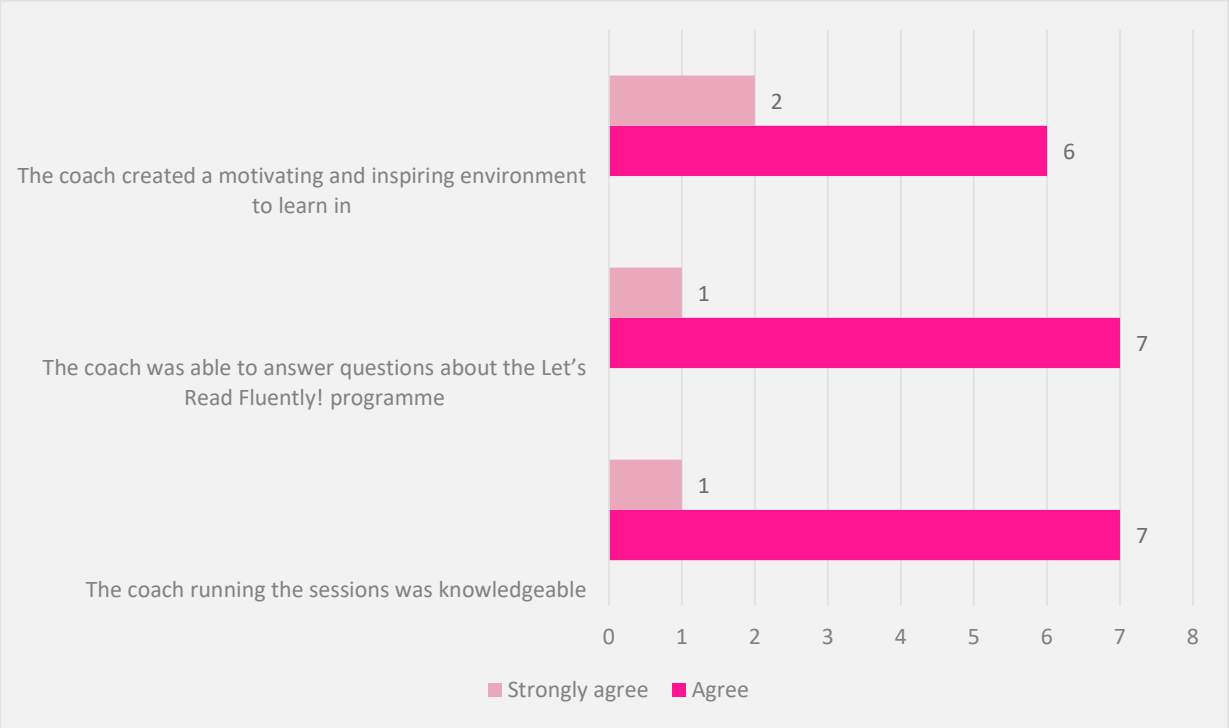
Facilitators to engagement

One of the main facilitators to engagement in the training and coaching were the trainers and coaches who delivered it. All resource room teachers who completed the post-training survey gave positive feedback about the trainers. All resource room teachers also reported that trainers were knowledgeable and able to explain key concepts well, in addition to being able to answer questions about the LRF! C/U model programme. All respondents indicated that the trainers presented materials clearly and logically and were able to create a motivating and inspiring training environment, in addition to managing the timing of the session well.

In the post-intervention survey, coaches were again seen by resource room teachers as knowledgeable, able to answer questions about the programme, and able to create a motivating and inspiring environment to learn in (Figure 28). It was explained during the coaches focus group how they would prepare materials for the resource room teachers, including evaluation tools, student attendance sheets and training plans, which explained how to carry out activities. Coaches reported that their coaching sessions and WhatsApp groups provided a space for teachers to share advice with one another, which would facilitate learning and further engagement.

> ***"If you come up with a new idea and manage to find or touch on a solution that sparks new motivation in the students, share it with the WhatsApp group."*** Coaches focus group participant

Figure 28: Resource room teacher perception of the LRF! C/U model coaches



Source: Resource room teacher post-intervention survey, n=8

In general, the training location, duration, and time of sessions were also viewed as a facilitator to resource room teacher engagement. The post-training survey feedback suggested a positive outlook among respondents, who all (100%; 8/8) felt the facility was of good quality, well equipped, and of an appropriate size for the group. This was echoed in the focus group.

> ***"The training duration was two days, and we didn't feel pressured. The duration of two days was excellent, wonderful. The place was great, the location was great, and the services they provided were really excellent."*** C/U model resource room teacher focus group participant

The number and duration of breaks during the training was also viewed as appropriate by all respondents and ensured continued attention.

Barriers to engagement

A minority of resource room teachers in the focus group expressed downsides to the training and coaching related to the location and timing. Some felt the training location was difficult to get to, due to being in a remote location and starting early in the morning.

> ***"The location might be a bit far from us… and we had to be there at 9am… it would have been better, if possible, to be there at 10am at the earliest."*** C/U model resource room teacher focus group participant

In addition, other resource room teachers reported how the coaching sessions in particular often took place after working hours, which made it challenging for some teachers to attend. A preference for them to take place during working hours was clear.

> ***"The main obstacle, honestly, was the timing, because we are female employees and we have housework when we return home…I felt that it would have been better if it were during official***

> ***working hours, not after leaving school".*** C/U model resource room teacher focus group participant

RQ8: To what extent do resource room teachers develop sufficient skills and confidence through the training and coaching? (IPE)

There was clear indication that resource room teachers developed skills and confidence from the training and coaching. All (100%; 8/8) resource room teacher survey respondents felt that the two-day training, three coaching visits, and bi-weekly coaching sessions were relevant and useful for their work. Resource room teachers in the focus groups also highlighted that training was clear and that practical application aspects of training were useful.

> ***"When it was applied practically, meaning when [coach] applied it and started explaining to us practically how to do this, that's the part we benefited the most from."*** C/U model resource room teacher focus group participant

Resource room teachers were positive about the skills developed from the training and coaching. One view expressed during the resource room teacher focus group was that due to the helpfulness of the coaching sessions, teachers did not need to refer back to the teacher manual, which contained guidance on how to implement tasks. Coaching sessions were felt to equip teachers with the skills needed to implement LRF! C/U model.

> ***"The meetings that we did every two weeks were of benefit to us. When [coach] came to us to help us, there was no need for us to go back to the guide."*** C/U model resource room teacher focus group participant

Resource room teachers also spoke about their increased confidence as a result of the training and coaching. During the resource room teacher focus group, some attributed this increase in confidence to the repetitive tasks practised during training sessions; as they became more familiar with tasks, the more their confidence grew.

Equally, coaches suggested during their focus group that the bi-weekly coaching had a positive effect on teacher confidence. Using the space to discuss challenges and share experiences with one another helped them to learn from their peers.

> ***"The teacher faced the challenge but solved it…So, when they meet and exchange experiences and solutions, many of the problems were greatly reduced."*** Coaches focus group participant

Additionally, WhatsApp groups set up by coaches also provided a space to bolster teacher's confidence, when they had anxieties about delivery.

> ***"There were concerns and worries among the teachers, like, 'How am I supposed to do this? This is the first time I'm teaching this way. Will it work or not?'. We provided a lot of emotional support to the teachers and told them, 'We are with you on WhatsApp and in the field."*** Coaches focus group participant

Motivation as a result of training and coaching was less covered during focus groups, but the survey responses already presented in RQ2 indicate that all respondents saw an increase in their motivation to deliver the programme. For further information on resource room teacher outcomes as a result of the training and coaching, see RQ2.

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

Evidence from focus groups showed that resource room teachers need to be supported in various ways during the delivery of the LRF! C/U model. There was agreement that most support provided to resource room teachers should continue during any future delivery of the intervention, but participants also suggested other ways in which support would be helpful, as discussed below.

Support which should continue, for future implementation of LRF! C/U model

Both resource room teachers and coaches suggested that coach support via a WhatsApp group should continue in future, to maintain teacher engagement. The purpose of this group was to answer queries posed by teachers, and to provide a platform for teachers to discuss challenges.

> ***"Even when the teacher feels that she is not facing a major challenge with her students, she often messages us on WhatsApp asking, 'How can I support them? How can I increase their motivation? How can I make them read better?' The teachers' level of commitment to applying it was very high."*** Coaches focus group participant

In addition, hybrid coaching was felt to be beneficial. A balance of in-person visits and online sessions via Zoom were both viewed as helpful by resource room teachers. In-person visits allowed the coach to see the students and the difficulties they were having, whereas Zoom facilitated problem sharing and solving with other teachers.

Lastly, survey feedback on the existing training and coaching indicated that it should continue for future delivery of the LRF! C/U model. Post-training survey responses indicate that the length of the two-day training was deemed appropriate to cover the material needed for LRF. In addition, the post-intervention survey responses suggested the number and duration of coaching sessions provided sufficient support to teachers, and the duration and frequency of the bi-weekly coaching sessions also gave teachers the necessary support.

RQ10: Are there any key contextual factors that appear to facilitate or impede successful implementation of the LRF! C/U model? (IPE)

Several key contextual factors that facilitate and impede implementation of the LRF! C/U model were identified during focus groups and are outlined below.

Facilitators

Effective training for the *coaches* was deemed as vital to delivering the intervention well. Training which enabled coaches to gain a solid understanding of the LRF! C/U model, through a focus on the methodology and practice book, was deemed essential for successful training to be subsequently delivered to resource room teachers.

> ***"I think that these [training] sessions between us and an institution were very important in understanding the methodology, understanding the routine."*** Coaches focus group participant

Barriers

Lack of parental engagement

A lack of parental engagement was a key barrier to the implementation of the LRF! C/U model, as mentioned by coaches and resource room teachers throughout the focus groups. Parents not helping their child with homework or the practice book was a recurring theme throughout. Coaches attributed this to a number of reasons: the parent's culture, parents not having a standard of literacy that enabled them to help, or because parents believed exams were more important than homework.

> ***"Most parents, due to culture, level of education, and the level of attention, as well as the challenges facing the family, for many reasons why children do not receive support."*** Coaches focus group participant

Resource room teachers reported setting up a WhatsApp group to aid parental engagement, which had varied success. Stigma also played a role in parental engagement. Coaches and resource room teachers reported instances where parents withdrew students from the programme. Some parents questioned why their child was in the resource room, linking it to students having a weak ability, whilst others thought that their child did not require additional support.

Similar sentiments to the above were shared from usual practice focus groups. Participants suggested that some students and parents linked the use of the resource room to inability, and as a result, families refused to let students use the resource room. For this reason, the resource room was not used often.

Teachers delivering usual practice shared the importance of family input.

> ***"If everyone doesn't come together – the resource room, the teacher, the principal, and the parents – it's hard for all of us to succeed in one mission."*** Usual practice focus group participant

These difficulties were worsened by students who were struggling facing ridicule outside of school, reducing engagement in the resource room further.

> ***"After leaving the class, he is mistreated or put in embarrassing situations because of his performance – whether it's because he couldn't write or couldn't read."*** Usual practice focus group participant

Teacher and student absence

Absence from both students and teachers was a further barrier affecting the implementation of the LRF! C/U model. When resource room teachers were absent, students would miss whichever LRF! C/U model session was planned to take place. This was problematic as planned sessions could not be repeated, leaving students with a gap in their learning. Student absence was also highlighted by resource room teachers; one view was that some students were involved with working alongside their families, rather than attending school. This meant they did not receive the full intervention. This may also explain why students did not have enough time to use the practice book at home, as highlighted by the coaches.

Resources in public schools

The pressured resources in public schools were viewed as a barrier to the implementation of the LRF! C/U model. In particular, teachers were viewed as not having much available time, considering they already taught long classes, and were then tasked with teaching LRF!, a methodology which requires a lot of teacher effort. Coaches suggested that the programme needed to be more realistic in its outlook on resources available in public schools, in order for the intervention to be successful.

> ***"The conditions for implementation in schools, are not easy…We know that the resources in public schools are very limited... the teacher is already burdened with long classes".*** Coaches focus group participant

Feasibility of the efficacy trial

RQ11: 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 usual practice groups? (IE and IPE)

School and participant retention

All 16 schools remained in the pilot between baseline and endline testing. The student attrition rates were 0% and 3% in the usual practice and C/U model groups, respectively. More than half of eligible students therefore completed the outcome testing in both groups. All eight resource room teachers remained in the pilot from training to the end of implementation and completed both the post-training and post-intervention surveys (one resource room teacher was not available for the focus group discussion).

Data collection

Integrated started baseline EGRA data collection with the16 participating primary schools on February 2nd 2024 and completed data collection within five working days. Similarly, the endline EGRA data collection started on May 15th 2024 and was completed within four working days. On average, it took 13 minutes to complete an EGRA assessment with the participating students.

Suitable design of LRF! C/U model materials and training

Coaches and resource room teachers reported mixed views on whether the LRF! materials were helpful. Coaches spoke positively about the LRF! C/U model as a whole.

> ***"The application, the materials, the interaction, and the pacing of the content were good."*** Coaches focus group participant

In terms of materials, resource room teachers had mixed views on the structure followed in the practice book, which went from reading letters, to syllables to words; some liked it and thought it made logical sense to start with easier elements, while others thought this structure was rigid. There were also mixed views related to the teacher guide; some resource room teachers suggested that the teacher guide was useful in reinforcing learnings from the practice book, whereas others did not use it, finding support from their coach more beneficial. The latter did not seem to be a critique of the teacher guide, but a positive outcome of the coaching.

As detailed in RQ7 and RQ8, training and coaching were considered suitable by resource room teachers.

> ***"The methodology was clear and smooth, so I don't think it needs more than that. We didn't feel like there was anything missing."*** C/U model resource room teacher focus group participant

RQ12: What does the pilot tell us about the feasibility of the resources of an efficacy trial e.g., measurement instruments or specific equipment used (including the PTI diagnostic tool)? (IE and IPE)

Data from the post-training survey indicated that the training materials were appropriate and meaningful; all respondents agreed that the materials were clear, coherent, and understandable.

Views on the practice book were mixed; despite all resource room teachers reporting in the post-intervention survey that they used it in all of their sessions, there were various comments throughout the focus groups that suggested there were changes that could be made. These related to the simplifying of content, the formatting of the book, and the inclusion of pictures.[49] Further information on this is detailed under RQ14.

In terms of the measurement instruments, these also divided opinion. As reported under RQ6, resource room teachers and classroom teachers agreed that there were some students who shouldn't have been identified who were, and vice versa. In addition, the tool was not always perceived to be effective at screening out those with learning difficulties, which was essential, as the LRF! C/U model was not designed to support these students. Identification of the correct students was an area identified as needing improvement in the previous trial, and despite the changes made to the process, one that continues to present a challenge.

As part of the IE, we conducted an additional exploratory analysis to assess the accuracy of the coarse-grain tool in identifying students eligible for the C/U model based on the PTI assessment. We found that the coarse-grain tool by itself had low overall accuracy and low ability to correctly identify students who would benefit from the C/U model. Based on our sample, our findings suggest the coarse-grain screening tool does not have good psychometric properties for identifying students eligible for the C/U model, so it should not be used as the sole measure in any future trials of the LRF! C/U model. However, it is also important to note that findings from teacher focus groups indicate instances where the identification tests were not taken seriously or not fully completed, as detailed under RQ6 earlier. This potentially reduces the accuracy of these measures and limits our ability to make strong claims about our findings. Therefore, we recommend caution when interpreting these findings and suggest further analysis. More detailed findings can be found in Appendix 2.

The EGRA is a widely used and robust measure of early reading literacy and was used as an additional tool in the current pilot to identify eligible students for the intervention. A detailed assessment of how meaningful the different screening tools were in selecting the appropriate students was not possible within the limited scope of this pilot (i.e., it was not possible to undertake EGRA with all students in the selected classrooms).

Assessing readiness for trial

RQ13: What changes, if any, are needed to the logic models? (IPE)

Several key changes to the logic model (Figure 2) were identified, based on the data from the IPE data collection. We suggest that adaptations to the Activities specified in the logic model should be made in order to achieve the Outputs, which in turn are essential to achieving the Outcomes.

First, more work needs to be done with parents in order for them to be 'engaged and supportive', as noted under Outputs. Lack of parental engagement was a key factor highlighted throughout the focus groups, and without it the impact of the programme was felt to be limited. The current Activities (awareness raising meeting for parents, and WhatsApp messages being sent to parents) will need to be built on to improve future engagement.

Second, the Activities used for identifying the 'appropriate students' for the LRF! C/U model need to be adapted, or their implementation improved. Evidence from both this trial and the previous trial indicated that teachers often relied on their own judgement, viewing this as more accurate than the diagnostic tools.

[49] Some of the criticisms made of the practice book (e.g. lack of pictures, formatting) are part of the principles of the LRF! programme.

Classroom teachers, resource room teachers, and coaches all reported that some students selected for the intervention were not suitable, either because their literacy skills were already high, or too low to benefit from the approach.

RQ14: What changes, to the intervention, implementation models, support or materials need to be made? (IPE)

Changes to the intervention

Focus groups elicited various suggestions on changes that could be made to the intervention.

One point of agreement between resource room teachers and coaches was that the LRF! C/U model required more time to implement. Resource room teachers suggested that sessions could be longer or more frequent to better accommodate the listening component of the LRF! C/U model, which would take up a lot of the session, while coaches suggested that the intervention could last for longer than one semester. This was to give the student more time to digest the learning, and to spread out the three lessons taking place per week.

> ***"The struggling child's issue is a time crisis in the first place. They need more time and longer support."*** Coaches focus group participant

Adding to this, one resource room teacher noted that spreading out the days that the LRF! C/U model was implemented would mean that they would not have to teach sessions five days in a row. This was taking place due to the number of students identified as being eligible for the programme, but with a cap of six students per lesson. This may have been a contributing factor to teacher boredom and fatigue with delivery.

Similarly, both resource room teachers and coaches suggested that teachers and students became bored with the repetitive nature of the programme, within and between sessions. They suggested there should be increased variety in the activities delivered as part of LRF! C/U model.

> ***"It feels like a prison for the child due to the same routine, the same format, and the same activities. It becomes very boring for the child…I think presenting the book and the reading material with varied activities is better than keeping it in one form."*** Coaches focus group participant

Changes to implementation models

As in the previous pilot, both coaches and resource room teachers consistently suggested during focus groups the need for more flexibility in supporting students who were struggling. The provision of one-to-one support or small group support were both proposed, or incorporating alternative activities that could be used with these students within the teacher guide.

> ***"Strategies…that have been proven by scientific research are good…but they require some flexibility in application and adjustments to align with the needs of the struggling child."*** Coaches focus group participant

Additionally, there was a view from coaches that the duration of LRF! C/U model sessions should be adjusted based on individual student needs, with it being likely that students who struggle more would need more time.

> ***"I believe it's the biggest challenge facing the success of the project for struggling students. The struggling child's ability to progress in reading is fundamentally tied to the time needed for application."*** Coaches focus group participant

Changes to support

Some changes to support are discussed under RQ9 about the support required for resource room teachers during delivery of the LRF! C/U model. Further suggestions were given during the resource room teachers and coaches focus groups.

Coaches highlighted the importance of teachers receiving additional financial provision. This was particularly in relation to resource room teachers using their own resources to incentivise student engagement, with coaches suggesting these costs should be covered by the programme.

> ***"I felt that teachers buy some incentives for the students, whether it is small rewards or certificates sometimes, sometimes they bring them candies and similar items…I don't want the teacher to be overburdened, beyond their capacity, costs, or the pressure they're already under"*** Coaches focus group participant

Finally, in order to ensure teachers were appropriately supported, resource room teachers suggested the workload should be reduced, due to already long working hours. This ties in with suggestions made about spreading sessions out over a longer time period.

Changes to materials

Any suggestions of changes to materials related to the practice book were raised during focus groups.

The use of pictures in the practice book divided resource room teachers and coaches.[50] While some believed adding pictures would be distracting for students, others felt they would aid in their engagement and understanding, especially for those who struggle to depend on listening as a learning method.

> ***"Images were necessary... I always believe in the importance of sensory diversity for children from first to sixth grade."*** C/U model resource room teacher focus group participant

A minority of resource room teachers noted that they did not use the enlarged version of the practice book that they were given, due to issues with the formatting. They described problems such as empty pages.

> ***"There is a problem with the text formatting [in the enlarged book]. For example, on the first page, the text is crammed together, and on the second page down, there are only two lines and the rest of the page is empty."*** C/U model resource room teacher focus group participant

In discussion with the delivery team, it was explained that these examples were not formatting errors in the enlarged book but were consistent across the student practice book and the teacher one, resulting from the changes to font sizes and spacing recommended after the previous pilot.

Suggestions were made by resource room teachers around the revision of content in the practice book. One viewpoint was that the book should avoid repetition of similar words, as this caused confusion for students who were in the initial phases of learning them. In addition, starting the book with simpler content for lower ability students was recommended by resource room teachers. This was in terms of

[50] As mentioned earlier, the lack of pictures in the book is one of the principles of the programme.

shortening the length and meaning of stories, which were seen as difficult for some students to understand.

> ***"The negative thing was the length of the texts…, the length of the sentences and their difficulty because they were similar in pronunciation."*** C/U model resource room teacher focus group participant

In the previous pilot, recommendations were made around having two distinct practice books, with one aimed at those with a lower ability. This was not implemented for this pilot but appears to be a point that may need to be addressed moving forwards.

RQ15: What can we learn from the pilot about minimal detectable effect size estimates, pre- and post- correlations, intra-cluster correlations, and sample sizes? (IE)

For this pilot trial, 16 primary schools were recruited with eight schools in the C/U model group and eight in the usual practice group. There were 161 students eligible for the trial following the baseline and endline testing. The average number of eligible students per school was 10.1 students.[51] While the average number of classes per school was 4.1 classes, the average number of students per class was 2.4 students.

We undertook power calculations to estimate the sample size required for a future efficacy trial of the C/U model using PowerUp! (Dong & Maynard, 2013). Our power calculations were informed by both the previous and current pilot evaluations of the C/U model. We had the following assumptions for our power calculations:

- The minimal detectable effect size (MDES) is analysed for a three-level cluster randomized controlled trial with programme assignment at the school level for one primary outcome measure. Please note that having more than one primary outcome measure would increase the minimum sample size required for measuring the impact of the programme on selected outcome measures.
- Based on the findings from the current pilot evaluation, we assume an average of 2.4 students per class and 10.1 students per school.
- We use a Type I error rate of 0.05 and a Type II error rate of 0.20 (a power of 0.80).
- The student- and school-level correlations of the primary outcome measure between baseline and endline was estimated to be 0.735 and 0.261, respectively, from the previous pilot evaluation of the C/U model. These correlations are slightly lower for the current pilot evaluation and the student- and school-level correlations of the primary outcome measure are 0.694 and 0.167, respectively. Even though the lower correlations mean larger sample sizes are required, we use the lower correlations from the current pilot evaluation for our power calculations.
- The estimates of Intracluster Correlation Coefficient (ICC)[52] for the primary outcome measure from our pilot sample would be imprecise due to the low sample size as it can be seen by wide

[51] Where there were more than three classes per grade in a school, three classes were randomly selected for the evaluation. The number of eligible students per school could be higher if all eligible classes are included in the programme.

[52] The ICC measures similarity between units in the same cluster; in this case, students 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 reduces the amount of unique information each new observation contributes to the sample.

confidence intervals of the ICCs reported earlier. In the previous pilot evaluation, the school-level ICC was calculated to be 0.01 while the class-level ICC is calculated to be 0.12 for oral reading fluency. In the current pilot evaluation, the school- and class-level ICCs are 0.026 with 95% confidence interval of 0.00093 and 0.44. Given these figures and low sample size, we assume 0.10 school-level and 0.12 class-level ICC, a more conservative assumption which requires a larger sample size.

- Finally, we have a lower attrition rate in the current pilot evaluation compared to the previous pilot evaluation. We assume an overall 15% student-level attrition, which is a slightly more conservative assumption.

Table 25: Sample size calculation - primary outcome analysis for the C/U model

		Catch-up (C/U) Recruitment stage	Catch-up (C/U) Analysis stage after 15% student-level attrition
MDES		0.200	0.204
Pre-test/ post-test correlations	level 1 (student)	0.693	0.693
	level 2 (class)	0.00	0.00
	level 3 (school)	0.167	0.167
Intracluster correlations (ICCs)	level 2 (class)	0.12	0.12
	level 3 (school)	0.10	0.10
Alpha		0.05	0.05
Power		0.8	0.8
One-sided or two-sided?		2	2
Average number of students per school		10.1	8.4
Number of schools	Intervention	67	67
	Usual practice	67	67
	Total	134	134
Number of students	Intervention	677	563
	Usual practice	677	563
	Total	1,354	1,226

Based on the sample size calculation and its assumptions outlined above, the sample size required to detect a MDES of 0.2 standard deviations in a future efficacy trial of the C/U model is 134 primary schools. The number of schools required will be significantly higher if an MDES smaller than 0.2 standard deviations is required for the efficacy trial.

Summary of findings using the IPE Framework

In this section, we summarise the IPE findings detailed above in relation to each research question and IPE dimension of interest.

Table 26: Mapping of IPE dimensions, RQs, data collection methods and analysis

IPE dimension	Research questions addressed	Research and data collection method	Key findings
Fidelity: Dosage	RQ3, RQ5, RQ13	Resource room teacher focus groups Resource room teacher end of programme survey Classroom teacher end of programme survey Coaches focus group	● Nearly all teachers reported delivering 3 or more sessions of LRF! that were 30 minutes or more ● The majority of teachers reported delivering the intervention for 13 weeks or more ● Resource room teachers reported consistent use of the practice book and the pedagogical approaches associated with LRF! ● Coaches and teachers suggested that the sessions/programme could be longer to cover all the content
Fidelity: Quality	RQ5, RQ8, RQ9, RQ10	Resource room teacher FGDs Coaches FGD Resource room teacher post-training survey Resource room teacher end of programme survey Classroom teacher end of programme survey	● Teachers and coaches reported that training had been delivered to a high standard and resource room teachers were engaged and motivated by it ● The majority of resource room teachers agreed that students were engaged during the LRF! sessions in the post-intervention survey

			● In focus groups, coaches and resource room teachers suggested student engagement was more variable
Fidelity: Adaptation	RQ1, RQ3, RQ5, RQ13, RQ14	Resource room teacher FGDs Coaches FGD Resource room teacher end of programme and end of training survey Classroom teacher end of programme survey	● Resource room teachers and coaches reported using incentives and competitions with students and parents to boost engagement, especially with homework tasks ● Resource room teachers and coaches suggested that some students needed 1:1 support and they needed to tailor activities more to individual needs ● The majority of classroom teachers reported LRF! sessions taking place more than three times a week
Reach	RQ3, RQ6, RQ12, RQ13, RQ1	Resource room teacher FGDs Coaches FGD Resource room teacher end of programme survey	● All resource room teachers attended the two-day training session and the majority attended all the coaching and

		Classroom teacher end of programme and end of training survey Attendance data for teacher training and coaching	bi-weekly support sessions ● More than half of resource room teachers believed the correct students were identified for the programme; over a third reported the incorrect students being identified ● More than half of classroom teachers reported high-achieving students being identified for LRF! incorrectly; more than a third indicated students with poor literacy were not identified
Responsiveness	RQ2, RQ3, RQ7, RQ9, RQ10, RQ11, RQ13, RQ14	Resource room teacher FGDs Coaches FGD Resource room teacher post-training survey Resource room teacher end of programme and end of training survey Classroom teacher end of programme survey	● Coaches and teachers reported a high level of engagement with the training and coaching ● The majority of resource room teachers perceived student motivation towards reading to have increased after the programme ● The majority of classroom teachers reported students were motivated to attend LRF! sessions ● Coaches and teachers reported

			that both students and resource room teachers found the repetitive nature of the LRF! method became more boring as the sessions went on, making it more difficult for them to engage ● Coaches and teachers suggested that a more tailored intervention would help to engage those with more literacy difficulties
Perceived impact	RQ2, RQ3, RQ4, RQ13	Resource room teacher FGDs Coaches FGD Resource room teacher end of programme and end of training survey Classroom teacher end of programme survey	● Teachers and coaches reported improvements in resource room teachers' knowledge, confidence and reading to support students struggling with literacy through the LRF! programme ● Teachers reported improvements in students' confidence in reading ● Perceptions of improved literacy levels among students in the LRF! programme were more mixed among teachers

Usual practice	RQ1	Resource room teacher FGDs Usual practice classroom/Resource room teacher FGDs	● Classroom and resource room teachers reported varying levels of training in teaching literacy before participating in the intervention ● Resource room teachers reported particularly few previous opportunities for training ● According to teachers, identifying students in need of additional reading support usually relies more on teacher judgement, with some use of literacy tests in the older grades ● Teachers reported using a wide variety of tasks and texts to engage their students in reading rather than the repetitive structure of the LRF! sessions ● Teachers suggested that sending videos home to support students with their homework, rather than just a book, helps to engage parents and students at home

CONCLUSIONS

This pilot evaluation aimed to assess evidence of promise, feasibility of the intervention, feasibility of a trial, and readiness for trial of a revised version of the LRF! C/U model, based on recommendations from a previous pilot.

Below, we reflect on some of our main findings and how they differ from the previous pilot. In doing so, we note some limitations of the evaluation, as well as drawing out recommendations for future delivery.

Evidence of promise

Improvement in (pre)-literacy level

The impact evaluation employed an RCT design in preparation for a future trial, rather than with the aim of establishing evidence of the causal effect of LRF! on the intended outcomes. The pilot evaluation had a small sample size and so lower statistical power. Studies with lower statistical power have estimates which are highly imprecise. They also have higher likelihood of Type I errors, meaning that an evaluation may conclude that a programme has had an impact when in reality it had no impact. Therefore, the findings from the impact evaluation should be treated with extreme caution, as the study was underpowered to detect differences between the C/U model and the usual practice group.

Unlike in the previous pilot, evidence from the impact evaluation analysis shows indicative evidence of students who received the LRF! C/U model making greater improvements in oral reading fluency (i.e., the primary outcome measure) and on syllable identification and word decoding (i.e., secondary outcome measures) compared to similar students in usual practice schools. Data collected from teachers through focus groups and/or surveys was more mixed; they generally reported quite variable levels of improvement in literacy amongst their students, although it was noted that this could also be related to the students' initial literacy levels and parental support and engagement in homework.

The perceived variability may also relate to the differences seen in improvement across the EGRA subtests. For example, teachers may have been focusing on the higher-level literacy outcomes, such as reading comprehension and listening comprehension, which did not show indicative evidence of improvement between baseline and endline. These items are also the least targeted elements of the LRF! programme. These subtests may also not be sensitive to small improvements in comprehension, given that only one question is asked for children who can read up to 10 words. It may be that the LRF! C/U model has more initial or immediate impact on lower-level literacy, such as syllable identification or that the EGRA subtests are more sensitive to small changes in these aspects of literacy. A future trial may benefit from an additional longitudinal follow-up to assess whether there is a longer-term impact on higher-level literacy outcomes.

Changes to usual practice

The training received and the delivery of the LRF! C/U model was a significant change to usual practice for resource room teachers. The extended period of training and ongoing coaching, and the fact that it was for resource room teachers only, seems to have benefitted these teachers in developing their knowledge about supporting struggling readers, as well as their confidence and motivation to deliver LRF!, compared to the previous trial.

Students used the practice book throughout their lessons and regularly took it home. Minimal adaptations were made to the intervention itself, although resource room teachers did discuss their use of incentives

and competitions to engage students more, especially in homework. They also used strategies to engage and encourage motivation from parents.

Perceptions of the intervention and its effects

Coaches reported an improvement in the intervention materials compared to the previous trial, although there were still some suggestions to improve implementation in the future. Both coaches and resource room teachers felt the intervention should be made longer in order to cover all of the content, support student learning, and reduce workload for resource room teachers in one school term. This would be important for both the C/U and whole-class models[53] of the LRF! intervention.

Resource room and classroom teachers reported increased confidence amongst students receiving LRF! in reading, although there were mixed findings concerning students' motivation and engagement. The repetitive nature of the LRF! method and the amount of content to be covered in the time were both suggested to reduce students' engagement and motivation to attend the LRF! sessions. As in the previous pilot, resource room teachers suggested that more varied methods could be useful in engaging the students over a long period of time, and that there was a need for a more one-to-one, tailored approach for some students that they did not perceive to be possible within the rigid structure of the LRF! C/U model. They also suggested that including weekly exams could help to engage students more, as they would recognise the need to study and practise outside of the classroom.

One unintended negative consequence of participating in LRF! was that teachers reported students missing a range of other classes. Some teachers suggested that LRF! produced a significant disruption to the student's learning. Increasing the length of the intervention in future could potentially reduce this negative effect by reducing the number of sessions per week that students needed to attend outside of regular classes.

The coaches and classroom teachers were positive about the impact that the training had on resource room teachers' ability to deliver effective reading sessions to students 'in need' of additional literacy support, which is an important change from the previous pilot. Resource room teachers' perceptions of this improvement as a result of their training was more mixed, although this may be due to them feeling skilled in this area before the training was provided.

For resource room teachers, there were also negatives of delivering the intervention in terms of increased workload and fatigue, which may have led to some reduction in motivation over the course of the trial. In a future trial, it will be important to consider the best utilisation of resource room teachers in delivering the intervention, including the number of sessions per teacher, as well as the number of students they are supporting.

Overall assessment of evidence of promise

The revised LRF! C/U model shows some indicative evidence of promise, with improvements in specific pre-literacy skills for students receiving the intervention compared to those receiving usual practice. The revised training and support provided for teachers has generally improved resource room teacher knowledge and motivation compared to the previous pilot and implementation fidelity is high. There remain some concerns regarding the length and intensity of the intervention and the negative impacts of this on students and resource room teachers. As in the previous trial, the lack of variety in the methods and the need for a more tailored approach for less able students are key points to address in the future.

[53] The whole-class approach is when classroom teachers deliver LRF! to all students in the class.

Feasibility of intervention

Fidelity of delivery

The improved training seems to have had a positive impact on the fidelity of delivery, with most resource room teachers using the LRF! practice book and following the key pedagogical approaches in the majority of lessons. Teachers generally reported delivering at least three LRF! sessions per week over the 12–14-week intervention period.

As discussed in the previous section, there were mixed findings regarding the perceived engagement of students and the effectiveness of the intervention for all students in need of additional literacy support. Key contextual factors suggested to affect this included parental support and engagement, and regional differences. Individual-level factors were also discussed, including the teaching style of the resource room teacher, and the attention and initial literacy levels of the students. These factors will be important to consider in the IPE for future trials.

Identifying the most appropriate students for the intervention

Most teachers reported that there were students identified for the LRF! C/U model that they did not believe required the intervention, or that there were other students in the class who would have benefitted from it more

Teachers may also have judged literacy ability on different criteria to those used in the study, in which the primary measure was oral reading fluency. As in the previous trial, teachers tended to rely more on their own professional judgement than the scores from tests, including the coarse-grain screening tool they are required to implement at the beginning of each school year. If a screening approach is used for a future trial of the LRF! C/U model, it will be necessary to ensure consistency of implementation across schools and may require additional training for teachers on the utility and precision of the measures used.

It is possible that the lack of improvement across several secondary outcomes was due to the inappropriate students being selected for the study. In cases where students with SEND, for whom the intervention is not designed, have not been correctly screened out, they may show little improvement between baseline and endline. Equally, those with relatively higher scores at baseline may not show as much improvement as lower-achieving students over the course of the intervention. The perceived or actual lack of precision of the screening measures used to identify appropriate students for the intervention remains an issue for the LRF! C/U model.

Training and supporting resource room teachers

Resource room teachers were well supported throughout the intervention, from initial training to the end of delivery. Informal WhatsApp groups and continued coaching sessions allowed resource room teachers to develop a professional learning community and ask questions to each other, as well as the coaches. This promoted a collaborative problem-solving approach and seems to have allowed resource room teachers to further develop their skills, knowledge, and confidence after the initial training was over.

While teachers in the previous trial were generally positive about the training they received, both teachers and coaches believed that resource room teachers, in particular, would benefit from longer training and more focused support. This approach seems to have had a positive impact on the delivery of the intervention in the current pilot and should be continued in a future trial. This may be useful for teachers delivering the whole class model as well.

Overall assessment of feasibility of the intervention

The revised approach to training was successful and, given the fidelity of the delivery, these aspects of the intervention appear to be feasible. The remaining issue is that of identifying appropriate students for the intervention, which has not been addressed by the revised approach to screening taken in this pilot study. Based on this evidence, the C/U model is not feasible unless a more robust screening procedure can be developed and tested. It may be that the students with additional literacy needs can be supported just as well within the previously-piloted whole class approach [54] and that applying the successful aspects of the current pilot could improve training and delivery for all teachers and students within the whole-class intervention.

Feasibility of the efficacy trial

Feasibility of process components

School recruitment was a relatively smooth process and relied heavily on Integrated staff based in Jordan to liaise with schools to explain the study and arrange testing visits. However, it is important to note the differences between regions in recruitment, as this may affect the feasibility of a larger efficacy trial. Another important factor is the reachability of schools; over 60 percent of schools approached by Integrated for this pilot trial were unreachable due to incorrect or disconnected phone numbers. This suggests a need for a larger pool of schools for an efficacy trial, as similar challenges were encountered in the previous pilot evaluation. Therefore, it is essential to develop mitigation strategies to effectively reach a larger sample of schools.

Both school and participant retention were very high, indicating a significant improvement compared to the previous pilot evaluation of the C/U model.[55] As a result, sample size calculations for an efficacy trial can take a fairly conservative estimate of attrition.

Finally, Integrated collected the IPE data, which meant that surveys and focus group discussions could be conducted in Arabic. This was important to ensure that all teachers were able to participate. Future trials will need to ensure this is possible to maintain a representative sample and promote inclusivity in the research.

Feasibility of resources

The LRF! training materials, practice book, and intervention materials were all considered meaningful and appropriate. However, there were some suggestions for improvements, particularly for the practice book. These related to the simplifying of content, the formatting of the book, and the inclusion of pictures. These shortfalls, as discussed earlier, may indicate why most teachers suggested they would not continue to use the practice book as it currently exists after the LRF! trial ended.

The feasibility of the coarse-grain screening and PTI tools is unclear. Exploration of more widely used standardised measures of literacy and cognitive development will be necessary before an efficacy trial of the LRF! C/U model can take place. An example of this sort of standardised screening measure is the Al Arabiya Intelligence and Achievement Tests,[56] which is a psychoeducational battery of tasks tested widely

[54] The whole class model is delivered by classroom teachers to all students in the classroom.

[55] The overall attrition rate for the C/U model from randomisation to endline data collection was 30.1% in the previous pilot evaluation.

[56] https://riversideinsights.com/arabiya-intelligence-achievement

within the Middle East. It is based on the well-known Woodcock-Johnson-IV test battery used across countries and languages for educational testing.[57]

Overall assessment of feasibility of the efficacy trial

While the processes involved in setting up and delivering an efficacy trial appear to be feasible based on the current and previous pilot, there are still concerns over the screening tests used for student selection.

Assessing readiness for trial

Changes suggested by teachers

As in the previous pilot, teachers and coaches suggested improvements to the intervention and materials. A key point that emerged during both pilots was the need for more flexible, one-to-one support for students with additional literacy needs than they felt was possible within the LRF! C/U model. Given that this point has been highlighted in both pilots, we recommend that it is considered and reflected in future LRF! materials, in either the C/U or whole-class model.

Changes to key activities specified in the logic model, namely the screening process and the parent engagement activities, were also identified within our analyses. Incorporating some of the ideas and strategies used by the resource room teachers to engage parents in future trials is likely to support student progress in literacy.

Coaches and resource room teachers agreed that increasing the length of the intervention would have benefits for both students and teachers. If this adaptation can reduce workload on resource room teachers in a full-scale efficacy trial, it may be easier to recruit and retain schools in a larger sample.

Overall assessment of readiness for trial

According to teacher reports, there are a number of important but relatively simple changes that could be made to improve the delivery and impact of the intervention. However, the questions that remain regarding the screening process indicate that the LRF! C/U model is not ready for trial in its current form.

Limitations

There are key limitations to this study that should be taken into account when assessing the evidence.

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.

Given that we were piloting new elements of the LRF! C/U model, rather than a new intervention, we focused IPE data collection only on those elements that had changed. Specifically, these were the teacher

[57] https://riversideinsights.com/woodcock_johnson_iv

training and screening processes. This means that we did not collect data from students involved in the intervention or their parents, which may have provided a more rounded impression of responsiveness and reach of the LRF! C/U model than relying solely on teachers' and coaches' perceptions.

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

As in the previous pilot, the intervention has proved feasible to implement and some aspects that were highlighted for improvement have had the desired effect. These include improved training for resource room teachers, focusing on students in Grades 2 and 3, and making alterations to the practice book.

On the other hand, several aspects highlighted in the previous pilot remain, including:

- Coaches and teachers continue to perceive the time needed to successfully deliver LRF! sessions to be longer than the allocated 12-14 weeks.
- Coaches and teachers suggest that the practice book and method could be tailored to students depending on their needs.
- Additional activities are required to promote parental engagement in the intervention, which is expected to impact student motivation and attainment.
- The appropriate students for the C/U model are not being selected and alternative screening methods are required.

Due to the fundamental issues related to the screening process, we do not recommend that the LRF! C/U model is scaled up to an efficacy trial without additional piloting of appropriate screening measures. Included in this piloting should be assessments of:

- the specificity and sensitivity of the measure(s).
- the reliability and validity of the measure(s) in comparable populations.
- whether the measure(s) can be implemented consistently to a high standard by teachers, rather than specialists.
- whether the screening process is onerous for students, their parents, or their teachers.
- whether there are negative consequences of using the measure, e.g., stigma, etc.

At the current time, the stigma associated with using the resource rooms, and the relatively limited support that exists for students with SEND in Jordan may mean that LRF! is better implemented through a whole-class approach, rather than using the C/U model.

REFERENCES

Abazdi, H. (2012). How to Speed Up Arabic Literacy for Lower-Income Students? Some Insights from Cognitive Neuroscience. Global Partnership for Education. https://openknowledge.worldbank.org/server/api/core/bitstreams/9eddc59a-47fb-56e5-9a43-c4d888103b96/content

Abadzi, H. (2013). *Literacy for All in 100 Days? A research-based strategy for fast progress in low-income countries*. Washington, DC: World Bank.

Abadzi, H. (2017). Improving Arab students' academic achievement: The crucial role of rapid reading and grammar mastery in the early grades. Policy Paper, (20).

Abadzi, H. (2020). Skills to stay: Memory functions in 21st-century education. Cambridge: Cambridge University Press.

Abadzi, H., & Martelli, M. (2014). Efficient reading for Arab students: Implications from neurocognitive research. World Summit of Innovation in Education (WISE).

Alzubi, E. M., & Attiat, M. M. (2021). Language teacher practices predicting students' reading self-efficacy: Jordanian students' participation in PISA 2018. Cypriot Journal of Educational Sciences, 16(6), 3213-3231.

Al-Huri, I. (2012). The Impact of Diglossia in Teaching/Learning the Arabic Course in Sana'a Secondary Schools. Learning the Arabic Course in Sana'a Secondary Schools (June 1, 2012).

Brombacher, A., Collins, P., Cummiskey, C., Kochetkova, E., & Mulcahy-Dunn, A. (2012). Student performance in reading and mathematics, pedagogic practice, and school management in Jordan. Prepared for USAID by RTI International under the Education Data for Decision Making (EdData II) Project, Task order AID-278-BC-00019. Research Triangle Park, NC: RTI International.

Campbell, M. (2008). Appendix 3: Bootstrapping and Variance Robust Standard Errors. 10.1002/9780470755839.app3.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. Psychological Science in the Public Interest, 19(1), 5–51.

Chall, J. (1996). Stages of reading development (2nd ed.). Fort Worth, TX: Harcourt Brace

Christley, R.M. (2010). Power and Error: Increased Risk of False Positive Results in Underpowered Studies. The Open Epidemiology Journal. 3. 16-19. https://doi.org/10.2174/1874297101003010016.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.

Cortés-Albornoz, M. C, Ramírez-Guerrero, S., García-Guáqueta, D. P., Vélez-Van-Meerbeke, A., & Talero-Gutiérrez, C. (2023b). Effects of remote learning during COVID-19 lockdown on children's learning abilities and school performance: A systematic review. International Journal of Educational Development, 101, 102835. https://doi.org/10.1016/j.ijedudev.2023.102835

Dimova, S., Woodbridge, H., Duysak, D., Garwood, E., Ruddick-Trentmann, J., Rezaian, M., Koval, N., Phillips, N., Kuo, T., Fugard, A., Griggs, J., & Phillips, D. (2023). "Let's Read Fluently! Pilot Evaluation Report".

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]

Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R., & Kerr, K. (2019). Implementation and process evaluation (IPE) for interventions in education settings: An introductory handbook. London: EEF.

EEF (2022). Implementation and process evaluation guidance for EEF evaluations. London: EEF.

EEF (2021). Phonics. Available at: [https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit/phonics]

EEF (2022). Statistical analysis guidance for EEF evaluations. London: EEF.

EEF. (2023). Early literacy approaches. Available at: [https://educationendowmentfoundation.org.uk/education-evidence/early-years-toolkit/early-literacy-approaches]

Eviatar, Z., & Ibrahim, R. (2014). Why is it hard to read Arabic? In E. Saiegh-Haddad, M.R. Joshi (Eds.), Handbook of Arabic literacy (pp.77-96). Springer.

Gupta S. K. (2011). Intention-to-treat concept: A review. *Perspectives in clinical research*, *2*(3), 109–112. https://doi.org/10.4103/2229-3485.83221

IEA. (2021). Countries' Reading Achievement. [https://pirls2021.org/wp-content/uploads/2022/files/P21_1-reading-achievement.pdf]

Lee, E. C., Whitehead, A. L., Jacques, R. M., & Julious, S. A. (2014). The statistical interpretation of pilot trials: should significance thresholds be reconsidered?. *BMC medical research methodology*, *14*, 41. https://doi.org/10.1186/1471-2288-14-41

Nusair, L., & Palmer, R. (2023). Weak foundations in Arabic literacy: Drivers of learning poverty in Jordan. International Journal of Educational Development, 99, 102774.

QRF. (September 2020). 'Exploring Jordan's performance'. [https://www.qrf.org/en/what-we-do/research-and-publications/pisa-2018-exploring-jordan%E2%80%99s-performance]

Ritchie, J., Lewis, J., Nicholls, C.M. and Ormston, R. (2013). Qualitative Research Practice: A Guide for Social Science Students and Researchers. London: SAGE Publications Ltd.

RTI (2012). Pupil Performance in Reading and Mathematics, Pedagogic Practice, and School Management in Jordan. [https://pdf.usaid.gov/pdf_docs/pa00jbnh.pdf]

RTI International (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 (2023). Girls have academic advantages and so do boys: A multicountry analysis of gender differences in early grade reading and mathematics Outcomes. https://eric.ed.gov/?id=ED636890

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

Rudnicky, A. I., & Kolers, P. A. (1984). Size and case of type as stimuli in reading. Journal of Experimental Psychology: Human Perception and Performance, 10(2), 231–249. https://doi.org/10.1037/0096-1523.10.2.231

Sainani, K. L. (2012). Dealing with non-normal data. *Pm&r*, *4*(12), 1001-1005.

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 (March 2023). 'Education in a Post-Covid World: Towards a RAPID Transformation. [https://eric.ed.gov/?id=ED627511]

World Bank (2019). Ending learning poverty: what will it take? Washington, D.C., World Bank. [https://openknowledge.worldbank.org/bitstream/handle/10986/32553/142659.pdf?sequence=7]

Zakaria, N. A. I., Saad, M., Rashid, M., & Nor, M. (2021). Systematic Review of Early English Literacy in ELL Children: What Do We Know from A Decade of Research. *3L: Southeast Asian Journal of English Language Studies*, *27*(4).

APPENDICES

Appendix 1: Additional Descriptive Analysis for Pre-literacy Items

The following section presents descriptive statistics for the LRF! pre-literacy items administered to students at baseline and endline testing. The pre-literacy items include the following:

1. Oral vocabulary: Task 1
2. Oral vocabulary: Task 2
3. Recognising Letter Names
4. Recognising High Frequency words

Table 27 below shows the mean pre-literacy scores at baseline and endline for the C/U model and usual practice group. For each pre-literacy item, paired t-tests were conducted to determine whether the mean change from baseline to endline is statistically significant from zero. The corresponding t-statistics and p-values are also presented in Table 27.

Table 27: Paired t-tests for pre-literacy items

	C/U model				Usual practice			
	Mean (SD)		*t*	Sig (2-tailed)	Mean (SD)		*t*	Sig (2-tailed)
Outcome	**Baseline**	**Endline**			**Baseline**	**Endline**		
Oral Vocabulary: Task 1	7.32 (2.89)	8.40 (2.11)	-3.36	0.0011	7.68 (2.01)	7.40 (2.22)	0.89	0.3770
Oral Vocabulary: Task 2	7.70 (2.60)	8.36 (2.04)	-2.34	0.0216	7.60 (2.18)	7.37 (2.30)	0.63	0.5336
Recognising Letter Names	2.58 (4.07)	8.35 (8.56)	-6.30	0.0001	3.78 (4.64)	3.28 (4.55)	0.78	0.4366
Recognising High Frequency Words	2.16 (3.26)	4.03 (4.73)	-3.81	0.0002	1.86 (2.29)	2.09 (2.83)	-0.74	0.4617

Oral Vocabulary (OV): Task 1 and Task 2

Oral Vocabulary (OV) Task 1 measures the total number of different food items a student can name correctly.[58] In the C/U model group, the baseline score has a mean of 7.32 (SD = 2.89) and the endline score has a mean of 8.40 (SD = 2.11). Students in the C/U model group name, on average, 1.08 food items more from baseline to endline, and this difference is statistically significant at 1%. OV Task 2 measures the total number of different animals a student can name correctly.[59] The C/U model group's baseline score has a mean of 7.70 (SD = 2.60) and the endline score has a mean of 8.36 (SD = 2.04). Similar to OV Task 1,

[58] The measure takes value between 0 and 10.

[59] The measure takes value between 0 and 10.

students in the C/U model group, on average, name significantly more animals from baseline to endline (given as p-values in Table 27). On the other hand, for students in the usual practice group, there is no statistically significant difference from baseline to endline in both tasks.

We further compared the mean scores of both OV Task 1 and OV Task 2 at endline for the C/U model and usual practice group. We find that students in the C/U model group, on average, named one food item more than students in the usual practice group, and one animal more. These differences are statistically significant at 1% level (p = 0.0045 for OV Task 1 and p = 0.0045 for OV Task 2).

Recognising Letter Names (RLN)

The Recognising Letter Names (RLN) item measures the number of correct letters a student can name per minute.[60] In the C/U model group, the baseline score has a mean of 2.58 (SD= 4.07) and the endline score has a mean of 8.36 (2.04). Students in the C/U model group, on average, name approximately six letters more from baseline to endline and this difference is statistically significant at 1% level. Similar to the other pre-literacy items, we do not observe a statistically significant difference from baseline to endline for students in the usual practice group. Furthermore, we compared the mean RLN score at endline for the C/U model and usual practice group. Students in the C/U model group, on average, name five letters more at endline compared to students in the usual practice group. This difference is statistically significant at 1% level (p = 0.0001).

Recognising High Frequency Words (RHFW)

The Recognising High Frequency Words (RHFW) item measures the number of correct words a student can read per minute.[61] In the C/U model group, the baseline score had a mean of 2.16 (SD= 3.26) and the endline scores had a mean of 4.03 (SD= 4.75). Similar to the other pre-literacy items, students in the C/U model, on average, name more correct letters at endline than baseline and there is no statistically significant difference from baseline to endline for students in the usual practice group. Finally, we compared the mean RHFW score at endline for the C/U model group and usual practice group. Students in the C/U model group read, on average, approximately two words more than students in the usual practice group and this difference is statistically significant at 1% level (p = 0.0014).

[60] The measure takes value between 0 and 18.

[61] The measure takes value between 0 and 5.

Appendix 2: Assessing the validity of the coarse grain screening tool

Classroom teachers involved in the LRF! programme administered the coarse-grain screening tool to all students in Grades 2 and 3 in Jordan during the 2023-2024 academic year. The tool is part of standard practice in Jordanian schools and is used to assess students' Arabic reading proficiency and identify students needing support. Teachers gathered data on each student's performance which was then compiled by QRTA. The data from the course-grain tool identified the lowest-performing 20% of students in each class based on their reading scores. The scores from the tool are a preliminary indicator of reading ability and do not offer detailed diagnostic information but serve as the first step in identifying students for more focused assessment. This group was then selected for further diagnostic assessment carried out by practitioners from PTI. If a student exhibited indicators of learning difficulties, delay in language development and global development delay during the assessment, they were not deemed as eligible for the C/U model.

The scores from the coarse-grain tool and the outcomes of PTI assessment were shared with NatCen. Using this data, we undertook further analysis to assess the accuracy of the coarse-grain tool in identifying students eligible for the C/U model. This Appendix summarises the findings of this assessment.

Method used to assess validity of screening tool

We assessed the test score validity for the coarse-grain screening tool against the PTI assessment using Area Under the Receiver Operating Characteristic (AUROC) curve and by calculating the Youden statistic. The AUROC curve measures how well the screening tool distinguishes between students who need the LRF! C/U model or not. The higher the AUROC score, the higher the accuracy of the tool in correctly classifying students.

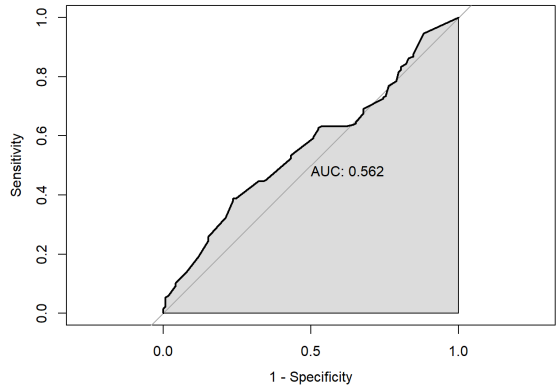
The Youden statistic balances the true positive rate (sensitivity) against the true negative rate (specificity) to select the optimal cut off score for identifying students who need the LRF! C/U model. Sensitivity, or true positive rate, quantifies how well the screening tool identifies true positives. In this context, it refers to how well the tool can classify students who truly need the C/U model according to the PTI assessment. Specificity, or true negative rate, quantifies how well the screening tool identifies true negatives. This shows how well the tool can classify students who truly do not need the C/U model according to the PTI assessment.

Furthermore, the ground truth is an indicator showing the true classification of a case, reflecting the actual status or outcome based on a diagnostic test. In this context, the ground truth is student eligibility for the C/U model based on the PTI assessment.

AUROC curve

The AUROC curve showed a value of 0.56. This indicates that the tool has low discriminative ability in differentiating between students who need to undergo LRF! C/U model (true positives) and those who do not (true negatives). Higher AUROC values (above 0.7) indicate good tool performance, while values close to 0.5 indicate that the tool performs slightly better than random chance.

Figure 29: AUROC curve



Furthermore, the sensitivity of the screening tool is 52%, which means that the tool correctly classified students who truly need the C/U model about 52% of the time, indicating low sensitivity of the tool. The specificity of the screening tool is 54%, meaning that the tool correctly classified students who truly do not need the C/U model about 54% of the time, indicating low specificity of the tool.

Youden statistic

A Youden statistic with a cut-off point above 0.5 is generally considered acceptable, as it reflects a better balance between correctly identifying true positives and true negatives. The optimal cut-off point determined by the validity assessment is a score of 0.42. This is the optimal threshold used to classify an instance as either positive or negative. This means that scores at or below 0.42 would identify students for further assessment.

At this cut off point, the true positive rate was 0.39, indicating that the tool correctly identified 39% of the students who needed C/U model. As shown in Table 28, the low sensitivity shows that many students who need the C/U model were missed (125 out of 204 were false negatives). However, the tool has relatively high specificity (0.76), showing it performs better at identifying true negatives, i.e., those who do not need the C/U model were identified 76% of the time. The false positive rate is low with 28 students being incorrectly identified as not needing intervention. The relatively low sensitivity coupled with a higher specificity suggests that the chosen cut-off point may be set too high, causing the tool to miss many true positives.

Table 28: Confusion matrix

	Predicted positive	Predicted negative
Actual positive	79 (True positive)	125 (False negative)
Actual negative	28 (False positive)	90 (True negative)

Summary

We undertook additional explanatory analysis to assess the accuracy of the coarse-grain tool in identifying students eligible for the C/U model. For this analysis, we used the scores from the coarse-grain tool and the outcomes of PTI assessment. Our findings from this exploratory analysis first showed that the AUROC value of 0.56 indicates that the coarse-grain tool has only slightly better discriminative ability than random chance (AUROC = 0.5). Furthermore, sensitivity (0.52) and specificity (0.54) are similarly low, suggesting that the tool has limited ability to identify students who truly need the C/U model (i.e., true positives) and students who truly do not need the C/U model (i.e., true negatives). Overall, our results demonstrate that the coarse-grain screening tool does not have good psychometric properties for identifying students eligible for the C/U model.

However, it is important to note that findings from focus groups with teachers (see section for RQ6) revealed instances where the identification tests were not taken seriously or not fully completed, potentially reducing the accuracy of these validity measures. Therefore, we recommend caution when interpreting these findings and suggest further analysis, such as replicating the analysis with full compliance with the testing procedure or employing alternative validity measures, to better understand the diagnostic performance of the coarse-grain tool.

Appendix 3: Adapted EGRA with additional pre-literacy 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
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	○ One shift ○ Morning Shift

	○ Evening Shift
7. Name of Evaluator	
8. Evaluator Code	
9. Grade	○ Second Grade ○ Third Grade
10. Division	
11. Child Number	
12. Child's Date of Birth	Month: _____ Year: _____
13. Child's Gender	○ Girl ○ Boy
14. Exam Start Time:	_____:_____ Choose One time slot: ○ Morning ○ Evening

Section 1 Print awareness	60 Seconds
— 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. Ensure the book is in the hands of the child.	
— Let's Begin — With the book in your hands, can you show me the front of the book? [Include here instructions to the enumerator to indicate if the child correctly or incorrectly identified the front of the book]. — Thank you. Now can you open the book to the first page and point to where we can begin reading the story? [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.	Take the book back from the child before moving on to the next section

Section 2 Oral vocabulary	60 Seconds
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. Clearly put a tick (/) in the box for each correct word. MULTI-SELECT 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	After 30 seconds, you will tell the child to 'stop'. The Early stop rule: If the child hesitates to name things you can eat after 5 seconds, say "thank you" and stop the exercise.
Now I would like you to name as many animals as you can. Clearly put a tick (/) in the box for each correct word.	After 30 seconds, you will tell the child to 'stop'.

MULTI-SELECT 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	The Early stop rule: If the child hesitates to name animals after 5 seconds, say “thank you” and stop the exercise.

Section 3 Recognise Letter Names	**30 Seconds**
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. Now let's do this exercise: tell me the name of this letter [and point to the first letter]: 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.” [[دار – زيت – خيوط – ملعب - شتاء]] Clearly put a tick (/) on any mistake the child makes. In the event that the child corrects him/herself, circle the sign (/) that you previously made for him/her. Put a tick (/) on the last letter the child identifies.	If the child hesitates to name the letter for more than 3 seconds, point to the next letter and say: “Let’s continue, please.” The Early stop rule:

	If you mark the first three answers as wrong and the child does not correct any mistakes, say “thank you” and stop the exercise.

Section 4: Read high frequency words	
— 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]". — Let's Begin — Can you try the next word? [add 5 high frequency words, use one high frequency word as the example]. Clearly put a tick (/) on any mistake the child makes. In the event that the child corrects him/herself, circle the sign (/) that you previously made for him/her.	

Section 5: Recognize Letter Sounds	60 Seconds
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." Now let's do this exercise: tell me the sound of this letter [and point to the letter K]: Good, the sound of this letter is "k." The sound of this letter is "K". Let's try another example: Tell me the sound of this movement [point to the aperture]: Well done, the sound of this movement is "-" The sound of this movement is "—" Did you understand what is required from you? 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? Let's Begin	After 60 seconds, you will tell the child to 'stop'. If the child hesitates to read the letter for more than 3 seconds, point to the next letter and say: "Let's continue, please." <u>The Early stop rule:</u> 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.
Clearly put a tick (/) on any mistake the child makes. In the event that the child corrects himself, circle the sign (/) that you previously made for him. Put a tick (/) on the last letter the child reads. Example: for K	

	10	9	8	7	6	5	4	3	2	1
(10)	نـ	ـه	حـ	ـة	ج	ـة	هـ	ف	تـ	بـ
(20)	ـِ	و	قـ	ز	س	خ	ـهـ	فـ	ة	ب
(30)	ق	ص	ب	ـغـ	ق	و	ز	خ	ـة	صـ
(40)	ذ	ظ	ي	نـ	و	ـهـ	ط	شـ	ض	ـعـ
(50)	ء	خـ	ش	ق	ن	ي	غ	د	بـ	غ
(60)	م	ن	ط	سـ	فـ	ث	ح	ذ	ض	ب
(70)	ـعـ	ط	ث	ق	ج	ـه	ضـ	ـه	ح	تـ

(80)	و	ص	جـ	د	ذ	خـ	ـهـ	جـ	ظ	ل
(90)	س	لـ	ذ	خـ	ث	ء	مـ	خ	عـ	ز
(100)	خـ	هـ	ـهـ	ضـ	ـعـ	ح	و	ذ	ب	ـُ

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**
▯ 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". ▯ 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" ▯ Let's try another example: read this passage [point to the passage]: Well done, we read this passage like this "C" The sound of this movement is "Su" ▯ Did you understand what is required from you? 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? ▯ Let's Begin	After 60 seconds, you will tell the child to 'stop'. If the child hesitates to read the letter for more than 3 seconds, point to the next letter and say: “Let’s continue, please.” ✋ <u>The Early stop rule:</u>
Clearly put a tick (/) on any mistake the child makes. In the event that the child corrects himself, circle the sign (/) that you previously made for him. Put a tick (/) on the last letter the child reads. Example: for K	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.

	10	9	8	7	6	5	4	3	2	1
(10)	دي	فو	بـِ	هِ	مي	زِ	وَقْ	حِب	تى	ظَ
(20)	ءُ	رُ	كُ	قو	رس	قِ	رِ	الْ	قَب	هٌ
(30)	مَ	إِ	كو	ذي	ظَ	جا	حَو	دي	هَ	يَحْ
(40)	طو	لى	ةِ	ذا	حًا	دًا	رو	تِ	رِ	هِمْ
(50)	نِ	دَ	بـِ	كَتْ	مِحْ	نٌ	مِنْ	صَوْ	ئِـ	دا
(60)	ها	لَتْ	قِ	ضِ	عِنْ	را	ةٍ	جا	دا	فِ
(70)	ضِ	صَوْ	أَلْ	أَنْ	عُصْ	ـهَ	سُ	حَتْ	مُ	رُ
(80)	قا	يَنْ	رى	خَ	قَةَ	دَ	عُ	طَ	ـهِ	خى
(90)	بَعْ	غَيْ	رٍ	كِنْ	إِ	تُ	مَزْ	مَنْ	عِنْ	عا
(100)	عا	رخ	حا	أَج	تَنْ	في	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
▯ This is a short story, focus well and read it correctly, aloud and as quickly as possible. - 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." - Start reading. ready? Let's Begin	Pull the text of the story in front of the child and ask them the questions below. Leave the child maximum 15 seconds to answer each question. 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.
After 60 seconds, you will tell the child to 'stop'. If the child hesitates to read the letter for more than 3 seconds, point to the next letter and say: “Let’s continue, please.” ✋<u>The Early stop rule:</u>	▯▯I will now ask you some questions about the story that I read. Answer the questions correctly.

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.				
Clearly put a tick (/) on any mistake the child makes. Tick the last word the child reads with (I)	Put a tick (X) in the box that corresponds to the child's answer, and then move on to the next question.			
		No answer	Not Correct	Correct
Dima is a student in the third grade. She likes to read books and writing stories 10	What does Dima like? Reading books and writing stories			
Dima went with her classmate Farah to the school library 18	Where did dima go with her classmate? To the school library			
Farah read a book about space, and Dima chose a story about birds 28	What did Farah read? A book about space			
Farah asked: Why do you like reading stories? 34	What did Farah ask her classmate? Why do you like reading stories?			
Dima answered confidentally: I dream about becoming a writer for children 42	Why does Dima dream about becoming a writer for children? Because she likes reading and writing, because she wants to be famous, to			

	have more networks, because she likes children			
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.				

Appendix 4: Success criteria, as detailed in the LRF! C/U model protocol

Dimension	Research Question	Criteria [linked to logic model]	Data Collection Methods	Anticipated sample size	Time of Data Collection
Evidence of Promise (EP)	EP1. In what ways, and to what extent, does the LRF! C/U model affect school, teacher, and pupil practice as compared to usual practice and learning?	EP1a. Resource room teachers use the updated LRF! practice book in more than half of their literacy lessons and follow the pedagogical approaches outlined through the teacher manual, training and coaching *[A3].*	FGD class/resource room teachers (usual practice)	8	During programme
			FGD resource room teachers (intervention)	8	End of programme
			Surveys with resource room teachers (intervention)	8	After training and end of programme
			FGD with coaches	2	Towards end of programme
		EP1b. Attendance data show that resource room teachers attended 2-day training and three coaching sessions, and the majority of additional bi-weekly meetings with coaches *[A1, A2]*	Attendance data for training	8	During programme
		EP1c. Students use the practice book in more than half of lessons and show motivation to take it home *[ST2]*	FGD resource room teachers (intervention)	9	End of programme
			Surveys with resource room / class teachers (intervention)	16 (8 per survey)	End of programme
			FGD with coaches	2	Towards end of programme

	EP2. How do teachers perceive the intervention and any changes that it has delivered?	EP2. More than half of classroom and resource room teachers perceive that there is a positive value in this intervention compared to/in addition to usual practice: - Resource room teachers report positive outcomes of the new training and support system put in place by QRTA on their delivery of literacy lessons compared to usual practice *[O1, LT1]* - Resource room teachers report increased knowledge, confidence and motivation to deliver effective reading sessions *[ST1]* - Resource room / class teachers believe that students involved in the CU programme have increased confidence and motivation to practice reading *[ST2]* - Resource room / class teachers agree that students involved in the LRF! C/U programme have improved (pre-) literacy level *[LT2]*	FGD resource room teachers (intervention)	8	End of programme
			Surveys with resource room teachers (intervention)	8	After training and end of programme
			Surveys with class teachers (intervention)	8	End of programme
	EP3. Is there evidence to support the revised logic model?	EP3a. Coaches and intervention teachers report improved resource room teacher ability to deliver more effective reading sessions to students 'in need' of additional literacy support *[LT1]*	FGD resource room teachers (intervention)	8	End of programme
			Surveys with resource room teachers (intervention) Surveys with class teachers (intervention)	8 8	After training and end of programme End of programme

		EP3b. Coaches and intervention class and resource room teachers agree that students involved in the LRF! C/U programme have improved (pre-) literacy level *[LT2]*	FGD resource room teachers (intervention)	8	End of programme
			Surveys with resource room / class teachers (intervention)	16 (8 per survey)	End of programme
			FGD with coaches	2	Towards end of programme
		EP3c. Results from the EGRA tests suggest that the LRF! C/U programme could improve oral reading fluency and specific sub-domains of literacy attainment[62] *[LT2]*	EGRA test	204	Baseline and endline
	EP4. Is there any evidence of unintended consequences (negative or positive) as a result of the implementation of the LRF! C/U programme?	EP4. Teachers and coaches report minimal or no negative consequences as a result of the implementation of the LRF! C/U programme on e) students participating in the programme f) other students in the class g) Resource room or class teachers h) parents	FGD resource room teachers (intervention)	8	End of programme
			Surveys with resource room /class teachers (intervention)	16 (8 per survey)	End of programme
			FGD with coaches	2	Towards end of programme
Feasibility of Intervention (FI)	FI5. Was the LRF! C/U model delivered as intended in terms of dosage, nature and	FI5. Resource room teachers and coaches report that the intervention was delivered as intended in terms of	FGD resource room teachers (intervention)	8	End of programme

[62] We need to be cautious about interpreting the results of the end line assessment given the small sample size

	quality? What modifications were made, with what implications?	d) Dosage: at least two out of three of the 30-minute LRF! C/U sessions are delivered per week for nearly all of the programme e) Nature: there is evidence that the practice book is being used regularly as part of the lesson, and teachers adopt all pedagogical approaches in each session ("I do", "we do", "you do") f) Quality: LRF! C/U sessions were effective and engaging for students 'in need' of additional literacy support *[D2, ST2, LT1]*	Surveys with resource room teachers (intervention)	8	End of programme
			FGD with coaches	2	Towards end of programme
	FI6. What is the learning about the use of the PTI diagnostic tool? How successful is it, in use, at identifying the most appropriate students for the C/U model?	FI6a. Teachers and coaches agree that the coarse-grain screening tool + PTI tool effectively identified appropriate students for the CU programme	FGD resource room teachers (intervention)	8	End of programme
			Surveys with resource room / class teachers (intervention)	16 (8 per survey)	End of programme
			FGD with coaches	2	Towards end of programme
		FI6b. Teachers and coaches report minimal or no negative consequences as a result of the implementation of the PTI tool with students (on students or parents) *[D1, O2]*	FGD resource room teachers (intervention)	8	End of programme
			Surveys with resource room / class teachers (intervention)	16 (8 per survey)	End of programme
			FGD with coaches	2	Towards end of programme

	FI7. What were the facilitators and barriers to engagement in the resource room teacher training and coaching sessions?	n/a	FGD resource room teachers (intervention)	8	End of programme
			Surveys with resource room teachers (intervention)	8	After training and end of programme
			FGD with coaches	2	Towards end of programme
	FI8. To what extent do resource room teachers develop sufficient skills and confidence through the training and coaching?	FI8. Resource room teachers and coaches perceive that the resource room teachers have the skills and confidence to effectively deliver the CU programme following the training and coaching sessions *[O1, ST1, LT1]*	FGD resource room teachers (intervention)	8	End of programme
			Surveys with resource room teachers (intervention)	8	After training and end of programme
			FGD with coaches	2	Towards end of programme
	FI9.What do we know about how resource room teachers need to be supported (coached) during delivery?	n/a	FGD resource room teachers (intervention)	8	End of programme
			Surveys with resource room teachers (intervention)	8	After training and end of programme
			FGD with coaches	2	Towards end of programme
	FI10. Are there any key contextual factors that appear to facilitate or impede successful	n/a	FGD resource room teachers (intervention)	8	End of programme

	implementation of LRF! C/U model?		Surveys with class / resource room teachers (intervention)	16 (8 per survey)	After training and end of programme
			FGD with coaches	2	Towards end of programme
Feasibility of Trial (FT)	FT11. 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 usual practice groups?	FT11a. Evidence that there is enough in place to allow the intervention to take place the following year at scale: a) there are enough participants trained to act as trainers/coaches b) school/participant retention rates during intervention and evaluation are high c) the intervention materials and training is suitably defined and developed d) the process of using the scores from coarse-grain tool and PTI tool to identify appropriate students for the CU programme is possible on a larger scale within the time period necessary for the efficacy trial (i.e. pre-EGRA testing and programme implementation) e) the process of using the scores from coarse-grain tool and PTI tool, followed by baseline and endline testing, is appropriate for the target group, (i.e., is not too onerous for this age group; is considered acceptable by parents/teachers) f) the working relationships / coordination between partners has been positive over the course of the pilot	Delivery team assessment	N/A	End of programme
			FGD resource room teachers (intervention)	8	End of programme
			Surveys with class / resource room teachers (intervention)	16 (8 per survey)	End of programme
			FGD with coaches	2	Towards end of programme

		FT11b. More than half of eligible students[63] complete the outcome testing in both intervention and usual practice groups	Delivery team assessment	N/A	End of programme
			EGRA data	204	Baseline and endline
	FT12. What does the pilot tell us about the feasibility of the resources of an efficacy trial, e.g. measurement instruments or specific equipment used? {including use of the PTI tool}	FT12a. The training materials, practice books and measurement instruments are appropriate and meaningful, i.e. the EGRA and PTI tests provide relevant data about literacy attainment and developmental difficulties, respectively, and are age- and context-appropriate	EGRA data	204	Baseline and endline
			FGD resource room teachers (intervention)	8	End of programme
			Surveys with resource room teachers	8	After training and end of programme
			Surveys with class teachers (intervention)	8	End of programme
			FGD with coaches	2	Towards end of programme
			Delivery team assessment	N/A	End of programme
			EGRA data	204	Baseline and endline

[63]Eligible students are defined as those who were included in the sample and received the intervention after applying inclusion/exclusion criteria from the coarse-grain screening tool, PTI, and EGRA assessments.

		FT12b. Any modifications to these tools are identified, based on evidence from the pilot, and are possible to implement if scaled up	FGD resource room teachers (intervention)	8	Towards end of programme
			Surveys with class/resource room (intervention)	16 (8 per survey)	End of programme
			FGD with coaches	2	Towards end of programme
			Delivery team assessment	N/A	End of programme
		FT12c. Delivery partners (Integrated and PTI) have sufficient capacity to conduct all the pre-programme assessments (PTI tool and EGRA assessments) in more schools within the time period necessary for the efficacy trial	Delivery team assessment	N/A	End of programme
		FT12d. Sufficient numbers of training materials and practice books, even after modifications, can be available by the time required	Delivery team assessment	N/A	End of programme
		FT12e. Funding is available for the efficacy trial	Delivery team assessment	N/A	End of programme
Readiness for Trial (RT)	RT13. What changes, if any, are needed to the logic model?	n/a	FGD resource room teachers (intervention)	8	End of programme
			Surveys with class / resource room teachers	16 (8 per survey)	End of programme
			FGD with coaches	2	Towards end of programme
	RT14. What changes to the intervention, implementation models, support or	n/a	Surveys with resource room teachers (intervention)	8	After training and end of programme

	materials need to be made?		Surveys with class teachers (intervention)	8	End of programme
			FGD with coaches	2	Towards end of programme
			FGD resource room teachers (intervention)	8	End of programme
	RT15. What can we learn from the pilot about minimal detectable effect size estimates, intra-cluster correlations, pre-and-post correlations and sample sizes?	n/a	EGRA	204	Baseline and endline

Appendix 5: IPE surveys; post-training, and post-intervention

Resource room teacher post-training survey

INTRODUCTION SCRIPT

Hello,

My name is [Name] from [Organisation], calling on behalf of the Queen Rania Foundation. I am calling to conduct a short survey with you about the training you were enrolled in for the Let's Read Fluently! programme.

We are contacting all the resource room teachers who were enrolled in the training to ask them about their experiences. The survey should take around 10 minutes to complete and I will read you the questions over the phone and enter your answers into the computer. The data will be anonymised and shared securely with the researchers analysing the project on behalf of the Queen Rania Foundation, and you will not be identified in any analysis or reports that are shared with schools. These reports will include group results from this survey alongside data from other surveys and focus groups. Your individual responses will not be shared with your school, or with the Let's Read Fluently! trainers and coaches.

You do not have to take part, and you can change your mind during or after the survey, and your data will be removed. Are you happy to continue with the survey now?

Section A

1. Teacher code:

2. In which region is your school?
    a. North of Jordan: Irbid, Mafrqa
    b. South of Jordan: Karak, Tafila, Maan, Aqaba
    c. Middle of Jordan excluding Amman: Madaba, Balqa, Zarqa
    d. Amman

3. How long have you been teaching in the Resource Room?
    a. Less than a year
    b. 1-2 years
    c. 3-5 years
    d. 6-10 years
    e. More than 10 years

4. Did you attend the 2-day training on Let's Read Fluently! delivered by QRTA?
    a. 1st and 2nd days
    b. 1st day only
    c. 2nd day only
    d. Less than one day
    e. No

IF A, GO TO 4.

IF B, C or D: If you did not attend the full 2 days of training, what was the reason for your absence?

a. My school could not release me for the training
b. I was unable to attend the training due to the geographic distance.
c. I was unwell or absent from work when the training took place
d. Something else

IF TEACHERS DID NOT ATTEND ANY OF THE TRAINING, MOVE TO END.

SECTION B: ABOUT THE TRAINING CONTENT AND MATERIALS

5. 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	5	4	3	2	1
The training materials were clear, coherent and easy to follow	5	4	3	2	1
The practical exercises were useful and helped me to understand the Let's Read Fluently! programme	5	4	3	2	1
The training was an appropriate length to cover all the material required for the Let's Read Fluently! programme	5	4	3	2	1

SECTION C: ABOUT THE TRAINERS

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

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
The trainers were knowledgeable and able to explain key concepts well	5	4	3	2	1
The trainers presented the materials clearly and logically	5	4	3	2	1

The trainers were able to answer questions about the Let's Read Fluently! programme	5	4	3	2	1
The trainers were punctual and managed the timing of the sessions well	5	4	3	2	1
The trainers created a motivating and inspiring training environment	5	4	3	2	1

SECTION D: ABOUT THE TRAINING CENTRE AND ADMINISTRATION

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

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
The training facility was in a suitable location	5	4	3	2	1
The training facility was of good quality and well-equipped for the training	5	4	3	2	1
The training facility was an appropriate size for the number of participants	5	4	3	2	1
The number and duration of breaks during the training was appropriate	5	4	3	2	1

SECTION E: ABOUT THE TRAINING OBJECTIVES

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

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
The training increased my understanding of how children learn to read	5	4	3	2	1
The training gave me a good understanding of the Let's Read Fluently! programme	5	4	3	2	1
The training helped me to develop the skills I need to	5	4	3	2	1

deliver the Let's Read Fluently! programme					
I feel confident after the training to deliver the Let's Read Fluently! programme	5	4	3	2	1
I feel motivated after the training to deliver the Let's Read Fluently! programme to students	5	4	3	2	1
Based on my training, I believe that Let's Read Fluently! programme will be effective in supporting students to improve their reading.	5	4	3	2	1

END OF SURVEY SCRIPT

That is the end of the survey, thank you for your time in completing it. If you have any questions about the survey or the data collected, you can speak to your Let's Read Fluently coach, who can put you in touch with the research team.

Resource room teacher end of programme survey

INTRODUCTION SCRIPT

Hello,

My name is [Name] from [Organization], calling on behalf of the Queen Rania Foundation. I am calling to conduct a short survey with you about how you have found different elements of the Let's Read Fluently! Programme, including the coaching, implementation of the sessions, and outcomes of the programme.

We are contacting all the resource room teachers enrolled in the 2nd pilot of LRF! to ask them about their experiences. The survey should take around 10 minutes to complete and I will read you the questions over the phone and enter your answers into the computer. The data will be anonymised and shared securely with the researchers analysing the project on behalf of the Queen Rania Foundation, and you will not be identified in any analysis or reports that are shared with schools. These reports will include group results from this survey alongside data from other surveys and focus groups. Your individual responses will not be shared with your school, or with the Let's Read Fluently! trainers and coaches.

You do not have to take part, and you can change your mind during or after the survey, and your data will be removed. Are you happy to continue with the survey now?

Note for interviewer (please read out question and all response options, rather than just the question)

SECTION 1 – BACKGROUND

1.1)Teacher code:

SECTION 2 – COACHING

2.1) How many coaching sessions did you receive on Let's Read Fluently!, offered by QRTA?

a) 1
b) 2
c) 3

IF C, GO TO 2.2

IF A or B: What was the reason for having less than three coaching visits?

a) Three coaching visits were not offered
b) Coaching visits could not be coordinated between QRTA and my school
c) I was unwell or absent from work when the coaching took place
d) Cancellation from QRTA coaches
e) Something else

COACHING VISITS: CONTENT AND MATERIALS

2.2) To what extent do you agree with the following:

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
The coaching visits were relevant and useful for my work	5	4	3	2	1
The number and duration of coaching visits gave me sufficient support for teaching the Let's Read Fluently! programme	5	4	3	2	1

BI-WEEKLY ONLINE MEETINGS

2.3) How many bi-weekly online coaching meetings did you attend?

a) 5-6
b) 3-4
c) 1-2
d) 0

If A, GO TO 2.4

If B, C or D: What was the primary reason for your absence from the bi-weekly online sessions?

a) My school could not release me for the coaching
b) I was unable to attend the coaching due to time pressure
c) I was unwell or absent from work when the coaching took place
d) I did not feel I required additional coaching sessions
e) I was not offered the bi-weekly coaching sessions
f) I had no internet connectivity/technical issues
g) Something else

2.4) To what extent do you agree with the following:

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
The bi-weekly sessions were relevant and useful to my work	5	4	3	2	1
The bi-weekly sessions gave me the support needed to improve my	5	4	3	2	1

teaching in the Let's Read Fluently! Programme					
The bi-weekly sessions were an appropriate duration and frequency	5	4	3	2	1
Bi-weekly sessions were adaptable according to my own questions and desired support	5	4	3	2	1

COACHING VISITS: ABOUT THE COACHES

2.5) To what extent do you agree with the following:

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
The coach running the sessions was knowledgeable	5	4	3	2	1
The coach was able to answer questions about the Let's Read Fluently! programme	5	4	3	2	1
The coach created a motivating and inspiring environment to learn in	5	4	3	2	1

COACHING SESSIONS AND BI-WEEKLY MEETINGS: ABOUT THE TRAINING OBJECTIVES

2.6) To what extent do you agree with the following:

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
The coaching and bi-weekly meetings have helped me to develop the skills I need to deliver the Let's Read Fluently! programme	5	4	3	2	1
The coaching and bi-weekly meetings have increased my feelings of confidence in delivering effective reading sessions	5	4	3	2	1
My motivation to deliver effective reading sessions to students has	5	4	3	2	1

increased as a result of the coaching and bi-weekly meetings					
Overall, I feel the training, coaching and bi-weekly meetings have given me the skills and knowledge I need to continue to deliver Let's Read Fluently if needed, without additional support.	5	4	3	2	1

SECTION 3 – IMPLEMENTATION

3.1) How many students were identified for the programme? [please write number below]

3.2) Do you believe the correct students were identified for the programme?

a) Yes
b) No

IF A, GO TO 3.4

If B: Why do you think children who were identified were not suitable for the programme?

a) They found the lessons too difficult
b) They were unable to focus
c) They were not supported with work at home
d) They were unable to regularly attend lessons
e) They did not require extra support with their literacy
f) Another reason

3.3) Out of the children participating in LRF, how many children do you think were **not** correctly identified for the programme?

a) 1-2
b) 3-5
c) More than 5

3.4) On average how many Let's Read Fluently! sessions did you deliver per week, per group?

a) More than three sessions
b) Three sessions
c) Two sessions
d) One session
e) None

IF A OR B, GO TO QUESTION 3.3

IF C, D OR E: Is there anything that prevented you from delivering more sessions?

a) I had other groups to work with
b) Other lessons are more important for LRF! students to attend
c) I needed more time to plan Let's Read Fluently! sessions
d) Students did not have their practice books
e) Students were absent
f) Something else

3.5) On average, what was the duration of your Let's Read Fluently! sessions, excluding commute/set up time?

a) More than 30 minutes
b) 30 minutes
c) Under 30 minutes

3.6) How many weeks were your Let's Read Fluently! sessions delivered for?

a) 14 weeks or more
b) 13 weeks
c) 12 weeks
d) 10-11 weeks
e) 6-9 weeks
f) 5 weeks or less

3.7) To what extent did you follow the 'I do; we do; you do' approach?

a) I used the 'I do; we do; you do' approach consistently across the sessions
b) I used theI do; we do; you do' approach most of the time
c) I did not use the 'I do' approach
d) I did not use the 'we do' approach
e) I did not use the 'you do' approach
f) I did not use the approach at all

IF A or B, GO TO Q 3.8

IF C, D, E OR F: Why was part of the 'I do; we do; you do' approach not followed?

a) The students did not understand the approach
b) The students were not engaged with the approach
c) The approach took too long to implement
d) I did not feel sufficiently trained in the approach
e) I did not feel that the approach was useful
f) Other

3.8) How many lessons, on average, did you cover between groups?

a) 36 lessons (100%)

b) 27-35 lessons (75%-100%)
c) 18-26lessons (50%-75%)
d) 9-17 lessons (25%-50%)
e) Fewer than 9 lessons

3.9) On average, what proportion of LRF! sessions did you use the practice book in?

a) 36 lessons (100%)
b) 27-35 lessons (75%-100%)
c) 18-26 lessons (50%-75%)
d) 9-17 lessons (25%-50%)
e) Fewer than 9 lessons

If A, GO TO 4.1

IF B, C, D OR E: What prevented you from using the practice book more?

a) I didn't find it helpful
b) I didn't understand how to use it
c) Students didn't like using it
d) Use of the practice book distracted students

SECTION 4 – STUDENT ENGAGEMENT

4.1) To what extent do you agree with the following:

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
Students were engaged with the Let's Read Fluently! sessions	5	4	3	2	1
Students could cope with the level of literacy required in Let's Read Fluently!	5	4	3	2	1
Students were confident in using their practice book in the Resource Room	5	4	3	2	1
Students were confident in using their practice book independently	5	4	3	2	1
Students used their practice book for more than half of lessons	5	4	3	2	1

Students were motivated to take the practice book home	5	4	3	2	1

4.2) How often, if at all, did the students take the practice book home?

a) Every day
b) A few times a week
c) Once a week
d) Once every two weeks
e) Not at all
f) Other

4.3) On average, in how many lessons did you assign homework in the practice book?

a) All lessons
b) More than half the lessons
c) At least one lesson a week
d) Less than one lesson a week
e) I did not assign any homework

4.4) What proportion of students generally completed the homework?

a) All of them
b) Over half
c) Around half
d) Less than half
e) None of them
f) Not sure

SECTION 5 – PERCEIVED OUTCOMES

5.1) In your opinion, have students who have taken part in Let's Read Fluently seen positive effects on their literacy? (Check all that apply, unless D, no positive effects)

a) Improved literacy levels
b) Increased motivation to practice reading
c) Increased confidence to practice reading
d) No positive effects on literacy seen

5.2) In your opinion, have there been other positive effects on students who have taken part in Let's Read Fluently? (Check all that apply, unless G, no positive effects)

a) Improvements in other academic capabilities

b) Increased confidence in other areas of learning
c) Improved attention or behaviour in other classes
d) Improved social skills
e) Improved attendance at school
f) Other
g) No other positive effects seen

5.3) In your opinion, have there been any negative effects on students who have taken part in Let's Read Fluently? (Check all that apply, unless G, no negative effects)

a) Reduced confidence in learning literacy
b) Reduced confidence in other areas of learning
c) Reduced attention/ more disruption in other classes
d) Increased bullying/stigma from other students in the class
e) Increased absences from school
f) Other
g) No other negative effects seen

5.4) In your opinion, have there been any negative consequences on you as a result of the Let's Read Fluently! programme? (Check all that apply, unless E, no negative effects)

a) Less time to focus on other students
b) Increased workload
c) More difficulty in interactions with parents
d) Other
e) No negative consequences

5.5) In your opinion, have there been any positive consequences on you as a result of the Let's Read Fluently! programme? (Check all that apply, unless E no positive consequences)

a) Better understanding of different approaches to teaching early literacy
b) Better understanding of how to identify students with reading difficulties
c) More able to provide appropriate literacy support to students
d) More engagement with parents about students' progress in reading
e) Other
f) No positive consequences

5.6) In your opinion, have there been any negative consequences on other students, not eligible for the LRF! programme, as a result of implementation? (Check all that apply, unless G no negative effects)

a) Increased levels of disruption
b) Reduced literacy progress
c) Reduced support that they would usually receive from you, the resource room teacher
d) Increased bullying/stigma from other students
e) Increased absences
f) Other
g) No negative consequence

5.7) In your opinion, have there been any negative effects on parents as a result of the LRF! programme? (Check all that apply unless G, no negative effects)

a) Added concern about their child's literacy levels
b) More time needed to give support to their children at home
c) Concern about their child's time in other lessons
d) Concern about the programme itself
e) Other
f) Unsure
g) None

5.8) At the end of the programme, will your school continue to implement the skills and approaches learnt from the Let's Read Fluently! programme? (Check all that apply)

a) My school will continue to use 'I do, you do, we do' approaches
b) My school will continue to use the practice book
c) My school will continue, but I'm unsure with which elements
d) Other
e) My school will not use Let's Read Fluently approaches after this semester
f) Unsure

END OF SURVEY SCRIPT That is the end of the survey, thank you for your time in completing it. If you have any questions about the survey or the data collected, you can speak to your Let's Read Fluently! coach, who can put you in touch with the research team.

Classroom teacher end of LRF! programme survey

INTRODUCTION SCRIPT

Hello,

My name is [Name] from [Organisation], calling on behalf of the Queen Rania Foundation. I am calling to conduct a short survey with you about how you have found different elements of the Let's Read Fluently! Programme, including the implementation of the sessions and perceived outcomes of the programme. We understand that you have not been delivering the sessions but have had contact with students and resource room teachers who have.

We are contacting all the classroom teachers enrolled in the 2nd pilot of Let's Read Fluently! to ask them about their experiences. The survey should take around 10 minutes to complete and I will read you the questions over the phone and enter your answers into the computer. The data will be anonymised and shared securely with the researchers analysing the project on behalf of the Queen Rania Foundation, and you will not be identified in any analysis or reports that are shared with schools. These reports will include group results from this survey alongside data from other surveys and focus groups. Your individual responses will not be shared with your school, or with the Let's Read Fluently! trainers and coaches.

You do not have to take part, and you can change your mind during or after the survey, and your data will be removed. Are you happy to continue with the survey now?

Note for interviewer (please read out question and all response options, rather than just the question)

SECTION 1 – BACKGROUND

1.1) Teacher code:

1.2) In which region is your school?

a) North of Jordan: Jarash, Ajloun
b) South of Jordan: Karak
c) Middle of Jordan excluding Amman: Madaba, Balqa, Zarqa
d) Amman

1.3) How long have you been a teacher for?

a) Less than a year
b) 1-2 years
c) 3-5 years
d) 6-10 years
e) More than 10 years

SECTION 2 – IMPLEMENTATION

2.1) How many times a week were Let's Read Fluently! students sent to the resource room for their reading sessions?

a) More than three times
b) Three times
c) Twice
d) Once
e) None
f) Unsure

2.2) On average, how long were these students out for the sessions?

a) More than 30 minutes
b) 30 minutes
c) Under 30 minutes
d) Unsure

2.3) How many weeks were the Let's Read Fluently! sessions delivered for?

a) 14 weeks or more
b) 13 weeks
c) 12 weeks
d) 10-11 weeks
e) 5-9 weeks
f) Less than 5 weeks

g) Unsure

2.4) Which lessons did students miss to attend Let's Read Fluently! sessions? (select all that apply)

a) Islamic Religion
b) Arabic
c) English
d) Maths
e) Science
f) Social studies
g) Art
h) PE
i) Other

2.7) To what extent have Let's Read Fluently! Sessions disrupted/not disrupted normal learning for students?

a) They have not disrupted normal learning at all
b) They have disrupted normal learning slightly
c) They have disrupted normal learning moderately
d) They have disrupted normal learning significantly
e) Unsure

SECTION 3 – IDENTIFYING STUDENTS FOR THE LET'S READ FLUENTLY! PROGRAMME

3.1) Were there any students from your class who were not identified for the Let's Read Fluently! programme, who you feel would have benefitted from it?

a) Yes
b) No
c) Unsure

If B OR C, go to 3.2

If A: How many students from your class do you think were **not** correctly identified for the Let's Read Fluently! programme ?

a) 1-2
b) 3-5
c) More than 5

3.2) Were there any students from your class that were identified for the Let's Read Fluently! Programme, who you feel did NOT require extra support with literacy?

a) Yes
b) No
c) Unsure

If B or C, go to 3.3

If A: How many students did not require extra support with literacy but were identified for the Let's Read Fluently! programme?

a) 1-2
b) 3-5
c) More than 5

SECTION 4 – STUDENT ENGAGEMENT

4.1) From what you have seen in your school, to what extent do you agree with the following?

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree	Don't know
Students were motivated to attend Let's Read Fluently! sessions	5	4	3	2	1	0
Students were motivated to take the practice book home	5	4	3	2	1	0

SECTION 5 – PARENTAL ENGAGEMENT

5.1) From what you have seen in your school, to what extent do you agree with the following?

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree	Don't know
Parents of students receiving Let's Read Fluently! had a positive perception of the programme	5	4	3	2	1	0
Parents of students receiving Let's Read Fluently! were supportive of their children i.e., encouraged the completion of homework	5	4	3	2	1	0

Parents of students not receiving the LRF! teaching had a positive perception of the programme	5	4	3	2	1	0

SECTION 6 – PERCEIVED OUTCOMES

6.1) In your opinion, have students who have taken part in Let's Read Fluently seen positive effects on their literacy? (Check all that apply, unless D, no positive effects)

a) Improved literacy levels
b) Increased motivation to practice reading
c) Increased confidence to practice reading
d) No positive effects on literacy seen
e) Unsure

6.2) In your opinion, have there been other positive effects on students who have taken part in Let's Read Fluently? (Check all that apply, unless G, no positive effects)

a) Improvements in other academic capabilities
b) Increased confidence in other areas of learning
c) Improved attention or behaviour in other classes
d) Improved social skills
e) Improved attendance at school
f) Other
g) No other positive effects seen

6.3) In your opinion, have there been any negative effects on students who have taken part in Let's Read Fluently? (Check all that apply, unless G, no negative effects)

a) Reduced confidence in learning literacy
b) Reduced confidence in other areas of learning
c) Reduced attention/ more disruption in other classes
d) Increased bullying/stigma from other students in the class
e) Increased absences from school
f) Other
g) No other negative effects seen

6.4) In your opinion, have there been any negative consequences on you as a result of the Let's Read Fluently! programme? (Check all that apply, unless F, no negative effects)

a) Less time to focus on current literacy curriculum
b) Less time to plan or deliver literacy lessons
c) More difficulty in interactions with parents
d) More disruptions during teaching time to allow students to attend the Resource Room
e) Other

f) No negative consequences

6.5) In your opinion, have there been any positive consequences on you as a result of the Let's Read Fluently! programme? (Check all that apply, unless H no positive consequences)

a) Better understanding of different approaches to teaching early literacy
b) Better understanding of how to identify students with reading difficulties
c) More time to plan or deliver literacy lessons
d) Less time spent on behaviour management during lessons
e) More engagement with resource room teachers about students' progress in reading
f) More engagement with parents about students' progress in reading
g) Other
h) No positive consequences

6.6) In your opinion, have there been any negative consequences on other students, not eligible for the LRF! programme, as a result of implementation? (Check all that apply, unless G no negative effects)

a) Increased levels of disruption
b) Reduced literacy progress
c) Reduced support they would usually receive from the resource room teacher
d) Increased bullying/stigma from other students
e) Increased absences
f) Other
g) No negative consequence

6.7) In your opinion, have there been any negative effects on parents as a result of the LRF! programme? (Check all that apply unless G, no negative effects)

a) Added concern about their child's literacy levels
b) More time needed to give support to their children at home
c) Concern about their child's time in other lessons
d) Concern about the programme itself
e) Other
f) Unsure
g) None

6. 8) To what extent do you agree that resource room teachers are better able to deliver effective reading sessions to students who are in need of additional literacy support, as a result of Let's Read Fluently?

a) Strongly agree
b) Agree

c) Neither agree nor disagree
d) Disagree
e) Strongly disagree
f) Unsure

6. 9) At the end of the programme, will your school continue to implement the skills and approaches learnt from the Let's Read Fluently! programme? (Check all that apply)
a) My school will continue to use 'I do, you do, we do' approaches
b) My school will continue to use the practice book
c) My school will continue, but I'm unsure with which elements
d) Other
e) My school will not use Let's Read Fluently approaches after this semester
f) Unsure

END OF SURVEY SCRIPT That is the end of the survey, thank you for your time in completing it. If you have any questions about the survey or the data collected, you can speak to your school's Let's Read Fluently! representatives, who can put you in touch with the research team.

Appendix 6: IPE focus group discussion guides

Focus Group Discussion Guide: Resource Room Teachers

(C/U model group)

'Let's Read Fluently!' (LRF) is designed to address issues with literacy among Arabic students. The intervention centres around a student practice book designed to reflect evidence about how Arabic reading fluency is best acquired.

Developed from the first pilot, LRF follows the Literacy Catch-Up (C/U) model which runs for 14 weeks, equating to one semester. C/U targets the lowest-achieving students in Grades 2-3 using the Early Grade Reading and Mathematics Program (RAMP) and Princess Taghrid Institute (PTI) tool. Prior to teaching, esource oom teachers receive two days of training. Subsequently esource oom teachers receive up to three follow-up coaching sessions with QRTA coaches, and are able to engage with bi-weekly meetings for support during the intervention. Teachers adopt an 'I do', 'we do', 'you do' pedagogical approach to learning, using the practice book. Lessons are delivered via small group tuition in three weekly 30-minute sessions.

The Queen Rania Foundation has commissioned a consortium which includes NatCen and Integrated to carry out a 2nd pilot evaluation of LRF. The aim of the evaluation is to explore delivery, understand the feasibility of different evaluation methods, and inform the planning of an efficacy trial. In particular, it aims to understand if the changes made to the student inclusion/exclusion criteria and to the training and support for esource oom teachers has addressed issues reported in the previous pilot, i.e., whether the correct students have been included in the intervention and the teachers feel able to effectively implement it.

In this 60-75 minute focus group with esource oom teachers, we will discuss:

- Resource oom teachers' understanding and perceptions of the programme
- Resource oom teachers' experiences of implementation to date, including barriers and enablers to delivery

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 - Hi, my name is [NAME]. Thank you for agreeing to take part in this focus group discussion with me today.

- Explanation of role: from an organisation called Integrated International. We have been commissioned by the Queen Rania Foundation (with NatCen) to conduct an independent evaluation of the pilot programme 'Let's Read Fluently!'. For the rest of this discussion, we will refer to it as LRF.

- Explanation of research – in addition to speaking with you and teachers at other participating schools across Jordan, we are also speaking to coaches who are delivering the coaching sessions as part of LRF, who have also conducted observations.

- During today's focus group discussion, we will discuss:
    - Your understanding of the programme
    - Your experiences of delivering the programme to date
    - Your insights into perceived impacts of the programme

- Your participation is voluntary – you can choose not to discuss any issue and can withdraw from the focus group at any point, without giving a reason.

- The information you provide will be used to write a findings report at the end of this school year which will be published on the QRF and NatCen websites.

- We are conducting this research independently and all information shared will be treated confidentially.
    - 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 your role and the fairly small sample of schools, others familiar with the programme may recognise you from what you have said.
    - At the end of the focus group, please let us know if you 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. This audio recording will be shared with the evaluators, who will have it translated and transcribed for analysis. Check that this is ok.

- The discussion today will last around 1 hour to 1 hour and 15 minutes, and we can take a break if required.

- FGD rules: Please respect everyone else's opinions; we value hearing from a range of perspectives. Only one person should talk at a time, and we ask that anything shared within this group today is not shared more widely.

- Do you have any questions before we begin?

Date of FGD:

School/s:

1. Participant introduction [5 minutes]

Aim: To ease participants into group discussion

Resource room teacher role

Please can you start by sharing with us what the role of a Resource Room teacher is?

- Is this generally the same across schools?

Understanding of programme

- How would you describe the LRF! programme?
  - What are the key aims?
  - What are/were your initial expectations for the programme? (PROBE: what did you initially think that LRF! may achieve?)

2. Reflections on LRF! in practice [25 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 student engagement with LRF! and the practice book.

Implementing the LRF! model in the classroom [5 mins]

- What did you find were the main differences between LRF! and your usual literacy teaching? (e.g. repeated practice, modeling, students' independent practice/reading. PROBE: what are

the main positives of LRF! compared to usual practice? What are the main negatives of LRF! compared to usual practice?)

- Have you encountered any issues with the practicality of delivery? (PROBE: locations of delivery, number of sessions delivered, availability of resources)

Students' engagement [10 mins]

- How have students responded to the LRF! approach in the classroom? (PROBE: How engaged have students been in LRF! lessons? Has this been different to usual practice?)
- What, if anything, about the LRF! programme has encouraged student engagement? (PROBE: Which elements of the programme do students tend to engage with the most? E.g. the teaching materials, the teaching methods, working in a small group, ability to practice at home?)
- What, if anything, about the LRF! programme has discouraged student engagement? (PROBE: Which elements of the programme do students tend to disengage with the most? E.g. the teaching materials, the teaching methods, working in a small group, ability to practice at home?). What did you do or would you like to do to improve engagement?
- Do you think the most students you are teaching in the Resource Room are the students that LRF! is designed for? (PROBE: Why do you think this?)
- Were there any negative consequences that you are aware of, that happened as a result of using the tools to identify students for the programme? (PROBE: For instance, with students or parents? What were these? Why do you think this happened, and how do you think this could be avoided?)

Using the practice book [10 mins]

- How have you found introducing the practice book to the class? (PROBE: Were there any difficulties when starting with the book?)
- How often is the practice book used in class? (PROBE: Could it have been used more? Why?)
- In your experience, is it achievable for students to use their practice book independently? (PROBE: If not, why do you think this is?)

- What about the practice book works well? (PROBE: What kind of impact have you seen it have on student's literacy levels?)

- Are there ways in which the practice book could be improved for use in the classroom? (PROBE: what effect do you think these changes would have on student engagement or outcomes?)

- Do you feel that the practice book has changed your approach to teaching literacy? (PROBE: Are there differences between the practice book and previous materials you used for teaching? Does it make literacy lessons any easier or harder to teach?)

- Are you finding that the practice book has been used at home?

    - **If yes:** What is helping students to use it at home? (PROBE: their own levels of motivation; support from parents; support from resource room teachers; access to the QR codes?)

    - **If no:** Can you share some of the reasons for this?

1. Perceived impacts [10 minutes]

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

- What effects have you seen from the LRF! programme on:

    - Students' reading ability? (PROBE: any issues with students mouthing words/ mirroring others rather than reading independently?)

    - Students' confidence with reading (e.g., attempting more reading tasks without asking for support, trying different types of reading tasks, greater participation in literacy lessons?)

    - Students' engagement with literacy learning? (PROBE: Are they motivated to participate in sessions, use the practice book, or take the practice book home?)

- Has there been a group of learners for whom LRF! has been more or less effective?

- Do you feel better equipped to deliver effective reading sessions having received the training and coaching and delivered the programme? (PROBE: Do you feel more confident? Have you developed more skills for teaching literacy?)

- Have you seen any unanticipated consequences of LRF, whether these be positive or negative? (PROBE: This could be related to learning or social effects on students who are participating in the programme, other students, resource room teachers, or parents?)

- Are you / your school likely to continue using LRF! to support children with literacy needs in the future? (PROBE: Why do you think this? If you do continue using it, will it be in the same format or would you make changes or combine it with other methods?)

2. Reflections on training and coaching [15 minutes]

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

Reflections on training

- It's likely that you've already answered some questions on the training you received from QRTA in the post-training teacher survey. We'd like to ask some more questions to get a more in-depth understanding.

  - Please can you share which parts of the training were particularly useful or less useful for your work? (PROBE: Why was this useful / What did it help you do?)
  - What helped you to engage in the training? Why? (PROBE: Relationship with those that delivered the training; practical arrangements such as timing, location or length of training; materials given etc.)
  - What made it harder to engage in the training? Why? (PROBE: Relationship with trainers; practical arrangements such as timing, location or length of training; materials given etc.)
  - Are there any parts of the training that you think should be changed for the future? (PROBE: What are these? Why?)

Feedback on QRTA's LRF! coaching sessions

- What helped you to engage in the coaches' sessions? Why? (PROBE: Did any of the following help? Your relationship with the coaches, practical arrangements such as timing, location or length of coaching sessions, materials given etc.)

- What made it harder to engage with the coaching? Why? (PROBE: Did any of the following hinder? Your relationship with the coaches, practical arrangements such as timing, location or length of training, materials given etc.)

- Are there any parts of the coaching that you think should be changed for the future? (PROBE: What are these? Why?)

- How useful was the teacher manual? (PROBE: were there specific parts that were more useful than others?)

- Can anything else be done to support Resource Room teachers during LRF! delivery?

5. Closing [5 minutes]

- Do you have any further comments about the LRF! programme that you think would be useful for us to know?

[Thanks and close of focus group discussion].

Focus Group Discussion Guide: Classroom Teachers and Resource Room Teachers
(Usual practice group)

'Let's Read Fluently!' (LRF) is designed to address issues with literacy among Arabic students. The intervention centres around a student practice book designed to reflect evidence about how Arabic reading fluency is best acquired.

Developed from the first pilot, LRF follows the Literacy Catch-Up (C/U) model which runs for 14 weeks equating to one semester. C/U targets the lowest-achieving students in Grades 2-3 using the Early Grade Reading and Mathematics Program (RAMP) and Princess Taghrid Institute (PTI) tool. Prior to teaching, esource oom teachers receive two days of training. Subsequently esource oom teachers receive up to three follow-up coaching sessions with QRTA coaches, and are able to engage with bi-weekly meetings for support during the intervention. Teachers adopt an 'I do', 'we do', 'you do' pedagogical approach to learning, using the practice book. Lessons are delivered via small group tuition in three weekly 30-minute sessions. In this model, the usual practice group of classroom and resource room teachers resume 'teaching-as-usual', without implementation of the LRF model.

The Queen Rania Foundation has commissioned a consortium which includes NatCen and Integrated to carry out a 2nd 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 efficacy trial.

In this 35–40-minute focus group with classroom and esource oom teachers, we will discuss:

Classroom and esource oom teachers' reflections on experiences of teaching literacy

Classroom and esource oom teachers' understanding of enablers and barriers to reading, including the achievement of student's literacy outcomes

Classroom and esource oom teachers' reflections on potential improvements to the current way of teaching literacy

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 – Hi, my name is [NAME]. Thank you for agreeing to take part in this focus group discussion with me today.

Explanation of role: from an organisation called Integrated International. We have been commissioned by the Queen Rania Foundation (with NatCen) to conduct an independent evaluation of the pilot programme 'Let's Read Fluently!'. For the rest of this discussion, we will refer to it as LRF.

Explanation of research – in addition to speaking with you, and classroom and Resource Room teachers at other participating schools, we are also speaking to coaches who are delivering the coaching sessions as part of LRF.

During today's focus group discussion, we will discuss:

- Your reflections on teaching students to read
- Your understanding of enablers and barriers to improving students' literacy levels
- Your insights into how literacy teaching could be improved

Your participation is voluntary – you can choose not to discuss any issue and can withdraw from the focus group at any point, without giving a reason.

The information you provide will be used to write a findings report at the end of this school year which will be published on the QRF and NatCen websites.

We are conducting this research independently and all information shared will be treated confidentially.

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 your role and the fairly small sample of schools, others familiar with the programme may recognise you from what you have said.

At the end of the focus group, please let us know if you 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. This audio recording will be shared with the evaluators, who will have it translated and transcribed for analysis. Check that this is ok.

The discussion today will last up to 35-40 minutes, and we can take a break if required.

- FGD rules: Please respect everyone else's opinions; we value hearing from a range of perspectives. Only one person should talk at a time, and we ask that anything shared within this group today is not shared more widely.

Do you have any questions before we begin?

Date of FGD:

School/s:

Participant background [under 5 minutes]
Aim: To ease participants into group discussion

Role/responsibilities

Please can you start by sharing with us what the role of a Resource Room teacher / classroom teacher is in your school?

Reflections on teaching students to read [15 minutes]
Aim: To explore experiences of teaching students to read, and to ascertain a better understanding of teaching-as-usual delivery

Typical approach [5 mins]

How are students in grades 2-3 typically taught to read in your school?

PROBE: Are there any particular exercises you use?

PROBE: Why do you teach this way?

Have you received training and support to help you teach students how to read?

IF YES - What has this looked like?

IF YES - Has this been effective in supporting student's learning?

IF NO - Would you find this useful in supporting teaching?

Supporting students with difficulty reading [10 mins]

How are students who have difficulty reading identified in your school? (PROBE: do you use a particular assessment to identify children with reading difficulties?)

In your experience, to what extent is it easy for teachers to identify the students who are struggling to read in comparison to their peers?

How is teaching adjusted for students who have difficulty reading?

PROBE: Is the resource room and resource room teacher used for those who have difficulty reading? (If yes, how often, how long, how many students use this facility?)

- Does this support help these students with their literacy? (PROBE: If not, why do you think this is?)
- What training do you receive to help children who have reading difficulties specifically?

Reflections on enablers and barriers to reading [15 minutes]

Aim: To better understand what enablers and barriers are common with teaching to read, including homework tasks.

Enablers and barriers to teaching to read [up to 10 mins]

How engaged are students with the current way of teaching to read? (PROBE: Which elements do students engage with the most? [use examples from their earlier answers about teaching to read])

What external factors help or provide barriers to teaching your students to read?

PROBES:

Time spent reading / not reading

Students having access / no access to books and reading materials

Parental/at home support / no parental/at home support

Anything else?

Engagement with homework [5 mins]

Do you set homework when teaching students to read? (PROBE: Why/ why not? How often?)

Are students able to complete homework independently, or with parental support, if needed? (PROBE: If no, what are the barriers to them completing homework?)

Potential improvements [less than 5 minutes]

Aim: To explore what changes could be made to better support students when learning to read.

Is there anything we have not covered today that you think is particularly effective when teaching students to read? (PROBE: class size, use of supporting materials)

Any final comments?

[Thanks and close of focus group discussion]

Focus Group Discussion Guide: Coaches

'Let's Read Fluently!' (LRF) is designed to address issues with literacy among Arabic students. The intervention centres around a student practice book designed to reflect evidence about how Arabic reading fluency is best acquired.

Developed from the first pilot, LRF follows the Literacy Catch-Up (C/U) model which runs for 14 weeks, equating to one semester. C/U targets the lowest-achieving students in Grades 2-3 using the Early Grade Reading and Mathematics Program (RAMP) and Princess Taghrid Institute (PTI) tool. Coaches engage in training and coaching sessions and additional bi-weekly meetings to support esource oom teacher's implementation of LRF in the resource room. Teachers adopt a 'I do', 'we do', 'you do' pedagogical approach to learning, using the practice book. Lessons are delivered via small group tuition in three weekly 30-minute sessions.

The Queen Rania Foundation has commissioned a consortium which includes NatCen and Integrated to carry out a 2nd pilot evaluation of LRF. The aim of the evaluation is to explore delivery, understand the feasibility of different evaluation methods and inform the planning of an efficacy trial. In particular, it aims to understand if the changes made to the student inclusion/exclusion criteria and to the training and support for esource oom teachers has addressed issues reported in the previous pilot i.e., whether the correct students have been included in the intervention and the teachers feel able to effectively implement it.

In this 75-minute focus group with QRTA coaches, we will discuss:

Coaches' perceptions of the programme

Coaches' observations of implementation to date, including barriers and enablers to delivery

Coaches' insights into perceived impacts to date

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 – Hi, my name is [NAME]. Thank you for agreeing to take part in this focus group discussion with me today.

- Explanation of role: from an organisation called Integrated International. We have been commissioned by the Queen Rania Foundation (with NatCen) to conduct an independent

evaluation of the pilot programme 'Let's Read Fluently!'. For the rest of this discussion, we will refer to it as LRF.

- Explanation of research – in addition to speaking with you, we are also speaking to Resource Room teachers at other participating schools, some who are delivering LRF, and some who are delivering usual teaching.

- During today's focus group discussion, we will discuss:
    - Your perceptions of the programme
    - Your observations of the implementation of LRF! to date
    - Your insights into perceived impacts

- Your participation is voluntary – you can choose not to discuss any issue and can withdraw from the focus group at any point, without giving a reason.

- The information you provide will be used to write a findings report at the end of this school year which will be published on the QRF and NatCen website.

- We are conducting this research independently and all information shared will be treated confidentially.
    - 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 your role and the fairly small sample of schools, others familiar with the programme may recognise you from what you have said.
    - At the end of the focus group, please let us know if you 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. This audio recording will be shared with the evaluators, who will have it translated and transcribed for analysis. Check that this is ok.

- The discussion today will last up to 75 minutes, and we can take a break if required.

FGD rules: Please respect everyone else's opinions; we value hearing from a range of perspectives. Only one person should talk at a time, and we ask that anything shared within this group today is not shared more widely.

- Do you have any questions before we begin?

Date of FGD:

School/s:

---

1. Participant background [5 minutes]

Aim: To ease participant into conversation, explore role and responsibilities, their understanding of the programme.

Role/responsibilities

- Please can you start by sharing with us what the role of a QRTA coach is for LRF?

Understanding of programme

- How would you describe the LRF! programme?

  What are the key aims?

  What are/were your initial expectations for the programme? (PROBE: What did you initially think that LRF! may achieve?)

Reflections on training [10 minutes]

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

Feedback on training that coaches received [5 mins]

- What was the format and delivery of training for coaches?
  How long was the training?

Were there specific elements of the training which you found more or less helpful?

Were there any gaps or issues with the training? (PROBE: Issues with the people who delivered the training, practical arrangements, materials)

Has the training that you received increased your feelings of confidence in delivering LRF! to teachers? (PROBE: How/ Why?)

Reflection on training that coaches delivered to Resource Room teachers [5 mins]

When thinking about the training that you delivered to Resource Room teachers, what do you think worked well? (PROBE: length of training; materials, any teaching strategies etc.)

What worked less well? (PROBE: length of training; materials, any teaching strategies etc.)

Are there any parts of the training that you think should be changed for the future? (PROBE: What are these? Why?)

Reflections on coaching [10 minutes]

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

Reflection on coaching sessions for Resource Room teachers

For the coaching sessions that you delivered, what do you think has worked well? (PROBE: practical arrangements such as timing, location or length of coaching; materials etc.)

What worked less well? (PROBE: practical arrangements such as timing, location or length of coaching; amount of coaching sessions you were expected to deliver etc.)

Has coaching been a useful addition to the initial training session?

Were the majority of coaching sessions attended?

If no: What were the reasons for non-attendance, that you are aware of?

Reflection on bi-weekly meetings

- What was the format and delivery of the bi-weekly meetings for Resource Room teachers? (PROBE: location; duration; frequency; regularity)

- Do you think that the bi-weekly meetings have been useful for Resource Room teachers?
  What do you think has worked well?
  What do you think has worked less well?

Were the majority of bi-weekly meetings attended?

If no: What were the reasons for non-attendance, that you are aware of?

- Can anything else be done to support Resource Room teachers during delivery?

Reflections on LRF! in practice [25 minutes]

Aim: To explore how LRF! was implemented in the resource room from the perspective of coaches. Also, to explore student engagement with LRF! and the practice book.

Implementation of the LRF! model in the resource room [10 mins]

- What are the main differences you've noticed between LRF! and usual literacy teaching? (PROBE: what are the main positives of LRF! compared to usual practice? What are the main negatives of LRF! compared to usual practice?)

- From your sessions and meetings with Resource Room teachers, has the frequency and regularity of LRF! sessions been implemented in the resource room as intended? (PROBE: have sessions been 30 mins long, and have they been delivered 3 times a week up to now?)

- From your observations of sessions, how closely have LRF! approaches (I do, we do, you do) been followed in the resource room?

How successful do you feel the implementation of the LRF! programme has been in the resource room? (PROBE: Have you witnessed any issues with implementation in the resource room? e.g., pace of content, materials, engagement, etc.)

How successful do you feel the implementation of the LRF! programme has been in terms of working with schools to a) identify students for the intervention, b) deliver the intervention? (PROBE: have you witnessed any issues with implementation e.g., obtaining students' RAMP scores from teachers, the interactions between PTI and schools, organising training/coaching sessions, etc.)

Student engagement [10 mins]

- According to your observations of classes, how have students responded to the LRF! approach in the resource room? (PROBE: How engaged have students been in LRF! lessons?)

- What, if anything, about the LRF! programme has encouraged student engagement? (PROBE: Which elements of the programme do students tend to engage with the most?)

- What, if anything, about the programme has discouraged student engagement? (PROBE: Which elements of the programme do students tend to disengage with the most? How could these be changed to improve engagement?)

- Do you think the students that are being taught in the Resource Room are the students that LRF! is designed for? (PROBE: Why do you think this?)

- Were there any issues that you are aware of, that happened as a result of using the tools to identify students in need of additional support? (PROBE: For instance, with students or parents? What were these? Why do you think this is happened?)

Using the practice book [5 mins]

How have you observed the practice book being used in the resource room? (PROBE: How often have you seen it be used/ Have you observed any difficulties?)

In your experience, is it achievable for students to use their practice book independently? (PROBE: If not, why do you think this is?)

- What about the practice book works well? (PROBE: What kind of impact have you seen it have on student's literacy levels?)

- Are there ways which the practice book could be improved for use in the resource room? (PROBE: i.e., if so, how?)

Perceived impacts [10 minutes]

Aim: To explore perceived benefits of programme for students and teachers.

- What effects have you seen from the LRF! programme on:
  Students' reading ability? (PROBE: have there been any issues with students mouthing words/ mirroring others rather than reading independently?)

  Students' confidence with reading? (e.g., attempting more reading tasks without asking for support, trying different types of reading tasks, greater participation in literacy lessons?)

  Students' engagement with literacy learning? (PROBE: Are they motivated to participate in sessions, use the practice book, or take the practice book home?)

- Has there been a group of learners for whom LRF! has been more or less effective? (PROBE: Why do you think this is?)

- As a result of the training and coaching, do teachers seem to have increased skills and confidence to effectively deliver the LRF! programme? (PROBE: Why do you think this? What examples have you seen?)

- Have you seen any unanticipated consequences of LRF, whether these be positive or negative? (PROBE: This could be related to students who are participating in the programme, other students, Resource Room or classroom teachers, or parents?)

Closing

Do you have any further comments about the LRF! programme that you think would be useful for us to know?

[Thanks and close of focus group discussion].

Appendix 7: Steps during the PTI assessment process

During the PTI assessment, four key aspects were evaluated by a committee of lead therapists to determine a child's eligibility for the LRF! C\U model. These aspects are detailed below:

1. **Parents report:**

A phone call screening assessment with parents with the purpose of:

- gathering consent to assess the child.
- basic information is collected (name, birthdate, nationality, phone number, grade, school).
- understanding of the family structure and the student's position in the family.
- Assessment of maternal pregnancy and childbirth conditions.
- evaluation of the child's developmental, motor, and language milestones. - Inquire about the child's academic performance and the parent's assessment (good/average/below average).

2. **Student Assessment Process - Stage 1: Learning Difficulties Test**

This test examines for auditory discrimination, comprehension of same-meaning sentences, following instructions, understanding others' speech, and understanding pronouns. It also evaluates difficulties in understanding and comprehension, speech capabilities, and reading and writing skills. Below is a more detailed description of its items:

1. Listening Difficulties Diagnosis, including:
   - Auditory Discrimination Test.
   - Comprehension of Sentences with similar or different linguistic structures carrying the same or
   different meanings.
   - Multi-step Instruction Comprehension Test.
   - Understanding Others' Speech.
   - Understanding Connected and Concealed Pronouns in Sentences.
2. Comprehension and Understanding Difficulties Diagnosis, including:

- Ability to understand the cause-and-effect relationship.
- Ability to draw important conclusions.
- Problem-solving ability.

4. Reading Difficulties Diagnosis, including:
   - Reading ability, fluency, and error-free reading.
   - Ability to analyze words into their letters.
   - Ability to assemble letters into words.
5. Writing Difficulties Diagnosis, including:
   - Writing ability, fluency, and error-free writing.
   - Ability to spell words and sentences.

3. Student Assessment Process - Stage 2: Language Development Test

- Evaluate linguistic, receptive, and expressive skills.
- Examine visual and auditory discrimination.
- Assess visual and auditory perception.
- Evaluate visual and auditory memory.

Below is a detailed description what this entails:

1. Visual Response Evaluation:

- Visual Discrimination: The ability to successfully distinguish between images from a set of options.
- Visual Stability: The capability to differentiate between similar images within a set.
- Visual Closure: The capacity to perceive the complete shape or object when presented with partial information.
- Visual Memory: The skill of remembering images after a brief exposure.
- Spatial Relationships: Understanding the positioning of objects in space, including distinguishing between right and left.
- Visual Discrimination of Shape and Background: The ability to focus on a specific object while ignoring other visual stimuli.
- Visual Sequencing: The aptitude to accurately recall and arrange a sequence of visual stimuli.

2. Auditory Response Evaluation:

- Attention to the Sound Source: Demonstrating attentiveness and responsiveness to auditory stimuli.
- Distinguishing Between Sounds: Successfully recognizing and differentiating between various auditory stimuli.

- Linking Sound with its Source: Appropriately associating sounds with their corresponding sources or actions.
- Auditory Memory Skills: Demonstrating the ability to memorize and recall auditory information, including names and sequences of items.

3. Speech Organs Evaluation:

- Face: Assessment indicates a symmetrical appearance of the face.
- Teeth: Evaluation reveals teeth that are functionally and structurally sound.
- Jaws: Whether both upper and lower jaws are determined to be functionally and structurally sound.
- Lips: The assessment notes symmetrical lips that are both functionally and structurally sound.
- Tongue: Whether both the structure and function of the tongue are deemed to be sound.
- Palate: Examination finds the palate to be within the normal range of appearance and function.

4. Language Skills Evaluation:

- Receptive Language:
- Observation Ability: the child's capacity to observe stimuli in various directions.
- Responsiveness: the child's ability to respond appropriately when called by name.
- Vocabulary Differentiation: the child's skill in differentiating common vocabulary items within predefined categories.
- Recognition of Colors and Shapes: the child's ability to distinguish between basic colors and geometric shapes.
- Following Instructions: evaluates the child's capability to follow simple instructions consisting of one or more requests, aligning with their chronological age.
- Expressive Language:
- Timely Expressive Language Skills: whether there is any delay in the child's expressive language skills.
- Naming and Description Abilities: the child's capacity to name images of vocabulary items within predefined categories and answer questions related to personal information.
- Description Skills: the child's ability to describe simple activities, events, facial and body parts, clothing items, qualities, and internal events and emotions according to chronological age.

- Tool Identification: It examines the child's capability to identify tools by naming their functions.
- Speech Evaluation:
- Speech Pronunciation: determines if there are any speech pronunciation issues present.
- Speech Fluency: the individual's speech fluency to identify any potential problems.
- Voice Examination: examines the individual's voice for any issues or abnormalities.

4. Therapist observations during assessment, diagnosis and recommendations

These could include observations of:

1. Language proficiency and expression.
1. Vocabulary acquisition and usage.
2. Reading comprehension.
3. Writing ability and proficiency.
4. Fine motor skills, including handwriting.
5. Attention span and concentration.
6. Visual communication skills.
7. Social interaction and interpersonal skills.
8. Building rapport and establishing relationships with the assessors.