

**Let's Read Fluently!**  
**Pilot Evaluation Protocol**  
**Evaluator: NatCen**  
**Principal investigator(s): Enes Duysak**

---

PROJECT TITLE	Let's Read Fluently! Catch-Up model Pilot Evaluation
DEVELOPER (INSTITUTION)	Queen Rania Teacher Academy (QRTA)
EVALUATOR (INSTITUTION)	National Centre for Social Research (NatCen) and Integrated International
PRINCIPAL INVESTIGATOR(S)	Enes Duysak
STUDY PLAN AUTHOR(S)	Enes Duysak, Hayley Leonard, Abbi Rennick, Charlotte Bessant, Rebecca Parker, Anjhana Damodaran
PUPIL AGE RANGE AND KEY STAGE	Ages 7 to 9 (Jordanian Grades 2 to 3)
NUMBER OF SCHOOLS/ SETTINGS	16
NUMBER OF PUPILS	164

**Study plan version history**

VERSION	DATE	REASON FOR REVISION
1.1		
1.0 [original]		

## Study rationale and background

### *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 primary school-aged children reaching oral reading fluency benchmarks of 46 words per minute was low, at 13.2% and 19.1%, respectively (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 pupils 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 enrol in school, the reality is that they are tasked with absorbing it alongside developing their more colloquial literacy at home. This diglossia<sup>1</sup> 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 (Abadzi, 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<sup>2</sup>, 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 pupils 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 children 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 pupils in a class, while the C/U model provides targeted support for selected pupils who are identified as falling behind their classmates in MSA Arabic literacy.

Approaches to early literacy teaching similar to the LRF! model in Cambodia, the Gambia and Egypt have shown evidence of promise. Similarities in approaches 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 receiving teaching in an intervention classroom, namely, being able to read more letters and making fewer errors than their peers. This suggests the LRF! model as a whole may have a positive impact on Arabic reading fluency among early grade pupils (Eckert et al., 2020).

---

<sup>1</sup> Diglossia is when there are multiple varieties of the same language throughout a community.

<sup>2</sup> 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 NatCen, in collaboration with Integrated International, 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, 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 protocol describes the evaluation design of a recommissioned pilot of the LRF! C/U model. It aims to understand whether a revised delivery approach, addressing the limitations of the previous pilot, can now show evidence of promise. The new pilot will assess the feasibility of the LRF! C/U intervention, 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 is funded by the Queen Rania Foundation (QRF) and supported by the Education Endowment Foundation (EEF) and the BHP Foundation. The evaluation team is led by NatCen working together with Integrated International, a research organisation based in Jordan.

## Intervention description

This section describes the LRF! C/U model. Please see the previous pilot report for further information about the LRF! W/C approach (Dimova et al., 2023).

### Why

It is estimated that early readers in Arabic need a level of automaticity<sup>3</sup> in oral reading fluency<sup>4</sup> of 45-60 words per minute (RTI, 2012). This fluency allows working memory to be freed-up for comprehension. Data from the use of the EGRA tool in 2018 suggests that only around 19% of Grade 2 and Grade 3 pupils meet, or exceed, the lowest levels of this benchmark. Alongside that, a significant number of pupils in Jordan (16.6%) scored zero in oral reading fluency in 2018 (RTI, 2018).

Early grade literacy approaches can be beneficial in terms of helping pupils as they progress in grade level (EEF, 2023). The 2018 Program for International Student Assessment (PISA) found that for 15-year-old Jordanian pupils', 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 pupils performed at or around the average OECD reading score and two in every five performed below the minimum proficiency level in reading (QRF, 2020). The 2021 Progress in International Reading Literacy Study (PIRLS), which reports on reading levels in 57 countries, found that 10-year-old Jordanian pupils 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-

---

<sup>3</sup> Automaticity is defined as being able to complete a task with no conscious effort.

<sup>4</sup> 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.

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 (EEF, 2021). The importance of phonics is reflected in the EEF's Teaching and Learning Toolkit and other literature (Seidenberg, 2017; Castles et al., 2018). It is important to note that learning losses from the consequences of Covid-19, including sustained school closures, could still have an impact on performance (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 importance placed on 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 and repetition.

### Who

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

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

The C/U intervention targets pupils in Grades 2 and 3 who are among the lowest achieving 20% of students in terms of Arabic literacy. The LRF! C/U model has not been designed to provide support for pupils with Special Educational Needs and Disabilities (SEND).<sup>5</sup>

Pupils are identified for the LRF! C/U model support based on their performance in three assessments; the start-of-term coarse-grain screening tool<sup>6</sup>, an additional diagnostic assessment administered by the Princess Taghrid Institute (PTI), and the EGRA assessment.

### Selection of pupils for LRF! C/U model support

Pupils are identified for inclusion in the intervention through the following process:

- At the start of the school year in September 2023, classroom teachers carry out the **coarse-grain screening tool** with all pupils in their class. This is a part of usual practice in Jordanian schools at the start of the school year.
- Where there are more than three classes per grade, three classes are randomly selected for the evaluation.
- Pupils who score in the lowest 20% of their class in Grades 2 and 3 in the coarse-grain tool will undertake an additional **diagnostic assessment** carried out by PTI.

---

<sup>5</sup> The description used in QRTA communications with teachers is that LRF! is suitable for 'pupils who are academically behind but don't suffer from mental or physical illnesses.'

<sup>6</sup> The Logic Model in Figure 1 below refers to this as a 'RAMP score'. The correct terminology will be used in the updated logic model in the pilot evaluation report.

The purpose of this additional screening is to identify pupils 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 is used as the final screening criterion to identify pupils with reading proficiencies that make them eligible for the programme. Pupils scoring 29 correct words per minute or less for the **ORF sub-scale of EGRA will be eligible for the programme**. This is in line with the four categories of reading proficiency for Jordanian students in Grade 2 and 3 published by RTI International (2023); those scoring 29 correct words per minute or less are categorised into the lowest reading proficiency category of ‘non-readers’.

Following this screening process, pupils 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, will be selected for inclusion in the LRF! C/U sessions.

The pupil screening process is described in greater detail in the *Recruitment* section below.

### **What**

The LRF! C/U model involves a practice-focused pedagogy and pupil practice book developed by cognitive psychologist Dr Helen Abadzi and the Al Qasimi Foundation<sup>7</sup> 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 pupils build ‘low level’ neurological functions: rapidly distinguishing letter shapes, chunking and decoding sounds and words. Altogether, the programme comprises the following elements.

#### **Awareness raising sessions**

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

#### **Teacher training and coaching**

QRTA carries out a two-day training for resource room teachers on how to use the LRF! C/U method and practice book. The training aims 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,
- how to support the involvement of parents/carers – for example, encouraging pupils’ use of the LRF! practice book at home, and supporting parents in this (for example, through WhatsApp messages).

The training includes 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 receive three coaching visits and access fortnightly online meetings. The coaching visits are delivered by QRTA staff, using a coaching model that has been designed to enhance implementation effectiveness.

---

<sup>7</sup> <https://www.alqasimifoundation.com/>

## **Delivery of LRF! C/U sessions**

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

LRF! C/U sessions are provided during the normal school day. Schools decide the scheduling of sessions, which will be agreed between classroom teachers and resource room teachers in each school. The Ministry of Education (MoE) have given permission to use one 'free activity period'<sup>8</sup> 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 pupils to practise 'reading' using the echo reading method<sup>9</sup> ("we do"). These two steps are intended to be completed in the first 10 to 15 minutes of the session. Following this, pupils are asked to independently work through the pupil practice book, taking each item in turn and with their finger on the text sounding out the letter, or word ("you do"). At this stage, 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 is intended to last 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 pupils who have been selected for the programme receive a copy of the LRF! C/U model practice book, which has been updated based on results from the previous pilot. Pupils are encouraged to take it home for extra practice with their parents/carers.

## **Parental engagement**

Schools and resource room teachers also engage with parents and caregivers to involve them in the LRF! C/U model and encourage them to support pupils at home.

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

## **How**

Pupils 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.

---

<sup>8</sup> There is a set curriculum for the free activity period that teachers implement with students; this primarily focuses on enhancing personal skills and values.

<sup>9</sup> Echo reading is when students will repeat out loud what the teacher has read.

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

- Small font sizes negatively affect letter identification, so the book uses large font sizes and spacing.
- The Arabic script is dense and visually complex, with students often identifying words and understanding their meaning slowly (Abadzi, 2012).<sup>10</sup> The 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 children 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).
- Pupils need to see meaning in text, so real words and sentences are introduced as soon as possible.
- The use of pictures in the text are minimised to ensure pupils learn letter sounds, rather than guessing.

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

---

<sup>10</sup> 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).



لَزِمَ مَدَدُ مَوْزُ رَمَزُ  
 رَمَلُ مَوْرَدُ مَوْلِدُ زَمَزَمُ  
 مَزْمَارُ الْمَالُ لَذِيذُ الْوَادِي  
 الْوَالِي دَلِيلُ الْوَزِيرُ مَاذَا  
 مَذَارُ الْوَالِي رَمَى زَمَزَمُ  
 مَالِي دَلِيلُ مَلَأَ رَامَ  
 مُرُورُ الْوَزِيرُ مَرَامُ مِيلَادُ

48

الْمَالُ رَمَلُ رُمُوزُ لَذِيذُ  
 مَدَدُ مَزْمَارُ الْوَادِي لَامُ  
 رَمَزُ رَمَلُ مَالُ مَوْلِدُ  
 رَامِي مَوْرَدُ مُرَادُ مَارِدُ  
 مَلَأَ مَدَى لَزِمَ مَوْزُ

اقْرَأ الْجُمْلَةَ وَالْعِبَارَاتِ الْآتِيَةَ:

الْمَوْزُ لَذِيذُ

49



اقْرَأُ الْجُمْلَ وَالْعِبَارَاتِ الْآتِيَةَ:

يَعْدُو الْمُهْرُ وَرَاءَ الْأَرَانِبِ.

أَعْطَانِي وَالِدِي هَدِيَّةً يَوْمَ مَوْلِدِي.

مَتَى أَزْهَرَتِ الْوُرُودُ؟

أَزْهَرَتِ الْوُرُودُ يَوْمَ الْأَرْبَعَاءِ.

مَتَى رَنَّ الْمُنْتَبَهُ؟

رَنَّ الْمُنْتَبَهُ بَعْدَ الْغُرُوبِ بِثَلَاثِينَ

ثَانِيَةً.

110

أَوَدُّ أَنْ أَنَامَ بِهُدُوءٍ تَامٍّ.

لِمَاذَا رَمَى وَايِلُ ثِيَابَهُ؟

رَمَى وَايِلُ ثِيَابَهُ لِأَنَّهَا مُهْتَرِئَةٌ وَرَثَّةٌ.

هَلْ تَوَدُّ أَنْ تَتَنَاوَلَ الْغَدَاءَ؟

نَعَمْ، أَوَدُّ أَنْ أَتَنَاوَلَ الْغَدَاءَ.

مَاذَا تَتَنَاوَلْتُ عَلَى الْغَدَاءِ؟

تَتَنَاوَلْتُ الْأَرْزَ الْمُبَهَّرَ مَعَ اللَّبَنِ.

لَا أَتَنَاوَلُ الْمِيَاهَ الْغَازِيَةَ أَبَدًا.

111

اقْرَأُ الْكَلِمَاتِ الْآتِيَةَ:

وَزَرَ دَرَزَ دُرُرُ

وَرَدَ رَدَدَ وَرْدُ

دَوَّرَ وَزَّرَ وَزَّرُ

30

رَدَدَ دَوَّرَ دُرُرُ

وَزَّرَ دَرَزَ وَرَدَ

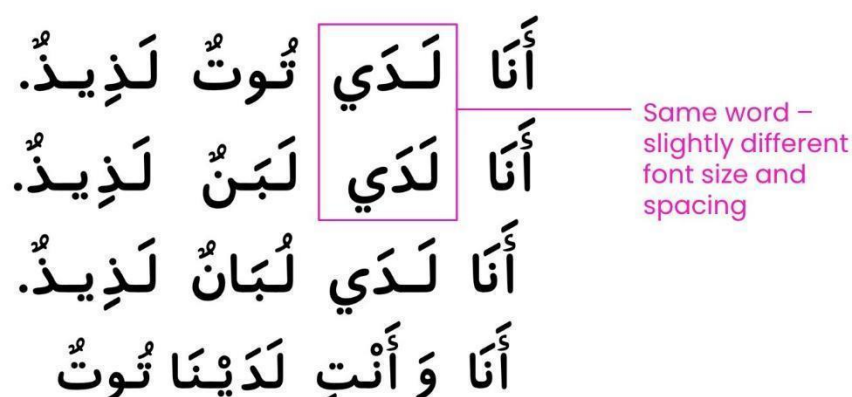
وَرْدُ وَرَدَ دُرُرُ

وَزَرَ وَرْدُ دَوَّرَ

دَرَزَ رَدَدَ وَزَّرَ

31

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



### When

The intended delivery of the LRF! C/U model is 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 is set, resource room teachers have some flexibility over how they facilitate it. They are encouraged to draw on their professional judgement to tailor instructions according to pupils' needs. They are expected to ensure that they progress through the content of the practice book, whilst ensuring that pupils are able to adequately master each lesson as they do so.

### Rationale for the revised pilot

Some of the key findings relevant 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 pupils' needs while still completing the syllabus.	The maximum capacity of pupils in each resource room session has been capped at six by the Ministry of Education.
Some resource room teachers found delivering the LRF! C/U model challenging, due to working with pupils facing difficulties with their learning. Resource room teachers also typically have less experience than	The training for resource room teachers will be extended to two days instead of one. Classroom teachers will no longer be trained to deliver the LRF! C/U model. The training will be for resource room teachers only.
	Beyond the three coaching sessions that

classroom teachers.	were already part of the intervention, resource room teachers will now also receive additional support through fortnightly online meetings. QRTA will establish a Professional Learning Community (PLC), which resource room teachers can draw ongoing support from during the intervention via online meetings.
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. This has not been implemented for this pilot.	The reading practice book was developed based on the feedback given by Dr. Helen Abadzi and teachers from the previous pilot phase. Changes include replacing some syllables, words and phrases, changing fonts, font size and spacing.
LRF! C/U was too difficult for pupils in Grade 1.	The C/U model is now implemented with pupils in Grades 2 and 3 only. Grade 1 is no longer included.
The pace of content was too fast for struggling pupils, 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 pupils, leaving some feeling frustrated.	
The coarse-grain screening tool alone did not identify the pupils most suited for the C/U intervention. The tool identified a very high proportion of pupils struggling with literacy. The evidence suggested that in some cases teachers responded to this by selecting pupils for the intervention based on their own judgement, which they felt was more accurate than the diagnostic tool. In the end, C/U pupils 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 pupils for inclusion in the C/U model has been updated. Rather than relying on scores from the coarse-grain tool alone, an additional diagnostic assessment tool developed by PTI, along with EGRA scores, will be used to help screen out those for whom LRF! C/U support is not considered suitable.

This pilot evaluation will focus on the aspects of the C/U model that have changed, rather than re-examining consistent elements of the programme where the pilot evidence has already been gathered in 2021-22 (Dimova et al., 2023).

### *Updated logic model*

The LRF! C/U logic model has been updated in response to findings from the previous pilot. These changes are reflected in an updated logic model, shown in Figure 1 below. This was discussed during a workshop in August 2023 attended by key members of QRF and the evaluation team. Moving from left to right, its components are:

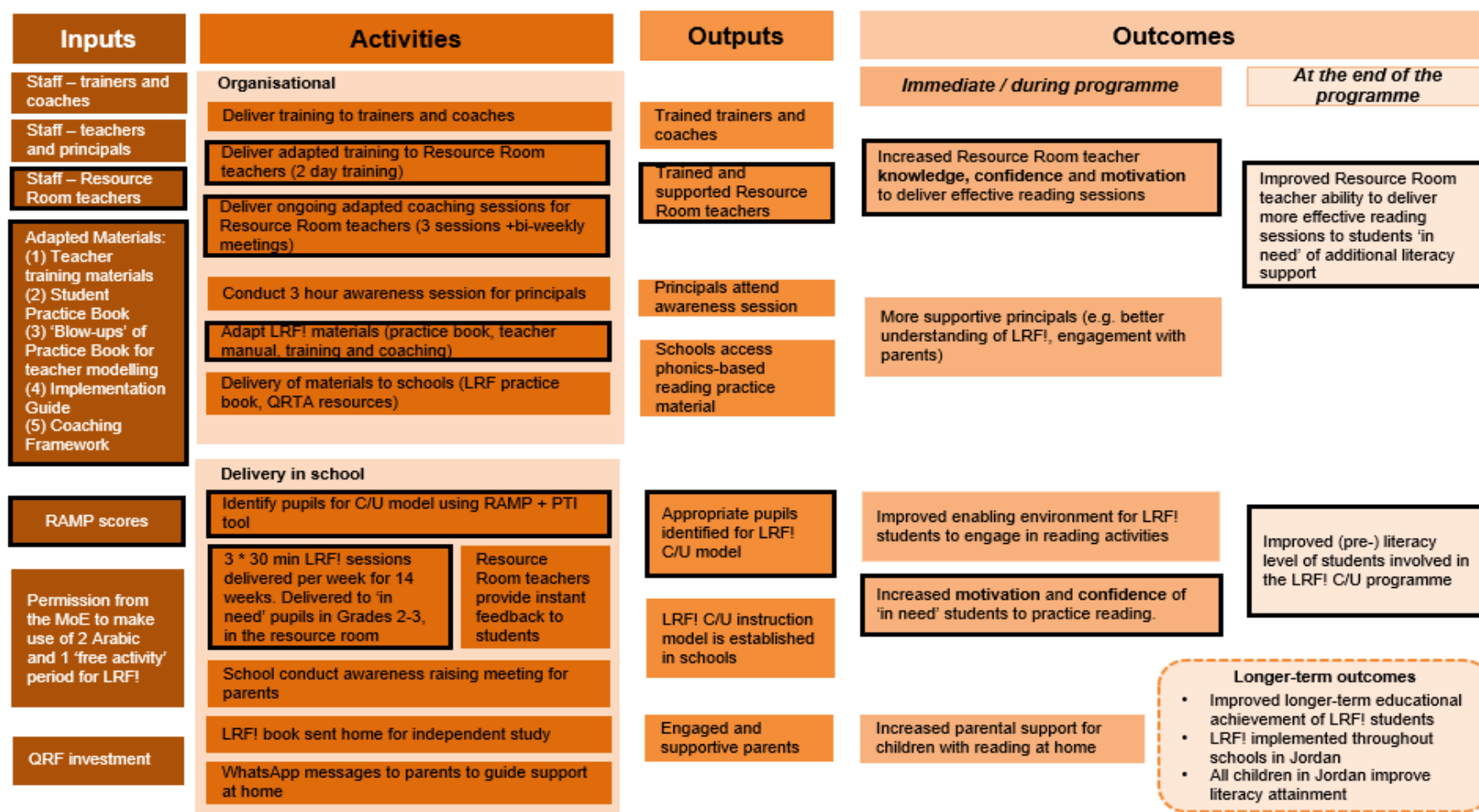
- **Inputs** describe the resources required to develop and deliver the C/U model.

- **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 future. This captures the longer-term vision of the C/U programme.

The boxes of the revised logic model that are **outlined in black** show the key components of the revised C/U logic model that will be the focus of the upcoming pilot evaluation work in 2023-24. 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 have been removed from the logic model.

Figure 1 Updated LRF! C/U logic model

## C/U refined 2<sup>nd</sup> pilot – Logic model v3



**Notes:**

- 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
- 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 have been 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 do not need to be addressed again.

We will be triangulating findings during analysis and report writing; findings from different data sources and across the IE and IPE will be combined in order to offer a nuanced understanding of the LRF! C/U model. For full details of how each research question will be addressed, including the data collection method, sample size, and timing of data collection, please see [Appendix A](#).

### *Evidence of promise*

1. In what ways, and to what extent, does the LRF! C/U model affect school, teacher, and pupil 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 pupils 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)

## Impact Evaluation Methods

### *Trial design*

The pilot evaluation is a two-arm cluster randomised controlled trial (RCT), with schools as the unit of randomisation and pupils 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 children were screened for their eligibility for the LRF! C/U model. Eligible pupils will take part in baseline and endline EGRA testing and be included in the IE analysis for this pilot.

Note that as a pilot evaluation, this study is not intended to provide a robust estimate of the causal impact of the LRF! C/U programme. This will be the focus of a future efficacy trial, if the findings of the pilot suggest that the C/U model is feasible to deliver, shows promise and is ready to be evaluated in a future trial. Consequently, while this pilot evaluation involves comparing literacy outcomes between C/U and usual practice, this evidence will be treated as indicative of the potential promise of this intervention, and not as a robust causal estimate.

The study design is summarised in Table 2.

**Table 2: Study design – 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)	<b>Source:</b> EGRA Grade 2 assessment with the addition of a set of pre-literacy items <b>Instrument:</b> EGRA + pre-literacy tool <b>Subdomain:</b> Oral reading fluency
Secondary outcome(s)	Variable(s)	Arabic literacy attainment and specific sub-domains of Arabic literacy attainment
	Measure(s) (instrument, scale, source)	<b>Source:</b> EGRA Grade 2 assessment with the addition of a set of pre-literacy items <b>Instrument:</b> EGRA + pre-literacy tool <b>Sub-domains:</b> Arabic literacy attainment, letter sound identification, syllable identification, reading comprehension, word decoding and listening comprehension
Baseline for primary	Variable	Oral reading fluency



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

### Recruitment

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

- were not part of the previous pilot.
- had students in Grades 2 and 3.
- were a single shift school.<sup>11</sup>
- were not taking part in any other literacy interventions, other than the Reading and Writing Project that has been carried out in Grades 1-3 in all MoE schools in Jordan since 2015.
- had a resource room teacher.
- were not delivering blended teaching.<sup>12</sup>
- were not in a Syrian refugee camp.<sup>13</sup>
- were from the middle, near north and near south<sup>14</sup> regions.

The recruitment process involved these steps:

- QRF shared an updated list of schools in Jordan, including school characteristics and contact information.

<sup>11</sup> In Jordan, schools may operate on one shift or two shifts. Schools operating on two shifts (morning and afternoon shift) have different group pupils in morning and afternoon shift, while schools operating on one shift have one group of pupils during the whole day.

<sup>12</sup> 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.

<sup>13</sup> 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 would include schools that have 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

<sup>14</sup> The North and South regions were narrowed to near north and near south for logistical reasons.

- NatCen researchers cleaned the data and excluded schools that were not eligible for the trial according to the criteria above. We 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.

School recruitment took place in November 2023. We recruited 16 schools for the pilot. 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.

### **Randomisation**

Schools recruited for the pilot were randomly allocated to either C/U 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). 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 have four geographical regions altogether: (1) Amman, (2) middle (excluding Amman), (3) near south and (4) near north. Given that we have four geographical regions, and a school could be either in a rural or urban area, we have eight strata altogether.

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

### **Participants**

As discussed above, pupils were identified for the C/U model 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 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<sup>15</sup> developed by the Early Grade Reading and Mathematics Programme (RAMP)<sup>16</sup> 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

---

<sup>15</sup> 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 \(rti.org\)](https://www.rti.org/jordan-remedial-study-reading-diagnostic-assessment-tool-and-stimulus-sheet)

<sup>16</sup> 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 maths 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](https://pdf.usaid.gov/pdf_docs/PA00THHW.pdf).

20% of pupils in each class based on these scores. They shared this information with QRF, PTI and NatCen.

For the lowest performing 20% of pupils only, practitioners from PTI then carried out an additional screening tool between December 2023 and February 2024. This involved first having a telephone call with parents or carers to ask about the pupils' developmental history, lasting for between 15 and 45 minutes. Then PTI carried out a language development test<sup>17</sup> and learning difficulty test<sup>18</sup> with pupils. The tests were carried out by PTIs' specialist team and lasted about 30-45 minutes per pupil, with breaks as required by the pupil. If a pupil exhibited indicators of delay in language skills using the language development and learning difficulties tests, they were identified by PTI to have delays in language development and were not recommended for the LRF! programme. Additionally, if a PTI therapist noticed indicators of global developmental delay using these tests and/or the pupil was reported to have ability indicators of global developmental delay by parents/carers during the screening phone call, these pupils were also not recommended for the LRF! programme. Pupils who had learning difficulties that were not indicative of global developmental delay or language delay were recommended for the LRF! programme as the programme can work 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 pupils were screened using the PTI tool and those with potential special educational needs or learning disabilities excluded, pupils took part in EGRA testing, described in the following section, in January and February 2024. 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. As LRF! helps pupils' reading fluency, scores from the ORF were used to identify those eligible for the C/U model. Pupils who are categorised as non-readers and beginning readers based on their ORF score (i.e., those who scored 29 correct words per minute or less) were deemed eligible for the programme.

### **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 is described in detail in the previous pilot report (Dimova et al., 2023). During the set-up meetings, we agreed to have the same tool that was used in the previous pilot.

As LRF! C/U mainly focuses on improving pupils' reading fluency, NatCen and Integrated International agreed with QRF and EEF that the primary outcome measure would be Oral Reading Fluency (ORF).<sup>19</sup> This will be obtained by administering the EGRA+pre-lit. As part of EGRA testing, students will be given a short story and asked to read it within one minute. The story will consist of 42 words. The ORF will be measured as the number of correct words read per minute, while ensuring standardisation if students read the passage in less than 60 seconds.

---

<sup>17</sup> 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.

<sup>18</sup> 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, evaluate difficulties in understanding and comprehension, speech capabilities and reading and writing skills.

<sup>19</sup> The updated C/U logic model (Figure 1) will be changed in the final report to reflect changes to the primary outcome measure.

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 would be ideal to assess the effect of the programme on reading skills that could be predictive of future successful reading (RTI International, 2015). We will include the specific sub-domains of Arabic Literacy Attainment (ALA) as measured by the EGRA as secondary outcome measures. We will produce separate metrics for these sub-domains from the EGRA. For the timed sub-domains of the EGRA, the per-minute score will be calculated as explained by the EGRA toolkit (RTI International, 2015). This scoring system allows differentiating pupils who have got the same number of correct items but finished in different time periods. The per-minute scores will be calculated based using the following formula:

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

The first sub-domain of the EGRA is Letter Sound Identification (LSI). It assesses a pupil's ability to associate sounds with letters. In this task, pupils are given 100 letters and asked to read the letter sounds within a minute. A per-minute score LSI will be created to assess pupils' ability on LSI. Pupils also complete a sub-domain of the EGRA which is used to assess pupils' ability to identify syllables. In this sub-domain, pupils are given 100 Arabic syllables to read within one minute. The number of correct syllables is used to create a per-minute score, which reflects pupil's ability to identify syllables. Pupils also complete a sub-domain of the EGRA which is used to assess pupils' decoding ability (i.e., the sub lexical route of word processing (RTI International, 2015). This sub-domain of the EGRA assessment includes a list of 50 one- and two-syllable nonwords to read within one minute. The number of correct nonwords is used to create a per-minute score for decoding nonwords. Furthermore, once pupils complete reading the short story for the ORF measure, they are asked five questions related to the short story to assess their Reading Comprehension (RC). They are 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 is Listening Comprehension (LC). This sub-task of EGRA includes a passage read by the assessor and is followed by oral comprehension questions answered by a pupil. Pupils are 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 has changed from ALA to ORF, the ALA will be part of the secondary outcome measures for consistency with the previous pilot evaluation. The ALA will be a single metric formed from the sub-domains of the EGRA. Every sub-domain of the EGRA has a different scale as the number of questions varies across the sub-domains. Each sub-domain will be standardised to have a mean of zero and standard deviation of one. We will then take the average of the standardised scores to create the single metric for the ALA. Due to differing scales on each sub-domain, some would have higher influence on the average score without standardisation and therefore it is necessary to standardise to ensure equal influence of each sub-domain on the average score. We will also estimate Cronbach's alpha to measure the internal consistency of the measure.

EGRA+pre-lit is administered at both baseline (February 2024) and endline (May 2024). Baseline primary and secondary measures will be identical to the endline primary and secondary measures, respectively. All eligible pupils in both C/U and usual practice schools are included in baseline and endline assessments.

## Analysis

### Primary analysis

In line with the EEF analysis guidance<sup>20</sup>, the primary outcome analysis will follow an Intention-To-Treat (ITT) approach<sup>21</sup>. The analyses will use a three-level multilevel model to account for the clustering of pupils (level 1) in classes (level 2) within schools (level 3). Random assignment to conditions is at the school level (level 3). This model includes school and class-level random intercepts and accounts for the baseline ORF and the stratification variable. 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 pupils (*i*) are clustered in classes (*j*) within schools (*k*). The intervention effect will be estimated by  $\beta_2$ ,  $\beta_3$  represents a vector of strata fixed effects for the schools (i.e., their geographical location),  $\beta_0$  is the intercept,  $\beta_1$  is 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<sup>22</sup>, other additional covariates will not be considered. The analysis will be carried out using Stata 17 or higher. Both the syntax used, and outputs of analysis will be saved as a record of the process. The difference in the means between the intervention and usual practice groups at endline will be expressed as a standardised effect size using Hedges' *g* with 95% confidence intervals. Following EEF guidelines, the numerator will come from the unstandardised effect estimate given by  $\beta_2$  in the multilevel model specified above, which will be adjusted for baseline score and strata. The denominator will be the unconditional pooled standard deviation in the primary outcome at endline. The formula is provided below:

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

where  $n_1$  and  $n_2$  are the number of pupils in both groups and  $s_1^2$  and  $s_2^2$  are the within-group variances in outcomes at endline. The primary analysis will also include detailed descriptive analysis: histograms, means, quartiles, and SDs, for all measures, groups, and time points. We will also report school and class-level Intraclass Correlation Coefficients (ICCs) alongside 95% confidence intervals in analyses. We will use the following model with a random effect by school,  $v_k$  a classroom-level random effect,  $u_{jk}$ , and the residual term,  $e_{ijk}$ . The model will have no fixed effects for strata so it will be arithmetically possible that the variance of  $v_k > 0$ :

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

### Secondary analyses

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

---

<sup>20</sup> [Evaluation design | EEF \(educationendowmentfoundation.org.uk\)](https://www.educationendowmentfoundation.org.uk/evaluation-design)

<sup>21</sup> 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>

<sup>22</sup> [Evaluation design | EEF \(educationendowmentfoundation.org.uk\)](https://www.educationendowmentfoundation.org.uk/evaluation-design)

For all defined secondary outcomes (e.g., letter sound identification, syllable identification, and reading comprehension), we will follow the ITT approach similar to that of the primary outcome analysis. This statistical model will take the form of a multilevel model, where pupils are clustered in classes within schools, and accounts for stratification factors used at randomisation.

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

where pupils (*i*) are clustered in classes (*j*) within schools (*k*). The intervention effect was estimated by  $\beta_2$ ,  $\beta_3$  represents 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.

For all secondary outcome measures, we will report confidence intervals at the 95% level and the effect size using Hedges' *g* as previously described.

## Implementation and Process Evaluation Methods

An Implementation and Process Evaluation (IPE) will be carried out to address the evaluation domains set out in Table 3 below. Our IPE domains of interest are informed by the EEF's framework for IPE's (EEF, 2022; 2019). The domains we will assess are fidelity (the extent that those implementing LRF! C/U adhere to the model), including dosage (how much of the intended C/U model of LRF! has been delivered), quality (how well different components of the LRF! C/U model have been delivered), and adaptation (any changes made to the LRF! C/U model). In addition, we will investigate reach (the rate and scope of participation in the LRF! C/U model), responsiveness (the degree to which participants engage with the LRF! C/U model), perceived impact (the perception of teachers and coaches as to whether the LRF! programme has achieved its intended outcomes), and usual practice (what usual literacy teaching looks 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 3, and are described in more detail below.

**Table 3: Mapping of IPE dimensions, RQ's, data collection methods and analysis**



IPE dimension	Research questions addressed	Research and data collection method	Data analysis methods
Fidelity: Dosage	RQ3, RQ5, RQ13	RR teacher FGDs	<i>Qualitative; framework analysis</i>
		RR teacher end of programme survey	<i>Quantitative; descriptive analysis</i>
		Classroom teacher end of programme survey	
Fidelity: Quality	RQ5, RQ8, RQ9, RQ10	RR teacher FGDs	<i>Qualitative; framework analysis</i>
		Coaches FGD	
		RR teacher post-training survey	<i>Quantitative; descriptive analysis</i>
		RR teacher end of programme survey	
	RQ5, RQ8	Classroom teacher end of programme survey	
Fidelity: Adaptation	RQ3, RQ5, RQ13, RQ14	RR teacher FGDs	<i>Qualitative; framework analysis</i>
		Coaches FGD	
		RR teacher end of programme survey	<i>Quantitative; descriptive analysis</i>
		Classroom teacher end of programme survey	
Reach	RQ3, RQ6, RQ12, RQ13	RR teacher FGDs	<i>Qualitative; framework analysis</i>
		Coaches FGD	
	RQ7	RR teacher post-training survey	<i>Quantitative; descriptive analysis</i>



	RQ3, RQ6, RQ12, RQ13	RR teacher end of programme survey	
		Classroom teacher end of programme survey	
	RQ1	Attendance data	
Responsiveness	RQ3, RQ7, RQ9, RQ10, RQ11, RQ13, RQ14	RR teacher FGDs	<i>Qualitative; framework analysis</i>
		Coaches FGD	
	RQ7, RQ10, RQ11, RQ14	RR teacher post-training survey	<i>Quantitative; descriptive analysis</i>
	RQ2, RQ3, RQ7, RQ9, RQ10, RQ11, RQ13, RQ14	RR teacher end of programme survey	
	RQ2, RQ3, RQ10, RQ11, RQ13, RQ14	Classroom teacher end of programme survey	
Perceived impact	RQ2, RQ3, RQ4, RQ13	RR teacher FGDs	<i>Qualitative; framework analysis</i>
		Coaches FGD	
		RR teacher end of programme survey	<i>Quantitative; descriptive analysis</i>
		Classroom teacher end of programme survey	
Usual practice	RQ1	RR teacher FGDs	Qualitative; framework analysis
	RQ1	Usual practice classroom/RR teacher FGDs	

We will gather information about **the number of participants attending LRF! C/U training and coaching sessions** from attendance registers collected during training. This information will be supplied by QRTA and will offer an indication of the reach and take-up of training.

**Focus Group Discussions (FGDs)** will be carried out with groups of intervention resource room teachers, and groups of usual practice teachers including both resource room and classroom teachers. Further FGDs will be held with coaches. FGDs will help explore perspectives around the

intervention and its context, and will be used to deepen and triangulate findings. Each FGD with teachers and coaches will last around 40-75 minutes and will be conducted either in person or online. The mode of delivery that yields the highest attendance will be selected. All focus groups will be led by trained enumerators. If groups take place online, enumerators will encourage all respondents to make use of various methods of communication, including raising a virtual hand to speak, and utilising the chat function to offer opinions. In both online or in-person groups, enumerators will be responsive to the group dynamics and ensure that all voices are heard and opinions respected.

The aims of the resource room teacher FGDs include: understanding perceptions around the challenges and opportunities presented by the LRF! C/U model, opinions about the suitability of the LRF! C/U approach and resources, perceptions of pupil use of the practice book, any adaptations they have made to the LRF! C/U model and any perceived changes in learning outcomes or behaviours among pupils due to the intervention. The usual practice teacher FGD aims to better understand current practice in teaching reading and in identifying reading difficulties, as well as any current barriers and enablers for pupils learning to read. This context may have changed since the previous pilot due to educational reforms and provides an important comparison to the LRF! programme.

**Surveys of resource room teachers** implementing the LRF! C/U sessions will be undertaken to collect data on their perspectives on training and coaching, delivery to pupils, dosage, and resources needed to implement the intervention. Each survey will take approximately 10 minutes to complete over the telephone with Integrated enumerators (one survey will take place after the initial training, and one towards the end of the programme). An additional **survey of classroom teachers** in the intervention group will take place towards the end of the programme to collect data on their perspectives on the consequences of the intervention on the children involved, as well any barriers to implementation.

## **IPE numbers and sampling**

### *Numbers and sampling for FGD*

Based on the overall sampling approach for the trial for the impact evaluation, there will be eight strata (four regions and two urban rural). Each stratum will have two schools, one intervention and one usual practice. Teacher focus groups in the intervention group will take place with one resource room teacher from each school (two focus groups of four resource room teachers each).

For the two usual practice group focus groups, there will be a mixture of class teachers and resource room teachers, consisting of two class teachers and two resource room teachers in each group (covering Grade 2 and 3 in each group).

Sampling for the focus groups will be carried out with attention to equity considerations, and inclusion of a diversity of viewpoints. An overview of participant numbers is outlined in Table 4 below.

**Table 4: Participants in FGDs**

Group	Teachers					Total teachers
	Number of FGDs	Near North	Near South	Middle (Madaba only)	Amman	
Intervention	2	1 urban	1 urban	1 urban	1 urban	8
		1 rural	1 rural	1 rural	1 rural	
Usual practice	2	1 urban	1 urban	1 urban	1 urban	8
		1 rural	1 rural	1 rural	1 rural	

Informed verbal consent will be given by participants at the start of each session. Those who no longer wish to take part or do not want to be recorded will have the ability to leave at this point. The option to not record the session will also be given.

#### *Numbers and sampling for the teacher survey*

All resource room teachers from the intervention group will be asked to complete a survey after the initial 2-day training, in addition to a separate survey towards the end of the programme. For the survey of class teachers, all Grade 2/3 teachers in intervention schools will be invited to participate to provide information about the impact of LRF! C/U on pupils. All surveys will take around 10 minutes to complete and will be conducted over the telephone with trained enumerators. Participants will provide informed verbal consent at the beginning of the telephone call.

#### **IPE analysis**

Data from the teacher surveys, collected using Alchemer survey software, will be exported to SPSS for descriptive analysis. Survey analysis will be conducted first, prior to analysis of FGDs. This will enable any top-level findings to be identified. The survey data will be used to triangulate findings from qualitative research methods, analysed as explained below.

The consortium will use NatCen's Framework approach (Ritchie et al., 2003), adapted for bilingual working, to manage data from focus groups. 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 which emerge from the data, we will assemble a matrix in which each row represents a FGD and each column a theme/ sub-themes. We will then summarise the qualitative data in the matrix, including illustrative verbatim quotes where appropriate. Once all data has been coded in this matrix, we will move onto analysis. This will involve a phase of 'detection', including studying what participants say about a particular phenomenon, listing these and then sorting them thematically. Once we identify different themes in

the data, we create higher-level categories that work as meaningful conceptual groupings for participants' views and experiences within and across schools.

Summaries will be produced in English and shared with the wider team for review. Core team members will then come together for an analytical planning meeting, where key themes, patterns and issues emerging from the qualitative data are discussed, and across-team verification of findings established. Findings will be reviewed in detail and mapped against the pilot RQs/LM in preparation for reporting.

## Ethics and registration

### *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 will be 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.
- We will ensure participants are not identifiable in the outputs.

### *Registration*

The trial was registered on the Open Science Foundation in April 2024 (<https://doi.org/10.17605/OSF.IO/J35WZ>). The trial registry will be updated with outcomes at the end of the project.

## Data protection

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

NatCen is fully accredited to ISO 27001 and subject to annual external audits of procedures to maintain accreditation. We also hold Cyber Essentials Plus Certification. We are fully general Data Protection Regulation (GDPR) compliant.

EGRA assessments are undertaken by Integrated, with pupils assigned a unique identifier. Test results are submitted to EGRA's Tangerine tool and a pseudonymised dataset transferred to NatCen via NatCen's Secure File Transfer Protocol (FTP) server. EGRA data are stored with back-end provider Prodigy and sent directly to authorised Integrated personnel.

Integrated will store the IPE data on a dedicated drive that can only be accessed by authorised personnel. All information stored, processed and/or transmitted by Integrated is protected in a manner consistent with contractual and legal restrictions proportionate to the level of sensitivity,

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

Data shared with NatCen is stored on NatCen's secure network, with access to the project folder restricted to authorised personnel only. The data is backed up, and NatCen carries 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.

A data retention period will be agreed with QRF. Once this period has expired, all relevant parties will securely erase project data (with explicit permission from QRF).

## Personnel

NatCen is the lead partner and accountable to QRF. NatCen leads on project coordination and management, evaluation design, analysis and reporting. Integrated is a subcontracted partner. They lead EGRA testing, IPE data collection and contribute to evaluation design, analysis and reporting. Members of the evaluation and delivery team are outlined in Table 5 and Table 6, respectively.

**Table 5: 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 6: Delivery team**

Name	Department/Project role	Role and team
Rola Said	Academic Affairs (Advisor)	Director of Programs
Feras Al Omari	Academic Affairs (Advisor)	Senior Education Specialist
Lubna Al Drainy	Academic Affairs (Technical project lead and coach)	Senior Education Specialist
Mohammad Salameh	Academic Affairs (Coach)	Teacher Educator
Sana Al Syouf	Project Management (manage the project plan and arrange logistics)	Project Officer
Fadi Al Mobaidein	Support Services (training and material procurement)	Procurement & Logistics Manager
Ala'a Saymeh	Finance (keep track of project financial expenses and reporting)	Senior Finance Officer



## Risks

We take a proactive approach to the management of risks, considering the likelihood and potential impact of key risk factors, as well as mitigations and contingencies (see Table 7).

**Table 7: Key risks, including the likelihood of occurrence and potential impact (High [H], Medium [M], and Low [L]).**

Potential Risk	Likelihood	Impact	Mitigation and contingencies
Slippage in research timetable	H	M	Detailed project timetables developed and regularly reviewed to identify problems early and focus activity on addressing them. There are options to expand the project team to catch up and to reschedule work if necessary.
Partnership does not work together successfully	M/H	H	We will work closely with QRF to assess the degree to which LRF! C/U model implementation, evaluation delivery and data collection are affected. We will work flexibly with QRF to adjust our analysis, approach, and timetable accordingly to ensure impacts can be measured in accordance with implementation. The corresponding caveats for interpretation, if present, will be added.
Major disruption to evaluation fieldwork resulting from public health emergency, extreme weather, or other national emergencies	M	M	We will work flexibly with QRF and participants to ensure data is collected safely. We will draw on our extensive experience of delivering research during major disruptions, such as the Covid-19 pandemic, which has involved collaborating with participants remotely, offering flexibility with data collection mode (online and phone), and being receptive to last-minute cancellations.
Fieldwork staff illness/unavailability/turnover due to various reasons	M	M	We have a large team of researchers and fieldwork assistants to draw on should availability unexpectedly change (changes will be made in agreement with QRF).
Difficulties engaging participants leading to failure to meet target recruitment rate	M	H	We will discuss recruitment strategies with QRF and work closely to complete recruitment on time during the inception stage.
School and/or pupil level attrition	M	M/H	This will be addressed by setting out the requirements for the trial at the outset and providing schools with clear instructions at the start of the project on what needs to be done and when. Our experienced field team will ensure as many children as possible are tested in

			each school at baseline and endline. Power calculations will account for expected attrition and will be informed by attrition levels from the previous pilot.
Unintended consequences	L	L/M	Unintended impact will be monitored and necessary action taken based on risks involved and ensuring the wellbeing of participants throughout.
Data security, confidentiality, and privacy	L	H	Robust data security measures will be implemented, following privacy rules and regulations.

## Timeline

**Table 8: Evaluation timeline**

	School Calendar	Proposed dates	Notes
	Semester 1	20 Aug 23 – 7 Jan 24	-
	Semester 1 exams	19 Dec 23 – 6 Jan 24	-
	Semester 2	21 Jan 24 – 26 Jun 24	-
	Semester 2 exams	1 Jun 24 – 25 Jun 24	-
Aug 23 – Feb 24	<b>Key milestones</b>	<b>Indicative deadline</b>	<b>Notes</b>
	<b>Pilot set up</b>		
	Logic model workshop	Aug 23	-
	Logic model agreed	Nov 23	-
	Coarse-grain screening tool carried out	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	-
	<b>Pilot Catch up Model implementation - Semester 2</b>		
	Teacher training and principal awareness sessions delivered	Jan 24	-
	EGRA baseline testing	Feb 24	-
	School resources delivered	Feb 24	-
	Coaching delivered	Feb – May 24	-
	Intervention delivered (12 weeks)	11 Feb – 16 May 24	Inc. Eid break
	FGDs with coaches	6-13 May	-
	FGDs with RR and class teachers (usual practice)	6-13 May	-
	End of programme surveys with RR and classroom teachers (intervention)	6-13 May	-
Jan 24 – Oct 24	FGDs with RR teachers (intervention)	12-16 May	-
	EGRA endline testing	19 May 24 – 23 May 24	-
	Efficacy trial decision made	Summer/Autumn 2024	-

## Appendix A – Summary of how each research question will be addressed

**Table 9: Summary of how each research question will be addressed**

Dimension	Research Question	Criteria [linked to logic model]	Data Collection Methods	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. RR 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/RR teachers (usual practice)	8	During programme
			FGD RR teachers (intervention)	8	End of programme
			Surveys with RR 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 3 coaching sessions, and the majority of additional bi-weekly meetings with coaches [A1, A2]	Attendance data for training	8	During programme
		EP1c. Pupils use the practice book in more than half of lessons and show motivation to take it home [ST2]	FGD RR teachers (intervention)	9	End of programme
			Surveys with RR / 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?	<p>EP2. More than half of classroom and RR teachers perceive that there is a positive value in this intervention compared to/in addition to usual practice:</p> <ul style="list-style-type: none"> <li>- RR 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]</li> <li>- RR teachers report increased knowledge, confidence and motivation to deliver effective reading sessions [ST1]</li> <li>- RR / class teachers believe that students involved in the CU programme have increased confidence and motivation to practise reading [ST2]</li> <li>- RR / class teachers agree that students involved in the LRF! C/U programme have improved (pre-) literacy level [LT2]</li> </ul>	FGD RR teachers (intervention)	8	End of programme
			Surveys with RR 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 RR teacher ability to deliver more effective reading sessions to students 'in need' of additional literacy support [LT1]	FGD RR teachers (intervention)	8	End of programme
			Surveys with RR teachers (intervention)	8	After training and end of programme
			Surveys with class teachers (intervention)	8	End of programme
		EP3b. Coaches and intervention class and RR teachers agree that students involved in the LRF! C/U programme have improved	FGD RR teachers (intervention)	8	End of programme

		(pre-) literacy level [LT2]	Surveys with RR / 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 <sup>23</sup> [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  a) students participating in the programme b) other students in the class c) RR or class teachers d) parents	FGD RR teachers (intervention)	8	End of programme
			Surveys with RR /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 quality? What modifications were made, with what implications?  FI5. RR teachers and coaches report that the intervention was delivered as intended in terms of a) Dosage: at least 2 out of 3 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	FGD RR teachers (intervention)	8	End of programme
			Surveys with RR teachers (intervention)	8	End of programme
			FGD with coaches	2	Towards end of programme

<sup>23</sup> We need to be cautious about interpreting the results of the end line assessment given the small sample size

		effective and engaging for students ‘in need’ of additional literacy support [D2, ST2, LT1]			
FI6. What is the learning about the use of the PTI diagnostic tool? How successful is it, in use, at identifying the most appropriate pupils for the C/U model?	FI6a. Teachers and coaches agree that the coarse-grain screening tool + PTI tool effectively identified appropriate pupils for the CU programme		FGD RR teachers (intervention)	8	End of programme
			Surveys with RR / 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 RR teachers (intervention)	8	End of programme
			Surveys with RR / 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 RR teachers (intervention)	8	End of programme
			Surveys with RR 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	FI8. RR teachers and coaches perceive that the RR teachers have the skills and confidence to effectively deliver the CU programme following the training and coaching sessions [O1, ST1, LT1]		FGD RR teachers (intervention)	8	End of programme
			Surveys with RR teachers (intervention)	8	After training and end of programme



	the training and coaching?		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 RR teachers (intervention)	8	End of programme
			Surveys with RR 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 implementation of LRF! C/U model?	n/a	FGD RR teachers (intervention)	8	End of programme
			Surveys with class / RR 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	Delivery team assessment	N/A	End of programme
			FGD RR teachers (intervention)	8	End of programme

		<p>within the time period necessary for the efficacy trial (i.e. pre-EGRA testing and programme implementation)</p> <p>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)</p> <p>f) the working relationships / coordination between partners has been positive over the course of the pilot</p>	Surveys with class / RR teachers (intervention)	16 (8 per survey)	End of programme
			FGD with coaches	2	Towards end of programme
		FT11b. More than half of eligible students <sup>24</sup> 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 RR teachers (intervention)	8	End of programme
			Surveys with RR teachers	8	After training and end of programme
			Surveys with class teachers (intervention)	8	End of programme

<sup>24</sup>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.

			FGD with coaches	2	Towards end of programme		
			Delivery team assessment	N/A	End of programme		
		FT12b. Any modifications to these tools are identified, based on evidence from the pilot, and are possible to implement if scaled up	EGRA data	204	Baseline and endline		
			FGD RR teachers (intervention)	8	Towards end of programme		
			Surveys with class/RR (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 RR teachers (intervention)	8	End of programme
					Surveys with class / RR teachers	16 (8 per survey)	End of programme

	RT14. What changes to the intervention, implementation models, support or materials need to be made?	n/a	FGD with coaches	2	Towards end of programme
			Surveys with RR teachers (intervention)	8	After training and end of programme
			Surveys with class teachers (intervention)	8	End of programme
			FGD with coaches	2	Towards end of programme
			FGD RR 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 B – Power Calculations

For this trial, 16 primary schools were recruited with eight schools in each intervention arm (eight in the C/U intervention and eight in the control group). There were 164 pupils eligible for the trial following the baseline testing. The average number of eligible pupils per school was 10.3 students while the average number of pupils per class was 4.3 students.

Power calculations for the primary outcome are based on the estimates from the previous pilot trial of the C/U model. The student- and school-level correlations of the primary outcome measure (i.e., Oral Reading Fluency) between pre-test and post-test is assumed to be 0.735 and 0.261, respectively. The school-level Intraclass Correlation Coefficient (ICC<sup>25</sup>) is assumed to be 0.01. Class-level ICC is assumed to be 0.12 for this primary outcome.

The calculations were undertaken using 'PowerUp!'<sup>26</sup> and indicate that the pilot trial of each intervention has statistical power of 0.2 to detect an effect of 0.2 standard deviations for the primary analysis.<sup>27</sup> Since this is a pilot study, it is underpowered, and so there is a high probability that no statistically significant effect will be found.

---

<sup>25</sup> The ICC measures similarity between units in the same cluster; in this case, pupils within the same classroom. Units within the same cluster may exhibit similarities due to being exposed to similar environmental characteristics. This must be accounted for when conducting sample size calculations, since similarity between units reduces the amount of unique information each new observation contributes to the sample.

<sup>26</sup> Dong, N., & Maynard, R. (2013). PowerUp!: A tool for calculating minimum detectable effect sizes and minimum required sample sizes for experimental and quasi-experimental design studies. *Journal of Research on Educational Effectiveness*, 6(1), 24–67.

<sup>27</sup> Statistical power indicates the probability that there will be a statistically significant difference between the treatment and control groups given characteristics of a study, e.g., sample size and minimal detectable effect size. 0.8 is commonly used as a benchmark for a statistical power in fully-powered efficacy trials.

**Table A1: Power calculations**

		Catch-up (C/U)
MDES		0.2
Pre-test/ post-test correlations	level 1 (pupil)	0.735
	level 2 (class)	0.00
	level 3 (school)	0.261
Intraclass correlations (ICCs)	level 2 (class)	0.12
	level 3 (school)	0.01
Alpha		0.05
Power		<b>0.20</b>
One-sided or two-sided?		2
Average cluster size		10.3
Number of schools	Intervention	8
	Control	8
	<b>Total</b>	16
Number of pupils	Intervention	82
	Control	82
	<b>Total</b>	164

Given the low power and the fact that this is a pilot evaluation, we suggest placing more emphasis on identifying unforeseen problems in the interventions, implementation, and how pupils are evaluated, rather than on estimating effect sizes.

## Appendix C – Adapted EGRA tool with additional pre-literacy items

### EGRA+prelit -Reading Assessment Tool

#### Reading Skills Assessment Tool for Primary Stage: Evaluator Instructions 2023

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.]

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].

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 form*



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

 60 seconds



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."



The Early stop rule:

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.

☐ 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 "K"] is "K".

☐ Now let's do this exercise: tell me the sound of this letter [and point to the letter "L"]:

☐ ✓: Good, the sound of this letter is "L"

☐ ✗: The sound of this letter is "L".

☐ Let's try another example: Tell me the sound of this movement [point to the letter "Sa"]:

☐ ✓: Well done, the sound of this movement is "Sa"

☐ ✗: 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




👉 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.

👉 Place the mark (I) on the last letter the child reads.

Example: K L Sa

	10	9	8	7	6	5	4	3	2	1	
(10)	ك	ص	ظ	خ	هـ	ز	د	ل	و	ج	
(20)	هـ	ص	ظ	ش	ج	نـ	ف	حـ	بـ	نـ	
(30)	س	و	ة	ر	د	قـ	خ	ح	فـ	ا	
(40)	ق	ثـ	بـ	ضـ	يـ	ح	هـ	ذ	غـ	خـ	
(50)	ع	هـ	غـ	ة	ز	و	ق	خ	بـ	ضـ	
(60)	ظ	و	ص	هـ	ط	نـ	يـ	سـ	ذ	عـ	
(70)	ن	خـ	بـ	ق	غ	يـ	ثـ	د	ء	ثـ	
(80)	ج	ط	حـ	خ	ضـ	هـ	عـ	ش	ثـ	غ	
(90)	ثـ	لـ	عـ	ف	مـ	خ	ذ	ء	س	ز	
(100)	ف	ع	ضـ	سـ	ح	ثـ	ط	بـ	م	ذ	
Remaining time of exercise time (number of seconds):											
Check this box (X) <input type="checkbox"/> 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.											

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

☐\*: We read this syllable as "C"


☐ 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

(10)	ك	ص	ظ	خ	هـ	ز	د	ل	و	جـ
(20)	هـ	ص	ظ	ش	ج	تـ	ف	حـ	بـ	نـ
(30)	س	و	ة	ر	د	قـ	خ	ح	فـ	ا
(40)	ق	ث	ب	ضـ	يـ	ح	هـ	ذ	غـ	خـ
(50)	ع	هـ	غـ	ة	ز	و	ق	خ	بـ	ضـ
(60)	ظ	و	ص	هـ	ط	نـ	يـ	سـ	ذ	عـ
(70)	ن	خـ	بـ	ق	غ	يـ	شـ	د	ء	تـ
(80)	ج	ط	حـ	خ	ضـ	هـ	عـ	ش	ث	غ
(90)	ث	لـ	عـ	ف	مـ	خ	ذ	ء	س	ز
(100)	فـ	ع	ضـ	سـ	ح	ث	ط	بـ	م	ذ

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



The Early stop rule:

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.

### Section 3: Reading Invented Words

 60 seconds

🔊 These are some invented words. Read as many of them correctly as possible. Do not read letter by letter, but read the whole word. For example, this invented word is "alfallat".

🔊 Now read the following word: [Point to the word "shalameeth "]:

🔊✓: Well done, "shalameeth"

🔊✗: "Shalameeth " correctly, say: This invented word is "shalameeth "

🔊 Let's try another word now: Read this word [Point to the word "nasiba"]:

🔊✓: Very good, "nasiba"

🔊✗: This invented word is "nasiba"

When I say "start", read the words accurately and as quickly as possible. We will start from here and continue in this manner [Point to the first word in the first line, and follow with your finger the words in the entire first line.] Are you ready? Let's start.

🔊 Clearly mark (/) any mistake the child makes.

🔊 If the child corrects himself, put a circle around the (/) mark you previously made for him.

🔊 Put the mark (J) at the last word the child read.

Example: alfallat shalameeth nasiba

	5	4	3	2	1
(5)	تَحْمُ	أَمَشْنُ	را	تاري	ضا
(10)	مَحْبُ	سا	صَالِبُ	داف	دَفْ
(15)	جِهَا	صَالِدُ	قِمَاسِي	قاط	رِلْمُ
(20)	بُجِي	يَمَاجِي	قَبِيرُ	أَطِي	تَشْبِيرُونَ
(25)	أَحِي	شَاو	مَاصِي	فُدَاسَا	قَدَحْنُ
(30)	يَمَضُ	سَعِيمَةُ	دَلِي	أَشِبُّ	سِي
(35)	خَابَةُ	تَوَلْ	مِيهِ	عَاصِلُ	شَمَدَ
(40)	أَفَا	بَلُخُ	أَغِي	سَلْعَبُ	انْفَيْصَنَ
(45)	سَمُهُ	خَمَبُ	جُدُهُ	قَبِسَهُ	خَنَاءَ
(50)	سَحَتَ	فَعُ	عَنَيْسَمُ	نَبَرُ	أَقِي

🔊 Mark the remaining time from the exercise duration (number of seconds).

🔊 Mark (x) in this box ☐ if you stopped this part of the evaluation because the child did not read any of the words in the first line correctly.



**After 60 seconds, say 'stop'.**











If the child hesitates to read a word for more than 3 seconds, point to the next word and say: "Let's continue, please".



**Early Stop Rule:**

If you mark (/) all the answers in the first line because they are wrong and the child does not correct any of his mistakes, say "thank you" and stop the exercise. Mark (x) in the box at the bottom of the page and move on to the next exercise.

<p><input type="checkbox"/> This is a short story, focus well and read it correctly, aloud and as quickly as possible.</p> <p>When you're done, I'll ask you some questions about what you've read. Did you understand what is required of you? - When I tell you, "Let's begin."</p> <p>Start reading. ready? Let's Begin</p>	<p> Pull the text of the story in front of the child and ask them the questions below.</p> <p> Leave the child maximum 15 seconds to answer each question.</p> <p>Ask the question corresponding to each line the child has read until you reach the line with the mark ( J ), which indicates where the child stopped reading.</p>
<p> 60 seconds</p> <p> After 60 seconds, you will tell the child to 'stop.'</p> <p> If the child hesitates to read the letter for more than 3 seconds, point to the next letter and say: "Let's continue, please."</p> <p> <u>The Early stop rule:</u></p> <p>If you mark all the answers in the first line as wrong and the child does not correct any of his mistakes, say "thank you" and stop the exercise. Put an (X) in the box at the bottom of the page and go to the next exercise.</p>	<p><input type="checkbox"/> I will now ask you some questions about the story that I read. Answer the questions correctly.</p>


 Clearly put a tick (/) on any mistake the child makes. Put the mark (J) at the last word the child read.	 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 confidentially: 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			
⌚ Mark the remaining time from the exercise duration (number of seconds).				
⏏ Mark (x) in this box <input type="checkbox"/> if you stopped this part of the assessment because the child did not read any word in the first line correctly.				

### Section 5: Listening Comprehension


**The evaluator reads the following text aloud once only and slowly (about one word per second). Tell the child:**


 I will read you a short story aloud, only once. After that, I will ask you some questions. Please listen carefully and answer them correctly. Do you understand what is required from you?


"Abu Sa'id woke up early in the morning, energetic; to go to his farm. He had his breakfast, then wore his light cotton coat. And when he opened the door, he stopped saying: 'Subhan Allah! What a beautiful view! The land is a green carpet.' Abu Sa'id went back and woke up his sons, calling out: 'The crops have sprouted, come and look at the green herbs growing.' The sons got up happily and went out to the field. Then they gathered some flowers."

	No answer	Wrong answer	Correct answer
Who woke up early in the morning? <b>Abu Sa'id.</b>			
Where did Abu Sa'id want to go? <b>To his farm.</b>			
How did the sons get up? <b>Happily.</b>			
What did the sons gather from the field? <b>Some flowers.</b>			
In which season did the story take place? <b>In spring.</b>			

### Section 6: Silent Reading of Text

 This is a short story, focus well and read it correctly for two minutes. When you finish, I will ask you some questions about what you read. Do you understand what is required from you? When I say "Let's start", begin reading. Ready? Let's start.

 120 seconds

 After 120 seconds, say "stop".

Let the child read for 120 seconds and alert the child if they get distracted from reading.



Mark (X) in the box that corresponds to the child's answer, and then move on to the next question.

		No answer	Wrong answer	Correct answer
Saad went on a school trip to the zoo.	Where did Saad go? <b>To the zoo.</b>			

The students wandered into the garden and saw a monkey eating a banana.	What was the monkey eating? Banana.			
Saad thought of offering a banana to the monkey, he searched for his bag and didn't find it.	What did Saad think? To offer a banana to the monkey.			
Saad told the teacher that he lost his bag. The teacher smiled and pointed to his shoulder.	Why did the teacher smile at Saad? Because she found his bag on his shoulder.			
Saad laughed shyly and thanked her.	Why did Saad laugh shyly? Because his bag was on his shoulder.			

## 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. (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 21<sup>st</sup>-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).
- 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., Cumiskey, 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.
- Chall, J. (1996). Stages of reading development (2nd ed.). Fort Worth, TX: Harcourt Brace
- 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.
- Cortés-Albornoz, M. C. et al. (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>]
- EEF (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.  
<https://www.springer.com/gp/book/9789401785440>
- 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](https://pirls2021.org/wp-content/uploads/2022/files/P21_1-reading-achievement.pdf)]
- 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://earlygradereadingbarometer.org/files/EGRA%20in%20Jordan.pdf>]
- RTI (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.
- 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).