EXECUTIVE SUMMARY

LET'S READ FLUENTLY!
PILOT EVALUATION REPORT

Sashka Dimova, Hannah Woodbridge, Enes Duysak, Eliza Garwood, Julia Ruddick-Trentmann, Mansor Rezaian, Nedjma Koval, Natasha Phillips, Tien-Li Kuo, Andi Fugard, Julia Griggs, Daniel Phillips

مؤسسة الملكة رانيا
QUEEN RANIA FOUNDATION

10 May 2023

ABOUT THE EVALUATOR

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

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

The lead evaluator was Sashka Dimova.

ACKNOWLEDGMENTS

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

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

EXECUTIVE SUMMARY

I. PROBLEM: THE ISSUE OF DEVELOPING ARABIC LITERACY

A strong foundation in literacy is a crucial element predicting educational success. Evidence shows that early literacy difficulties can persist, limiting children's ability to achieve their potential (Brombacher et al., 2012). Pupils in the countries that use the Arabic language and script for instruction are performing at a low level in international and internal examinations (Eckert et al, 2020). For example, results using the Early Grade Reading Assessment (EGRA) conducted in Jordan since 2012 have shown that primary school-aged children are failing to reach reading comprehension benchmarks (RTI International, 2018). In addition, PISA results from 2018 estimated that 52% of Jordanian 10-year-olds are unable to read and understand a short age-appropriate piece of text. Existing evidence suggests that it is very unlikely that pupils will make up for learning loss during the next stages of their education, leaving these children at a significant disadvantage throughout their schooling and life (World Bank, 2019). Therefore, providing the right support in the early years of schooling is essential for reducing this 'performance gap'.

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

II. POTENTIAL SOLUTION: THE "LET'S READ FLUENTLY!" INTERVENTION

'Let's Read Fluently!' (LRF!) is an intervention that aims to support children in developing foundational literacy skills in Arabic through a practice-focused, phonics-based pedagogy and reading practice book. In the model, pupils are taught to process written text more quickly by firstly repeating individual letters and words to the point of automation. This is intended to enable them to decode reading faster, in order to read more fluently and free up working memory to recall important information and think critically. Time engaged in practice and

receiving timely feedback (namely, reinforcement and corrections) are seen as important predictors of reading ability.

Within 30-minute sessions, teachers follow an "I do, We do, You do" pedagogy, that entails:

- First, using large versions of the textbook, the classroom teacher introduces the letter-sound, or letter combinations, and models how to 'read' it. ("I do").
- Then students practice 'reading' together using either the choral or echo method ("we do"). These two steps are to be completed in the first 10 to 15 minutes of the session.
- Finally, learners independently work through the pupil practice book, sounding out letters or words with their finger following the text ("you do"). At this stage of independent pupil practice, the teacher's role is to encourage students and provide corrections.

The final, "you do" stage of independent practice with teacher feedback should be around 15-20 minutes, which is half to two-thirds of the session. This is a key feature of the LRF! model, as research in cognitive science indicates that individuals need to independently and repeatedly practice decoding to develop the automaticity needed for fluent reading (Abadzi & Martelli, 2014).

The **practice book** is also a key element of the LRF! model as it 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, including:

- New letter shapes are introduced slowly, one by one.
- Small font sizes negatively affect letter identification, so the book uses large font sizes and spacing.
- The Arabic script is dense and complex, and so it creates a higher cognitive load for new readers than other languages. The book and the LRF! model in general place importance on repetition and teacher feedback.
- It follows a phonics-based approach in which pupils gradually decode words using their phonics knowledge rather than using other clues or seeking help. This also means that the use of pictures is minimized to encourage pupils to learn letter sounds, rather than guessing.
- Pupils need to see meaning in text, so real words and sentences are introduced as soon as possible.
- The pupil practice book stresses repetition of patterns, alongside lots of practice in recognising them.
- Invented words are also included for each new letter that is introduced. Invented words give pupils the opportunity to practice phonics and to improve pupil's ability to recognise the most common sounds for letters.

All pupils receive a copy of the practice book and are expected to take it home for extra practice with their parents/carers. Teachers support this form of parental engagement in two ways: (1) by raising awareness through an introductory meeting with parents, and (2) by encouraging support of students' practice through communications via existing channels (e.g. WhatsApp messages to parents).

III. THE PILOT EVALUATION: BACKGROUND

This executive summary highlights the implementation and results of a pilot evaluation of LRF! conducted in Jordanian primary schools in 2021-2022. The purpose of the pilot evaluation was to test the intervention's feasibility and evidence of promise as well as assess its readiness for an efficacy trial. The pilot was also intended to provide preliminary evidence on the impact of LRF!, the mechanisms of change, and lessons to inform future scale-up.

The pilot addressed five research questions related to three key evaluation pillars. (See Table 1)

Table 1: Evaluation Pillars and Research Questions

Evidence of Promise	Feasibility of Intervention	Readiness for Future Trial
1. Is there preliminary evidence that LRF affects intended outcomes?	2. Was the intervention feasible to deliver? 3. What factors facilitate or impede success implementation?	4. What is the feasibility of a future randomized controlled trial (RCT)? 5. Are the outcome measures appropriate and feasible to measure?

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

IV. HOW WAS THE PILOT CONDUCTED?

In the pilot evaluation of LRF!, two different approaches of the intervention were delivered:

a **Whole Class (W/C) model**, delivered to Grade 1 pupils in their home classroom by their grade-level teacher, and
a **Literacy Catch-Up (C/U) model**, which targeted and was delivered in a resource room by a support teacher to the lowest-achieving 20% of pupils in grades 1, 2, and 3.

As the C/U model targets the lowest achieving 20% of pupils in a class, it is intended to be

delivered to groups of five-to-six pupils with similar literacy learning needs. In effect, this is a form of extra small group tuition, using the practice book as the learning material.

To assess LRF!'s potential impact and provide insights that can inform a future trial design, the evaluation was designed to:

- be an experimental study, with schools randomly allocated to one of the pilot arms or a control arm (W/C, C/U, or control).
- include a comprehensive implementation and process evaluation (IPE) using focus groups discussions with teachers, parents and pupils, interviews with school stakeholders, classroom observations, a teacher survey, and a pupil survey

Given that a small number of schools were participating in the pilot, the evaluation was not designed to detect significant differences between the treatment and control schools. Instead, the potential impact of LRF! was measured by: (1) students' Arabic literacy attainment based on the Jordanian Early Grade Reading Assessment (EGRA) and a set of pre-literacy items (hereafter EGRA+pre-lit), and (2) students' letter sound identification, speed and accuracy of word decoding and reading comprehension derived from the EGRA+pre-lit sub domains.

The pilot evaluation started in July 2021 with the recruitment, selection, and randomization of schools. The evaluation recruited a total of 24 schools (W/C: 8; C/U: 8; Control: 8) that met the eligibility criteria. After a one-day training of teachers and a three-hour orientation with principals and supervisors, both delivered by the QRTA, the models were implemented in selected schools. The **W/C model** was delivered to Grade 1 children in Semester 2 (spring) of the 2021-22 academic year, while the **C/U model** was delivered to children in Grades 2 and 3 in Semester 1 (fall) of 2021 and to Grade 1 in Semester 2 (spring) of 2022. (See Table 2)

Table 2: Implementation of LRF! Pilot Models (2021-2022)

	Semester 1 – Fall 2021	Semester 2 – Spring 2022
Whole Class Model	_	Grade 1
Catch-Up Model	Grade 2 and 3	Grade 1

The selection of pupils for the intervention was carried out by teachers, using the coarse grained diagnostic tool developed by the Early Grade Reading and Mathematics Program (RAMP)[1] - already in use in Jordanian classrooms. Overall, the delivery of LRF! was aligned with the existing curriculum content as much as possible so that pupils had a coherent learning experience.

Across the 24 selected schools, the pilot involved a total of 587 pupils (W/C model = 180; C/U model = 114; Control Group = 294 pupils). While following the "I do, We do, You do" pedagogy, teachers had some flexibility over how they facilitated the sessions. However, the content of

[1]For more information on RAMP see https://pdf.usaid.gov/pdf_docs/PA00THHW.pdf

each session was set. Teachers were also encouraged to draw on their professional judgment about tailoring instruction according to pupils' needs. Teachers were expected to ensure they adequately progressed through the practice book, while at the same time ensuring pupils were able to adequately master each 'lesson' as they did so. Meanwhile, pupils in control schools underwent baseline and endline assessments but received teaching as usual.

V. OUTCOMES MEASURES AND DATA SOURCES

The Impact Evaluation (IE) analysis had two components:

- **Primary outcome pilot analysis** - The primary outcome analysis followed an intention-to-treat (ITT) approach, which compares pupils assigned to the treatment and pupils assigned to the control group, irrespective of whether pupils assigned to the treatment group actually receive the intervention. The analyses used a three-level multilevel model to account for the clustering of pupils (level 1) in classes (level 2) and schools (level 3). The treatment assignment was at the school level (level 3). This model included school and class-level random effects and accounted for the baseline EGRA+pre-lit. A separate model was estimated for each intervention type (the W/C approach and the C/U approach).

- **Secondary outcome pilot analyses** – The secondary outcome analysis involved first providing summary statistics and an unadjusted mean difference between the intervention and control group for secondary outcome measures (i.e. oral reading fluency, letter sound identification, syllable identification, and reading comprehension). The analysis also followed the ITT approach, similar to that of the primary outcome analysis.

Meanwhile, the Implementation and Process Evaluation (IPE) used several different data collection tools and sources. (See Table 3.) Unless otherwise stated, all data collection was administered by Integrated's team of trained enumerators.

- **Training Attendance Data** – QRTA gathered attendance data for each LRF! training session.
- **Classroom Observations** – Classroom observations were conducted in all sections implementing either the W/C or C/U intervention. These observations assessed pupil engagement with LRF!, pupil use of the practice book, and the administration of the 'I do', 'We do', 'You do' approach. All sections were observed twice, once a few weeks after starting the intervention and once again shortly before the intervention finished.
- **Teacher Survey** – The teacher survey covered experiences of LRF! training and coaching, engagement with parents, and perceptions of LRF! and was completed by the teachers of both LRF! models.
- **Pupil Survey** – The pupil survey consisted of multiple questions assessing engagement and interest in reading, access and usage of the internet, and reflections on LRF! for intervention pupils at the endline. All pupils participating in the evaluation, including those in the control

- arm, completed the pupil survey.
- **Focus Group Discussions** (FGDs) – FGDs with teachers, parents, and coaches took place with several stakeholders across all three arms of the intervention.

Table 3: Data Tools and Sources Utilized for the LRF! Pilot

LRF Schools	Business-as-Usual Schools	Coaching and Training Data
Classroom observations Teacher survey Pupil survey Teacher FGDs Parent FGDs	Pupil survey Teacher FGDs Parent FGDs	Training and attendance data Coaches FGDs

VI. FINDINGS

1. Evidence of Promise

Highlights

- The evidence from the impact evaluation suggests that delivery of the W/C approach could improve pupils' literacy attainment. In the case of the C/U approach, there is no evidence of improvement. However, the pilot RCT was small and was not designed to measure impact robustly.

- Perceived outcomes, as reported by teachers, coaches, and parents indicate improvement in pupils' literacy, engagement and confidence with reading for the W/C. Evidence for the C/U approach was mixed. Evidence also suggested some signs of a potentially adverse psychological impact on Grade 1 pupils in the C/U model.

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

Findings for the **C/U approach** were more mixed. The data indicated no differences in pupils'

literacy attainment between schools that implemented the C/U model and control schools. This was the case for pupils across all three grades. However, there is an indication that the C/U model may have been more suitable for children in higher grades. Some teachers observed improvements in pupil performance and confidence in reading, particularly amongst Grade 3 pupils in C/U schools. Coaches, like teachers, also observed more positive impact for pupils in Grades 2 and 3 but noted that for Grade 1 C/U pupils, the material was too difficult for them. Parents and pupil reports also suggest that LRF! could have been too difficult for children in C/U Grade 1.

Still, the findings from the evaluation should be treated with caution because, due to the small sample size, secure conclusions cannot be drawn from the impact estimates. Any difference in outcomes cannot be interpreted as being directly attributable to LRF!, but instead as showing indicative evidence of promise.

2. Feasibility of Intervention (in terms of implementation)

Highlights

- LRF! was feasible to deliver, and most schools delivered LRF! as intended with minor adaptations. These adaptations mainly included allowing more time for delivery through lengthening LRF! sessions, adjusting the pace for pupils with lower literacy ability, or providing individual coaching to C/U pupils who struggled with the content.

- The pilot identified a number of potential improvements needed to the interventions, including modifications to the training for C/U resource room teachers, the format for LRF! delivery, the content of the practice book, and the selection process for C/U pupils.

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

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

- **Training and coachin**g – Evidence suggested that resource room teachers may require additional support and thus adaptations to the training materials for C/U resource room teachers and additional coaching is recommended.

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

- **Content of the practice book** - Even though the practice book was seen as a useful aid, and all sessions were based on the book, the perceptions on its content were less positive. Parents, coaches, and teachers from both intervention arms perceived the material to be too difficult for some pupils in Grade 1 or those struggling with literacy.

- **Diagnostic tool** – For the C/U model, the majority of teachers and coaches did not think the diagnostic tool was able to identify the lowest performing pupils in Grade 1 or to account for learning difficulties. Coaches and teachers also voiced concerns about the test environment of the tool proving unsettling for some pupils in both intervention arms. Thus, feedback suggests that there is a need to reconsider the process and testing environment for selecting pupils. However, the results from the baseline EGRA assessment indicate that pupils who were selected for the C/U model based on the diagnostic tool had significantly lower EGRA scores in comparison to children in W/C. This suggests that the diagnostic tool was successful in selecting children with lower literacy skills.

3. Assessing the Feasibility and Readiness of LRF! for a Future Efficacy Trial

Highlights

- Evidence in relation to the evaluation procedures indicates that a clustered RCT design with allocation to the school level will be suitable for a future scaled evaluation.

- There were no significant challenges in relation to recruitment, randomization, and retention of schools. However, two notable stumbling blocks were a lack of up-to-date contact information for schools as well the fact that two of the eight C/U schools dropped out of the pilot after baseline data was taken. These obstacles were overcome but should be considered when designing a future trial.

- The inconsistencies in evidence of promise and the suggestions for improvements in relation to the feasibility of C/U suggest that further development is required before C/U is reassessed for readiness for trial. Particular attention must be paid to the training for resource room teachers and to adaptations needed for small group settings with lower ability students.

Overall, the results of the LRF! pilot indicate the intervention in its W/C form is likely feasible at the level of a full-scale efficacy trial. Some of the key evaluation procedures such as recruitment and randomization were executed well during the pilot study, suggesting that a clustered RCT

design with allocation at the school level will be feasible in the future. However, it was easier to recruit schools in the near south region in comparison to (1) schools in Amman, (2) the middle region excluding Amman, and (3) the near north region. Recruitment for any future evaluation should consider these regional differences. Finally, retention rates of schools were high for both the W/C and C/U approaches, and the outcome measures were seen as appropriate to identify progress in literacy.

VII. BARRIERS

Three barriers affected the implementation of both the W/C and C/U models as well as the teaching and learning in control schools. (See Table 4)

Table 4: Barriers Impacting the LRF! Pilot Conclusion

Barrier	Impact
Covid-19	• Due to the Covid-19 school closures prior to the pilot study, Grade 1 pupils had missed their kindergarten education. This made LRF! challenging for Grade 1 pupils as they lacked the foundational skills and knowledge. • Pupil absenteeism due to sickness was high during the pilot study. • To contain Covid-19, the government unexpectedly closed schools in late December (instead of mid January) for the winter break, which was extended until February. As a result, the C/U pilot with grades 2 and 3 was interrupted, with its final sessions and endline testing occurring after the winter break.
Staffing Challenges	• Teachers and coaches mentioned staffing challenges including absences of teachers or changes in school leadership. • Varying professional skills and experience among teachers.
Parental Engagement	• Coaches and teachers stated parental engagement with schools was very limited, with only a small number of parents interacting with teachers. • Various reasons for low engagement were given, including parental illiteracy, lack of capacity from parents, and perception that the school is responsible for their child's education.

VIII. CONCLUSION

While this evaluation provides evidence to suggest that implementation of the W/C and C/U approaches was successful, it also identified some recommended intervention adaptations or conditions that are needed to make LRF! succeed in an efficacy trial. Overall, there is indicative evidence that W/C is ready for trial conditional on modifications in respect to the time allocated for LRF! delivery, and in respect to the content of the practice book. While the delivery and evaluation partners felt that they have the capacity and knowledge to deliver and evaluate a scaled-up version of LRF!, they also felt that allowing sufficient time for adaptations will be key to the success of the evaluation. The pilot found that the C/U approach would require substantial changes before it can be reassessed for readiness for trial.

IX. REFERENCES

Abadzi, Helen. (2017) "Improving Students' Academic Achievement: The Crucial Role of Rapid Reading and Grammar Mastery in the Early Grades." Policy Paper No. 20. Sheikh Saud bin Saqr Al Qasimi Foundation for Policy Research. https://doi.org/10.18502/aqf.0047

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

Brombacher, A. et al. (2012) EdData II: National Early Grade Literacy and Numeracy Survey–Jordan Intervention Impact Analysis Report. Research Triangle Park, North Carolina, US: Research Triangle Institute (RTI) & USAID.

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

RTI International. (2018). 2018 Early Grade Reading and Mathematics Initiative (RAMP) Lot Quality Assurance Sampling Assessment (p. 9). RTI International.

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