Evaluation of a pilot programme to promote child-centred early childhood education and reduce over-age enrolment in Liberia

Early Learning Partnership Phase 2: Liberia

Donny Baum, Jacobus Cilliers, David Jeffery, Gloria Olisenekwu, Rachel Outhred, and Yvonne Capehart Weah

December 2021

Oxford Policy Management

About Oxford Policy Management

Oxford Policy Management (OPM) is committed to helping low- and middle-income countries achieve growth and reduce poverty and disadvantage through public policy reform.

We seek to bring about lasting positive change using analytical and practical policy expertise. Through our global network of offices, we work in partnership with national decision makers to research, design, implement, and evaluate impactful public policy.

We work in all areas of social and economic policy and governance, including health, finance, education, climate change, and public sector management. We draw on our local and international sector experts to provide the very best evidence-based support.

Oxford Policy Management Limited
Registered in England: 3122495

Level 3, Clarendon House
52 Cornmarket Street
Oxford, OX1 3HJ
United Kingdom

Tel: +44 (0) 1865 207 300
Fax: +44 (0) 1865 207 301
Email: admin@opml.co.uk
Website: www.opml.co.uk
Twitter: @OPMglobal
Facebook: @OPMglobal
YouTube: @OPMglobal
LinkedIn: @OPMglobal

Preface

This evaluation was carried out by OPM, with support from BRAC Liberia and the We-Care Foundation. The project manager was David Jeffery. The remaining team members were Femi Adegoke, Martina Garcia Aisa, Ekundyao Arogundade, Donny Baum, Jacobus Cilliers, Brian Law, Kirsty McLaren, Gloria Olisenekwu, Georgina Rawle, Ajala Tayo Stephen, and Yvonne Capehart Weah. Rachel Outhred and Oxford MeasurEd supported the analysis of the learning assessments, and prepared the exercise to set learning benchmarks. Mark Minford provided oversight of the cost sections of this report. Alina Lipcan and Zara Majeed provided support during the inception phase.

We are grateful for the support of the ECE Bureau in the Ministry of Education, and the support of the ministry as a whole. We are indebted to the principals, teachers, and children who assisted with this evaluation.

Executive summary

Context

Liberia has delivered substantial developments in early childhood education (ECE) over the last decade. Following the introduction of ECE in 2011, the Ministry of Education (MoE) has directed the implementation of multiple programmes and activities to improve access to quality ECE in Liberia. A key aspect of these efforts is a clear understanding of the systemic challenges in improving ECE in Liberia, and the identification of effective interventions that can be scaled within this context.

The Early Learning Partnership (ELP) is a multi-donor trust fund, managed by the World Bank, which works with countries to improve early learning opportunities and outcomes for young children, through both research and operational support. The ELP is concerned with the political economy and governance of early learning service provision. Specifically, the programme seeks to generate knowledge on what works to remove institutional and systemic barriers to the provision of ECE at scale, and on what reforms could create the right institutional incentives to improve ECE provision.

Oxford Policy Management (OPM) is managing the ELP research in Liberia, in partnership with the MoE. In Phase 1 (2017–18), we undertook an early learning system diagnostic that has provided insights into the early learning context in Liberia. This diagnostic identified two key obstacles to improving learning outcomes for ECE: the high prevalence of over-age children in ECE classrooms, and low levels of teacher training in ECE (specifically in child-centred pedagogies and use of the national curriculum). In ELP Phase 2 (2019–20), we evaluated a pilot of an intervention to address these two obstacles.

In August 2020, ELP Phase 2 was extended (across all ELP countries) in order to include research on the impact of the COVID-19 pandemic on ECE systems. The extended study explored the response of the MoE and other education stakeholders in seeking to mitigate the effects of COVID-19 on the education system, and the ECE subsector in particular, as well as the impact on ECE schools and teachers. The MoE's primary initiative to maintain continuity of learning during school closures was a Teaching by Radio (TbR) programme, although in interviews conducted for this study stakeholders noted that children in ECE would be unlikely to access such programming effectively.

This report provides the evaluation of the ELP Phase 2 intervention, and also provides a summary of the extended research on the MoE's response to the pandemic. The full report on the MoE's response to COVID-19 is available separately.

Intervention

In response to the findings from ELP Phase 1, the MoE, World Bank, and the research team developed terms of reference for an intervention that would provide workshop- and workplace-based training for teachers. Preschool teachers received support on child-centred pedagogies and use of the national ECE curriculum, while teachers supporting over-age children recently promoted from ECE received support in an Accelerated Learning

Programme (ALP). In addition, schools receiving the intervention were encouraged to promote children of primary school age from ECE to primary school, in line with national legislation.

This programme was delivered by a service provider (BRAC Liberia) selected in partnership with the MoE. The intervention for teachers was initially intended to be made up of (a) an initial workshop of five days; (b) two two-day refresher workshops; and (c) 8–12 school-based mentorship visits. Moreover, it was anticipated that teachers would deliver what they had learned in their classrooms for nine months (November 2019 to July 2020). However, the intervention design was significantly disrupted by school closures in March 2020, due to the COVID-19 pandemic. Consequently, teachers received only one refresher workshop and an average of two to three mentorship visits (in addition to the original five-day workshop). Moreover, teachers delivered the programme for only 4.5 months (November 2019 to March 2020).

Evaluation design

The design of this evaluation applies procedures of (i) random sampling, to ensure generalisation of the findings to the larger population of schools in the selected counties; and (ii) random assignment of sampled schools to treatment and control conditions, to provide a causal estimation of the treatment impact. According to this design, schools assigned to the control condition would operate without any changes to their ECE curriculum, pedagogy, and their approach to managing over-age students. In contrast, over the course of the 2019/20 academic year, schools assigned to the treatment condition would implement the ECE curricular and over-age student interventions described above. In total, 27 treatment schools and 27 control schools in Gbarpolu and Bomi countries participated in the study.

Students in both treatment and control classrooms were assessed prior to the initiation of the intervention (i.e. at programme baseline) in the areas of literacy, numeracy, and executive function. Following completion of the intervention (i.e. at programme endline), students were once again assessed in these areas, with differences in baseline-to-endline growth between treatment and control students representing the impact of the intervention on student learning. Additionally, teacher practices were measured at both baseline and endline to determine the impact of the programme on teacher pedagogy, teacher–child interactions, and classroom management. A survey of school principals collected information on receptivity to the intervention and implementation costs.

787 children, 82 teachers, and 54 principals were assessed and surveyed at baseline (November 2019) and endline (February 2021). This reflects a retention rate of 79% for children and 88% for teachers.

Key findings

The key findings of the evaluation were as follows:

- COVID-19 has had a significant impact on the fidelity of the intervention; chiefly, the number of mentorship visits was reduced from 8–12 to 2.8 per teacher.

- The intervention increased uptake of the national ECE curriculum, but this did not translate into changes in teachers' beliefs about child-centred learning or use of child-centred pedagogies.
- The intervention increased the proportion of otherwise over-age children promoted from ECE to primary school and reduced the proportion of over-age children enrolled in ECE. However, the intervention did not increase teachers' and principals' confidence in meeting the needs of over-age children, and teachers in treatment classrooms actually expressed lower confidence than control teachers in the likelihood that promotion would benefit over-age children.
- The intervention did not improve the learning outcomes in early literacy, early numeracy, and the executive function of ECE children. Nonetheless, learning outcomes improved in both treatment and control schools, despite school closures. Due to funding constraints it was beyond the scope of the study to investigate changes in learning outcomes for over-age children resulting from the intervention.
- The intervention costs on average US$ 14.96 per child (US$ 1,297 per teacher) if delivered according to the original design of one initial workshop, two refresher workshops, and eight mentorship visits, and assuming a student–teacher ratio (STR) of 30:1. This would be a significant addition to the current cost of ECE provision estimated in the 2016 Education Sector Analysis (ESA) (US$ 24 per child annually, but with an STR of 53:1), but would constitute a smaller share of the costs to the Government of providing ECE at the standard envisaged in national policy estimated in ELP Phase 1 (US$ 67 to US$ 223 per child annually, depending how much parents contribute to food and stationary).
- Despite the MoE's efforts to promote TbR during school closures, the majority of teachers (61%) were unaware of any distance learning activities offered to either children or families by their school, government, or any other organisations.

Recommendations

We offer the following recommendations regarding the ECE intervention and improving the resilience of the education system to future shocks.

Intervention

1. In addition to **improving access to the new ECE curriculum for teachers, complementary interventions will be needed to change teaching practices**. The uptake of the national curriculum by teachers and principals is unlikely to be sufficient to increase the use of child-centred pedagogies and to improve learning outcomes, without accompanying interventions such as training and support.
2. Any teaching training focused on changing teaching practices and teachers' attitudes towards child-centred pedagogies is likely to require a relatively **high dosage** to be effective. The eventual dosage used in this evaluation, which was limited by COVID-19, was likely too low to be effective, although further research would be needed in order to establish the efficacy of any higher dosage.

Emergency response

3. The MoE was correct to identify the need to ensure the continuity of salary payments as a priority. However, the success in achieving this continuity appears to be only partial. We recommend the **MoE prioritise making improvements in payroll administration in order to increase resilience to shocks** in the future.

4. Despite being aware of the limitations of TbR during the Ebola epidemic, the Education in Emergencies Technical Working Group (EIE TWG) saw no effective alternatives for responding to the COVID-19 pandemic at scale due to limited transport and communications infrastructure. **The MoE's ability to respond to future shocks is likely to be similarly constrained unless this infrastructure is improved.** To achieve this, development partners of the MoE may wish to prioritise funds for this purpose.

5. In addition to improving transport and communications infrastructure, **the MoE may benefit from enhancing its capacity to disseminate information during school closures effectively**. Otherwise, a lack of awareness of government activities will remain a barrier to the uptake of available services.

Further research

6. Despite the limited dosage and the major interruptions that were experienced there were still some indications of initial successes in the uptake of elements of the intervention. Hence an **additional study is recommended in order to assess the intervention in a situation without the interruption of the pandemic**, before taking any decision on scaling up the full intervention. Such a study may include an additional component on social behaviour change to target teachers' attitudes towards child-centred learning.

7. Given the success of the intervention in reducing over-age enrolment despite the interruption of the pandemic, the additional study may consider testing a **shorter intervention aimed specifically at reducing the number of over-age children in ECE**. This study should include an assessment of the impact of such an intervention on the learning outcomes of over-age children.

8. Since teachers' usage of the national curriculum did not correlate to a change in teaching practices, the additional study may consider the **inclusion of qualitative research to understand how teachers interpret and use the national curriculum**, and how teachers engage with training programmes that seek to change their attitudes towards child-centred pedagogy.

Table of contents

List of tables and figures

Tables

Figures

List of abbreviations

ALP	Accelerated Learning Programme
ECE	Early childhood education
ELDS	Early Learning Development Standards
ELP	Early Learning Partnership
EMIS	Education Management Information System
ESA	Education Sector Analysis
FCDO	UK Foreign, Commonwealth and Development Office
ICC	Intracluster correlation coefficient
IRT	Item response theory
MDES	Minimum detectable effect size
MELE	Measure of Early Learning Environments
MELQO	Measuring Early Learning Quality and Outcomes
MoE	Ministry of Education
OPM	Oxford Policy Management
SD	Standard deviation
STR	Student–teacher ratio
TbR	Teaching by Radio
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development
VPI	Vice Principals for Instruction
WASSCE	West African Senior School Certificate Examinations

1 Introduction

1.1. Context

The Early Learning Partnership (ELP) is a multi-donor trust fund, managed by the World Bank, which works with countries to improve early learning opportunities and outcomes for young children, through both research and operational support. The ELP is concerned with the political economy and governance of early learning service provision. Specifically, the programme seeks to generate knowledge on what works to remove institutional and systemic barriers to the provision of early childhood education (ECE) at scale, and on what reforms could create the right institutional incentives to improve ECE provision.

With the support of the UK Foreign, Commonwealth and Development Office (FCDO), the ELP has launched the ELP Systems Research programme. The programme's objectives are (1) to provide policymakers in a set of focus countries with actionable information to help guide the delivery of quality, equitable early learning at scale; and (2) to build the international evidence base in the emerging field of systems research in ECE. The programme is supporting one research team in each of the following focus countries/regions: Ethiopia; Jamaica; Liberia; the province of Punjab, Pakistan; and Tanzania.

Oxford Policy Management (OPM) is managing the ELP research in Liberia. In Phase 1 (2017–18), we undertook an early learning system diagnostic that has provided insights into the early learning context in Liberia. This diagnostic identified two key obstacles to improving learning outcomes for ECE: the high prevalence of over-age children in ECE classrooms,[1] and low levels of teacher training in ECE (specifically in child-centred pedagogies and knowledge of the national curriculum). In ELP Phase 2 (2019–20), we evaluated a pilot of an intervention that sought to address these two obstacles. In August 2020, ELP Phase 2 was extended (across all ELP countries) in order to include research on the impact of the COVID-19 pandemic on ECE systems. The extended study explored the response of the government and other education stakeholders in seeking to mitigate the effects of COVID-19 on the education system, and the ECE subsector in particular, as well as the impact on ECE schools and teachers.

This report details the evaluation of the pilot ECE interventions under ELP Phase 2, and includes the impact of COVID-19 on teachers and principals in piloted schools. In Section 1.3, we summarise key findings from ELP Phase 1 that provided the basis for the design of the intervention that was evaluated in this report. Subsequently (in Section 1.4), we outline the response of the Government and other education stakeholders in seeking to mitigate the effects of COVID-19 on the education system, and the ECE subsector in particular, in order to provide the context for this study's findings on the effects of the pandemic.

[1] For the purposes of ELP Phase 1 and Phase 2, 'over-age' was defined as being enrolled in ECE but being of primary school age (i.e. six years old or older). While the MoE defines 'over-age' for primary and secondary school as being 'greater than 3 years older than the appropriate age level for a grade', the definition used in the ELP is appropriate for ECE. This is because while primary school is compulsory, ECE is not. Consequently, a child of seven years in ECE is legally required to be in primary school.

1.2. Policy context

The policy context in respect of ECE in Liberia has developed substantially over the last decade (Kim *et al.*, 2022). In 2011, the Education Reform Act established a Bureau of ECE within the Ministry of Education (MoE), and mandated that children aged three to five years should attend ECE for at least two years. However, ECE was not made compulsory, and schools were permitted to charge fixed fees for ECE, although primary school remained fee-free. In 2012, the National Inter-Sectoral Policy on Early Childhood Development provided a framework for managing the ECE system in Liberia and set targets to expand access to ECE and develop a ECE training framework. In 2018, the Pro-Poor Agenda for Prosperity and Development set targets for the roll-out of a standardised curriculum and ECE training and for reducing student–teacher ratios (STRs) below 50:1 in all regions. The Pro-Poor Agenda also identified the need to monitor the delivery of ECE to ensure compliance with the curriculum and quality standards.

In line with these policy changes, multiple programmes and activities have been implemented to improve access to quality ECE in Liberia. Over the past decade the Bureau of ECE has led initiatives to improve awareness of and understanding of the importance of ECE, such as the Early Childhood Development and Community Education and Awareness Program. A new specialised ECE teacher qualification (the ECE C Certificate) has been introduced. Improving access to quality ECE was one of the nine core strategies of the Getting to Best Education Sector Plan 2017–21 (MoE, 2016b). During 2020, the MoE also developed draft Early Learning Development Standards (ELDS), which proposed benchmarks for early childhood learning which may be used in the future to support ongoing system strengthening.

In support of this agenda, a variety of projects have been implemented by partners working with the MoE. Among others, these have included system-building activities supported by the Open Societies Initiative for West Africa, training unqualified ECE teachers as part of the GPE Getting to Best project, and the Read Liberia kindergarten reading programme. Programmes have also been implemented to address over-age enrolment across the education sector. These include the Accelerated Quality Education project and other projects targeting out-of-school children. The goals of improving ECE and supporting age-appropriate enrolment have also been pursued through multiple other programmes, including programmes of the United Nations Children's Fund (UNICEF) and the Liberian Educational Advancement Programme, among many others.

1.3. Key findings from ELP Phase 1

In Phase 1 of the ELP System Research Programme, each country research team conducted a diagnostic of the early learning systems in their country. In Liberia, we identified two particularly severe challenges: low-quality ECE provision, and high prevalence of over-age enrolment in ECE classes. Our overall assessment of the early learning system at that time was as follows:

- The system was partially aligned to promote quality ECE in Liberia, with insufficient resources being the severest constraint. Although most children enrolled in ECE were over the age of six, most children could only successfully complete the easiest

assessment tasks in a test aimed at children between three and five years old. Moreover, teachers were found to engage in predominately rote teaching, with little time spent on child-centred activities.

- There were aspects of the ECE system that had the potential to contribute to quality ECE. The national government was supportive of ECE and had developed a national curriculum, and there were at least basic formal processes for gathering data on ECE and monitoring the performance of schools. Many parents engaged with their children's schools, and most principals had a teaching qualification and supported their teaching staff.

- However, the impact of these features was undermined by limitations elsewhere in the ECE system. Most significantly, schools were under-resourced and the vast majority of ECE teachers surveyed had no qualifications in ECE. Additionally, the MoE was reportedly disconnected from county and district governments, which in turn had limited resources and staff capacity. Principals appeared to have limited ability to sanction underperforming staff.

- There was little alignment in the system to promote age-appropriate enrolment in ECE. Although there was a national policy governing age-appropriate enrolment, it was unclear to what extent it was prioritised by government over competing concerns, and it was followed by only a minority of schools, and it was not adhered to by parents.

It is important to note that ELP Phase 1 was conducted in 2017. Many efforts to improve the ECE system have been implemented since then, including initiatives to expand teacher training. This summary therefore provides the context for the development of the Phase 1 intervention in 2019, rather than providing an updated account of the ECE system at the time of writing.

Section 2.1 links these findings to the objectives of the intervention piloted in this evaluation. A full report of the Phase 1 findings is also available online (OPM, 2019).

1.4. Key findings from a study on the COVID-19 response

COVID-19 presents a severe threat to ECE systems internationally, both through the near universal closing of schools and through the effects of an economic recession. An extension of the ELP Phase 2 study explored how the MoE has responded to the crisis, and, in particular, the relative prioritisation of ECE in this response. The findings were based on key stakeholder interviews with representatives from the MoE and development partners, as well as an analysis of government publications and press releases from 16 March 2020 to 31 January 2021 (OPM, 2021).

The MoE played an integral role in the coordination of the response to COVID-19 within the education sector, through convening and chairing an Education in Emergencies Technical Working Group (EiE TWG). The MoE responded quickly to the COVID-19 crisis: the first case of COVID-19 in Liberia was registered on Friday 13 March 2020, and schools were closed on Monday 16 March.

The MoE's primary response to maintaining continuity of learning school closures was the provision of a Teaching by Radio (TbR) programme, with different programmes for children of all ages. The equity limitations of such an approach were recognised, but this was believed to be the only viable intervention to maintain some measure of learning during school closures. In June 2020, the MoE directed a staggered approach to resuming classes after it was announced that the West African Senior School Certificate Examinations (WASSCE) would be held later in the year. Resumption of classes began with students in Grade 12, and was progressively extended from Grades 6 to 11. Children in ECE to Grade 5 were asked to remain at home, as it was widely believed that young children would not be able to comply with the necessary health protocols. Teachers for children in ECE to Grade 5 were asked to prepare homework assignments. All classes were directed to open for the new academic year on 1 December 2020. Schools opened at different times between December 2020 and February 2021, and in January and February 2021 the MoE increased efforts to distribute health and hygiene equipment to schools.

ECE appears to have received limited attention during the MoE's response to COVID-19. Interviewees in this study recognised that the effectiveness of the TbR programme for children in ECE was likely to be limited. Nonetheless, TbR was considered the only viable intervention during school closures. The WASSCE provided the impetus for schools to reopen, but these exams were not applicable to young children. Interviewees reported that this is a continuity of the lack of prioritisation of ECE prior to the pandemic, although the MoE disagreed with this assessment when reviewing a draft of this report.

Three factors were likely key in shaping the MoE's response to the COVID-19 pandemic. The MoE deferred to the instructions provided by the Ministry of Health (MoH) and the National Public Health Institution of Liberia, who were critical in determining when and how schools would reopen. A lack of funding within the MoE was also an important factor; although the MoE published an Emergency Response Plan (ERP), this required a budget of $32 million of which reportedly only half had been raised from donors at the time of writing the original report on the COVID-19 response (February 2021). This meant that many activities were not delivered, and that the selection of activities was determined in part by donors' sectoral interests. Finally, a lack of infrastructure reduced the range of viable responses to the pandemic. Long distances between schools, difficult terrain and poor-quality roads, and rising fuel costs provided substantial challenges to the distribution of resources and the delivery of in-person programmes. Interviewees in this study were aware that limited and inequitable access to radio was a significant limitation of the TbR programme.

1.5. Structure of this report

This report is structured as follows.

- Section 2 details the objectives of the study, and the research questions;
- Section 3 details the intervention being evaluated, including the selection of the service provider;
- Section 4 details the development of learning benchmarks in collaboration with the MOE, which form the basis of our analysis;
- Section 5 details the evaluation methodology;

- Section 6 reports the findings of the evaluation; and
- Section 7 summarises the findings against the research question, and discusses the implications.

This report is accompanied by a technical annex, providing further detail on the methodologies for the evaluation (Annex B) and the development of the learning benchmarks (Annex C).

2 Objectives

2.1. Objectives of the research

The objective of this study was to evaluate an intervention that addresses key misalignments in the ECE system identified during Phase 1. Specifically, these misalignments were that[2]:

- Although a national ECE curriculum has been developed, it is used in only a minority of schools (in the Phase 1 sample[3], 31% of principals reported using the national curriculum in particular, and 17% of ECE teachers reported using any curriculum at all);
- Although the national ECE curriculum prescribes play-based and child-centred learning, ECE teachers have not been provided with ECE-specific training in this regard (80% of teachers had no such training);
- Although children above the age of 6 should be enrolled in primary school, most schools have large numbers of over-age children enrolled in ECE (52% of children enrolled in ECE were 7-years or older);
- Although over-age enrolment in ECE has been identified as a challenge by the MoE, very few schools actively support over-age learners or feel equipped to do so (50% of principals and 76% of ECE teachers reported not offering additional support, and 80% of these principals said this was because they did not have the resources and 43% said that they did not know how to provide this support).

In order to address these misalignments, a two-pronged intervention was identified for Phase 2. This intervention entailed:

1. A pre-primary programme to improve the quality of ECE instruction in order to better prepare children for starting school;
2. A primary school programme to assist teachers support the learning of over-age children who will be moved from the pre-school to primary school[4].

Two reasons were primarily considered for including both prongs of the intervention. First, children in ECE are likely disadvantaged by absence of child-centred and play-based learning activities, and the presence of significantly older children in their classes. This is because teachers may struggle to teach to children with a wide range of abilities and needs (especially if not trained to do so), and, moreover, may default to 'rote teaching' if this was believed to be necessary to serve older children. This means that it was necessary to both

[2] Note that these findings are based on data drawn from randomly selected schools, teachers, and students from two counties in Liberia, and so should be considered indicative but not necessarily nationally representative.
[3] All percentages in this section refer to findings from the Phase 1 sample.
[4] In the original evaluation design, it was envisaged that this intervention would be 'Teaching at the Right Level'. However, as described in Section 3, in consultation with the MoE, an 'Accelerating Learning Programme' (ALP) that was currently in use was selected instead.

provide training in child-centred and play-based ECE, and to encourage schools to promote over-age students to primary school.

Second, there was a concern about the potential for unintended negative effects if over-age children were promoted to primary school without additional support being provided to primary school teachers to accommodate these students. This necessitated the inclusion of training for primary-level teachers as part of the intervention.

2.2. Research questions

Table 1 outlines the research questions as originally proposed for the ELP Liberia Phase 2 study. In addition, three supplementary research questions were added to understand the government response to, and the impact of, COVID-19 on school operations.

Table 1: Research questions

What is the impact of the revised ECE programme compared to existing ECE provision?
1. What impact do professional training for ECE teachers, child-centred and play-based learning, and classes being restricted to age-appropriate children have on the cognitive outcomes of children who are of the appropriate age to attend ECE (i.e. between the ages of three and six, inclusive)? Specifically, what is the impact of the intervention on children's (i) early literacy, (ii) early numeracy, and (iii) executive function in particular? 2. Does the revised ECE programme increase the uptake of the national ECE curriculum?
What is the impact of the intervention on the number of over-age children enrolled in ECE?
3. Does the intervention reduce the number of over-age children enrolled in ECE? 4. What proportion of over-age children are promoted to primary school?
What is the cost of the revised ECE programme and primary-level intervention compared to existing ECE provision?
5. What is the cost of training teachers in the new programme and providing them with the necessary resources (e.g. school visits, school/classroom volunteers)? 6. Will the programme entail any additional reoccurring costs, such as the hiring of new staff (e.g. volunteers used in the intervention, additional teachers for the movement of over-age students to primary), the development of materials, and the increased engagement with the community/parent–teacher associations?
Are teachers and principals likely to be receptive to receiving the two parts of the intervention if it is scaled up?
7. Are principals and teachers receptive to receiving additional training and support? 8. Are principals and teachers receptive to promoting over-age children to primary school? 9. Through these interventions, do principals and teachers feel equipped to use the national ECE curriculum? 10. Through these interventions, do principals and teachers feel equipped to support over-age children?
Additional research questions (ARQs): COVID-19 response
ARQ1: What steps has the MoE taken in response to COVID-19? What has shaped this response? ARQ2: What has the effect of COVID-19 been on student re-enrolment? ARQ3: How has COVID-19 affected ECE teachers' ability to continue to provide play-based, child-centred ECE in schools after the end of lockdown? What has the effect been on teacher retention, especially for 'volunteer' teachers?

Research question ARQ1 ('What steps has the MoE taken in response to COVID-19?') has been reported on separately, and is summarised in Section 1.4. ARQ2 and ARQ3 are addressed in the current report, along with the 10 original research questions.

3 Intervention

3.1. Intervention description

In summary, the diagnostic review of the ECE system in Liberia undertaken as part of ELP Phase 1 identified two critical challenges. First, although a child-centred national ECE curriculum had been developed, most preschool teachers were found to not be using it; and, indeed, only a minority had any formal training in ECE, with fewer having specific training on student-centred and play-based learning. Second, the majority of children in preschool classes were found to be of primary school-going age. Together, these challenges hinder teachers' ability to deliver quality, age-appropriate ECE.

The focus of ELP Phase 2 was the evaluation of a pilot to address these challenges. The MoE, the World Bank, and the research team developed terms of reference for an intervention that it was expected would provide the following:

- workshop- and workplace-based training to preschool teachers in child-centred pedagogies using the national curriculum;
- workshop- and workplace-based training to primary school teachers on how to support over-age children in their class; and
- encouragement to schools to promote children of primary school age from ECE to primary school, in line with national legislation.

The intervention's design was refined following consultations during development of the teaching and learning materials, and during piloting of the intervention. The initial intervention design is detailed below; the refinements are discussed in Section 3.3.

The intervention for preschool teachers consisted of the following:

- Developing teaching and learning materials, including play-based and child-centred lesson plans and supplementary materials, based on the national ECE curriculum framework.
- Training of preschool teachers in a five-day workshop, covering play-based learning, classroom management, and lesson planning.
- Providing preschool teachers with a lesson plan book, class registers, ECE supplementary reading books, and other teaching and learning materials.
- Conducting school-based mentor visits to each teacher, and conducting refresher training workshops.

The intervention for ECE teachers provided teacher planner workbooks, a blank lesson plan workbook, and a teacher guide based on the ECE curriculum. These materials had been previously developed by BRAC Liberia, with the support of UNICEF, and were approved by the ECE Bureau.

The intervention for teachers supporting primary school-aged students consisted of the following:

- Developing an instruction kit, focused on basic skills in mathematics and English for children of diverse ability without prior primary school education, based on the national curriculum.
- Training of primary school teachers in a five-day workshop, covering pedagogy and classroom effectiveness, with a focus on core subjects: English, mathematics, social sciences, and science.
- Providing each teacher with a teaching guide, other teaching and learning materials, and learning assessment tools.
- Conducting school-based mentor visits to each teacher, and conducting refresher training workshops.

Following consultations with the MoE's Bureau of Basic and Secondary Education, it was agreed that the Accelerated Learning Programme (ALP) was the most appropriate teaching programme to use for the primary school-level intervention. The materials used for these purposes were from Level 1 of the Accelerated Quality Education project, which had been developed by the MoE, United States Agency for International Development (USAID), and UNICEF.

The interventions were piloted in three schools in Grand Cape Mount county, which were not enrolled in the primary study. The pilot is described in Section 3.3 and the dosage of these interventions after the pilot is detailed in Section 3.6.

3.2. Selection of service provider

OPM and the MoE conducted an open bidding process for the service provider for this programme. In appointing a service provider, we undertook the following process:

- Terms of reference were prepared collaboratively, and advertised on the ministry's website.
- Four proposals were received in total.
- These proposals were graded individually by four OPM team members and by a committee within the MoE.
- An initial round of interviews was conducted by OPM with three of the service providers.
- The MoE and OPM agreed to shortlist two service providers, who were invited to present their proposals to a joint committee at the MoE in Monrovia.
- Following this process, BRAC Liberia was selected as the service provider for this intervention. This was on the basis of their demonstrable expertise in ECE in Liberia, and their history of collaborating with the MoE on ECE interventions using the national curriculum. Due to circumstances imposed by the COVID-19 pandemic,

BRAC Liberia was unable to provide the final refresher training and debrief to teachers. With approval from the MoE, this workshop was instead delivered by the We-Care Foundation.

3.3. Pilot

The intervention was piloted with ECE teachers from three schools in Grand Cape Mount county. Grand Cape Mount neighbours the counties where the full intervention would be carried out (Bomi and Gbarpolu counties), and, like the intervention counties, is also rural. The pilot comprised delivery of ECE training, delivery of teaching and learning materials, and a short series of mentoring visits.

Following the pilot, components of the intervention were refined following consultations between BRAC, MoE, and OPM. Due to concerns from principals that the promotion of over-age children *en-masse* from ECE to primary school would overwhelm current class sizes in primary school, schools were permitted to create a separate accelerated learning class to support students to gain the necessary skills to transition into primary grades within one year. All but two schools elected to use a temporary ALP class, rather than promote over-age children immediately.

In addition, initial consultations with the MoE and schools identified the importance of extending the training to Vice Principals for Instruction (VPIs). In Liberian schools the VPI plays a key role in overseeing teaching and learning. In an intervention such as this, where significant changes in teaching practices and class groupings were being introduced, it was vital that VPIs agreed with the proposed changes. Moreover, in the event of teacher absences or other issues, the VPI plays a key role in supporting continuity of instruction and programming. In schools dealing with understaffing, VPIs may play a key role in ensuring that instruction is provided where there are too few teachers.

Finally, revisions to teaching and learning materials were made following the pilot training. These were made to improve materials and to ensure that common rote learning methods were not included in supplementary materials being provided to teachers.

3.4. Intervention beneficiaries

Based on the evaluation design (Section 5), in which 27 schools were assigned to receive the intervention, it was anticipated that a maximum of 54 ECE teachers and 54 teachers supporting primary-aged children would be direct beneficiaries of the training programme (‘ALP teachers’) – i.e. a maximum of two ECE teachers and two ALP teachers.

However, several sampled schools had fewer ECE teachers. This meant that fewer teachers were initially enrolled in the training programme than had been budgeted for. Consequently, schools that had more than two ECE teachers were invited to include these additional ECE teachers in the training (although these teachers would not be included in the evaluation). In addition, as described in Section 3.3, VPIs (or the principal or another senior member of staff) were invited to participate in the training. In total, 99 teachers were trained as part of the intervention: 41 ECE teachers, 31 ALP teachers, and 27 VPIs. Teachers who were reported to be ‘volunteer’ (unpaid or paid through parent–teacher association donations) were included as teachers.

The indirect beneficiaries were the children enrolled in the classrooms of teachers trained in the intervention. Based on enrolment data collected at baseline, we estimate that the programme reached 1,179 children (744 in Bomi, 345 in Gbarpolu).

3.5. Description of training and support

The intervention consisted of three types of support to teachers: an initial training, school-based mentorship visits, and a refresher training workshop.

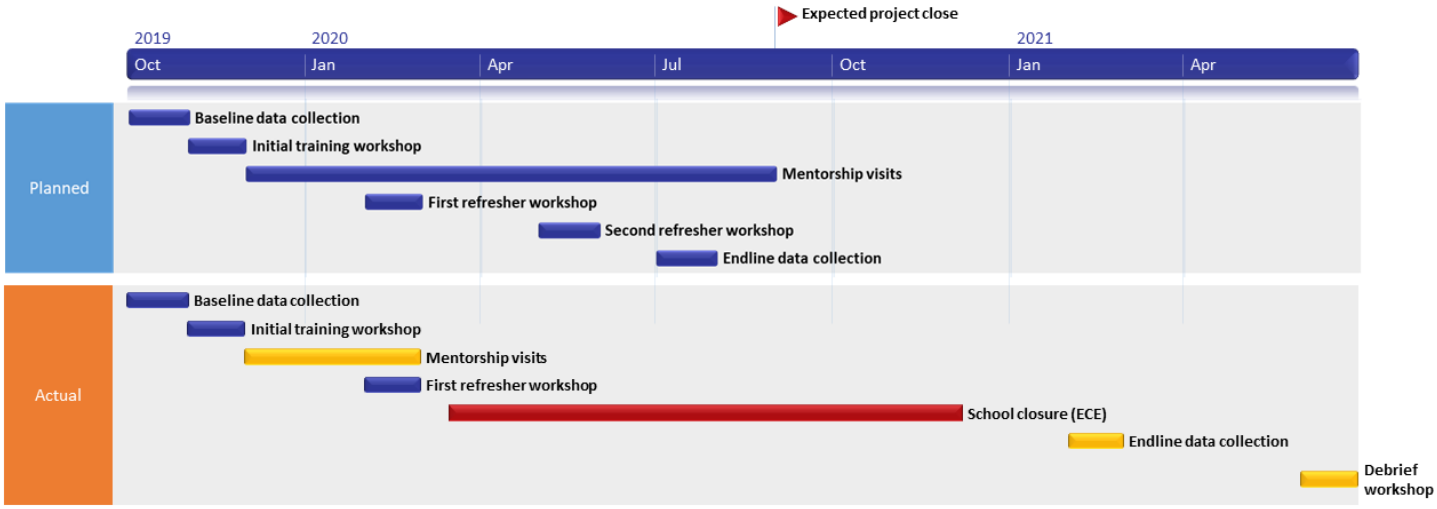
Table 2: Description of training and support

Type of support	Description
Initial training	Five days of initial in-person training was provided to ECE teachers (in the ECE programme) and the selected Grade 1 teachers (in the ALP programme). VPIs received a combination of ECE and ALP training, to allow them to support both sets of teachers. Training was delivered by master trainers nominated by the MoE. Master trainers had been trained to deliver each type of teaching programme and were recognised as specialists in delivering each type of training. Each teacher received either a copy of the ECE curriculum and teacher guide, or the ALP teacher guide and lesson book.
Mentorship	A comprehensive mentorship programme was intended to follow the initial training. Prior to COVID-19, the intervention was designed to support 8–12 mentorship visits per teacher, based on evidence from effective teacher support programmes that mentors should visit teachers at least once a month (World Bank, 2021, p. 2). In each visit, it was expected that teachers would have their classes observed by a mentor, and would receive feedback and guidance on their teaching after the class.
Refresher training	It was intended that teachers would participate in two refresher workshops (each lasting two days) over the course of the intervention, which would revise key concepts and enable teachers to discuss good practices and challenges with peers.

3.6. Dosage and timing

Figure 1 compares the planned timeline for the intervention and the actual timeline. Blue lines indicate planned activities, while orange indicates the actual activities that deviated from the original plan.

Figure 1: Expected versus actual timeline for intervention and fieldwork



The original intervention was planned to run from November 2019 to August 2020. Within this period, mentorship visits were due to run from December 2019 to July 2020. However, due to the COVID-19 pandemic, schools closed in March 2020. ECE classes did not resume formally until December 2020. This significantly affected the dosage of the intervention. Specifically:

- While the intervention was designed to provide 8–12 mentorship visits per teacher (from December 2019 to July 2020), due to school closures mentorship visits were suspended in March 2020 and **thus only 2.8 mentorship visits were provided per ECE teacher and 2.2 visits per ALP teacher.**[5]
- Although a second refresher workshop was intended to take place during the intervention (May 2020), due to school closures this was not possible and the **second workshop was cancelled**. In lieu of this, the We-Care Foundation delivered a debrief workshop for teachers in June 2021.
- Although the intervention originally anticipated that the programme would be delivered in classrooms for nine months (November 2019 to July 2020), due to school closures **the programme was only delivered for 4.5 months** (November 2019 to March 2020).
- The evaluation was originally planned such that endline data collection would take place immediately after the final mentorship visits (August 2020). However, due to COVID-19, **endline data collection took place 11 months after the final mentorship visits.**

In summary, COVID-19 severely affected the delivery of the intervention.

[5] Annex B provides details on the number of mentorship visits per teacher.

4 Learning benchmarks

4.1. Context and purpose

As will be described in Section 5, the primary purpose of the evaluation of children's learning outcomes is to assess the impact of the intervention. However, this analysis would offer additional benefit if children's learning outcomes could be compared to expected levels of learning by student age in Liberia. During the course of the evaluation, the MoE drafted ELDS in consultation with key stakeholders. While these are still pending approval, the ELDS provide a valuable framework for interpreting children's learning outcomes.

Drawing on the ELDS, the research team worked with the MoE and key stakeholders (listed below) in order to develop benchmarks of expected achievement in the research assessments for children of different ages. This process, and the resultant benchmarks, are described below.

4.2. Methodology

Oxford MeasurEd developed a set of online activities to elicit responses from key stakeholders in the Liberia education sector to develop descriptions of what learners know and can do at different ages, and to draw comparisons with desired curriculum competency expectations.

A secure microsite was developed, structuring an online benchmarking activity into four sessions:

- **Session 1** – Introduction;
- **Session 2** – Creating skills descriptors;
- **Session 3** – Setting age-specific Expectations
- **Session 4** - Presentation of Final Benchmarks

Each section included an introductory video explaining the aims of the session, a PowerPoint pack, and a survey function to elicit responses from government officials. The fourth session presented the final decisions based on the engagement with and approvals from the Ministry. This ensured that those who participated were informed of the final decisions made regarding benchmarking.

Eight key stakeholders in Liberia participated in the online activity, including representatives from the ECE Bureau, the Curriculum Department, the Liberia Early Childhood Professional Network, UNICEF and Accelerated Quality Education.

Participants were informed that the expectations should be informed by:

- Liberia ELDS;
- Expertise and experience of participants; and

- Global frameworks and other approaches, including the global learning framework developed to monitor SDG4.

Participants were asked to indicate when the child should be able to do each of the skills assessed in the test. For each descriptor each participant was asked if the child should be able to do the skills before their 3rd, 4th, 5th or 6th birthday. A detailed description of the benchmarking exercise is provided in Annex C.

The results were used to develop proficiency bands and these bands were reviewed and approved by the ECE Bureau.

4.3. ECE benchmarks

Based on the benchmarking exercise undertaken with key stakeholders in Liberia, the following proficiency ranges were agreed to describe those proficiencies tested in the assessment, and to benchmark these specific skills by what children are expected to know and to be able to do in language and literacy (Table 3) and mathematics and executive function (Table 4). These tables describe the specific tasks tested using the assessment tools and are listed from least to most difficult within each proficiency range. In Liberia, the medium of instruction in ECE includes both children's home languages and English, and the medium of instruction at primary school is English. However, the Education Reform Act of 2011 also allows for a local language to be used as the language of instruction at the basic education level (Grades 1–9). County School Boards are responsible for determining the local language used, based on the languages spoken in that region (USAID, 2021).

Table 3: Language and literacy proficiency level descriptions

Proficiency range	Description of the knowledge and skills of learners achieving within this range
Before a child's third birthday	Learners performing within this proficiency range can correctly name body parts, such as hand, ear, eye, and mouth, and can correctly name two or more things that can be eaten.
Before a child's fourth birthday	In addition to the skills above, learners performing within this proficiency range are able to identify the letters A, B, and D, recall specific pieces of information from a simple reading text read to the learner (listening comprehension), name one or more animals and things that can be eaten, and write the letter A.
Before a child's fifth birthday	In addition to the skills above, learners performing within this proficiency range are able to draw a figure recognisable as a circle, a rectangle, or an 'x', demonstrate early print material concepts by holding a book in the correct position, identify the letters C, K, N, and L, write the letters D, F, N, and H, write his or her name in full or in part, name six things or more that can be eaten, infer meaning in a simple listening comprehension

Proficiency range	Description of the knowledge and skills of learners achieving within this range
	exercise, and read high-frequency words like 'book' and 'apple'.
Before a child's sixth birthday	In addition to the skills above, learners performing within this proficiency range are able to: read high-frequency words like 'to' and 'red',[6] write words like 'cup' and 'tree', name eight or more things that can be eaten and animals, correctly sound letters including S and B, say one or more words starting with the sound H, S, and F, and read part or all of a four-word sentence.

Table 4: Mathematics and executive function Proficiency Level Descriptions

Proficiency range	Description of the knowledge and skills of learners achieving within this range
Before a child's third birthday	Learners performing within this proficiency range can count to two and six using stones.
Before a child's fourth birthday	In addition to the skills above, learners performing within this proficiency range are able to identify the longest or tallest of two objects, count from one to 11, identify the object with the greatest volume, and demonstrate limited inhibitory control.[7]
Before a child's fifth birthday	In addition to the skills above, learners performing within this proficiency range are able to (in order of difficulty) identify the highest of two objects;[8] memorise and then recall the correct order of two images previously presented on cards; identify the larger number of two one-digit numbers; identify the numbers two, five, seven, eight, and 10; identify a circle and triangle from a selection of shapes; use two puzzle pieces to replicate an image from a selection of four images; use physical objects to add two one-digit numbers summing to three or five; identify the smaller number of two one-digit numbers; count to 14 using stones; count from one to 30; correctly place at least two pieces of a three piece puzzle together; identify

[6] While these high-frequency words may not seem more difficult than 'book' and 'apple', psychometric analysis finds that they are slightly more difficult. The reasons for this are not known.
[7] Inhibitory control, also known as response inhibition, is a cognitive process and, more specifically, an executive function that permits an individual to inhibit their impulses and natural, habitual, or dominant behavioural responses to stimuli (a.k.a. prepotent responses) in order to select a more appropriate behaviour that is consistent with completing their goals. See Ilieva *et al*. (2015).
[8] For example, a ball in a tree is higher than another ball on the ground.

Proficiency range	Description of the knowledge and skills of learners achieving within this range
	the larger number of two two-digit numbers under 20; identify the smaller number of two two-digit numbers under 40; use physical objects to subtract two one-digit numbers; and count from 50 to 59 and demonstrate some inhibitory control.
Before a child's sixth birthday	In addition to the skills above, learners performing within this proficiency range are able to subtract a single-digit number from a two-digit number without using physical objects; identify the number 13; add two one-digit numbers summing to 10 using physical objects; solve a word problem requiring the addition of two one-digit numbers under five; identify the numbers 12, 14, 17, and 20; add two single-digit numbers without using physical objects; memorise and then identify in the correct order five images previously presented on cards; add a one-digit and two-digit number, using physical objects and subtract a one-digit number from a two-digit number, using physical objects.

It is important to note that the numeracy index contains items from both the numeracy and executive function constructs to match the assessments used in the evaluation. This is because relatively few items on executive function functioned well enough (in ELP Phase 1 and in piloted assessments in ELP Phase 2) to be included in the final assessments of ELP Phase 2. Since there were few items, and since these items were set within a mathematics context, they were included in the section on numeracy during the assessments.

These benchmarks are used in the analysis phase to describe the proportion of children in the early years meeting age-level expectations (as reported in Section 6.1.2).

5 Evaluation methodology

5.1. Overview

The design of this evaluation applies procedures of (i) random sampling to ensure generalisation of the findings to the larger population of schools in the selected counties; and (ii) random assignment of sampled schools to treatment and control conditions to provide causal estimation of the treatment impact. According to the design, schools assigned to the control condition were expected to operate without any changes to their ECE curriculum, pedagogy, or their approach to managing over-age students. In contrast, over the course of the 2019/20 academic year, schools assigned to the treatment condition were expected to implement the ECE curriculum and over-age-student interventions described in Section 3, above. The original intervention was planned to last nine months. However, due to COVID-19, 16 months elapsed between baseline and endline, and schools were closed for 11 of these months.

Students in both treatment and control classrooms were assessed prior to the initiation of the intervention (i.e. at programme baseline) in the areas of literacy, numeracy, and executive function. Following completion of the intervention (i.e. at programme endline), students were once again assessed in these areas, with differences in baseline-to-endline growth between treatment and control students representing the impact of the intervention on student learning. Additionally, teacher practices were measured at both baseline and endline to determine the impact of the programme on teacher pedagogy, teacher–child interactions, and classroom management. A survey of school principals collected information on receptivity to the intervention and implementation costs.

787 children (409 female, 378 male), 82 teachers (41 female, 41 male), and 54 principals (8 female, 46 male) were assessed and surveyed at baseline (November 2019) and endline (February 2021). This reflects a retention rate of 79% for children and 88% for teachers.

5.2. Site selection and sampling

5.2.1 Selection of intervention locations

The sampling frame for this study comprises existing government schools in two of Liberia's North Western counties: Gbarpolu and Bomi. These counties were selected in consultation with the MoE and BRAC. This consultation involved discussions on the feasibility and representativeness of counties in Liberia's North Western, North Central, South Central, and South Eastern regions. The MoE and BRAC identified counties in the North Western (Gbarpolu, Bomi, and Grand Cape Mount) and North Central (Nimba, Lofa) regions as 'preferred' sites for the intervention, with one county in the South Eastern region (River Cess) being identified as an 'acceptable' location. For BRAC, these preferences were based on their existing capacity in each region; for the MoE, these preferences were based on ensuring a relatively equal distribution of foreign assistance in the country.

Ultimately, the North Western region was selected as the study site, motivated by considerations of accessibility, affordability, and feasibility. Additionally, the identification of sites included the selection of potential counties, based on locations that are not currently receiving aid or where other external projects by the Global Partnership for Education or USAID are not being implemented. Of the three 'preferred' counties in the North Western region, Gbarpolu and Bomi were selected as the intervention counties, based on higher levels of poverty (average levels of absolute poverty are 60.5% in Gbarpolu, 64.3% in Bomi, and 53.7% in Grand Cape Mount, compared to a national average of 50.9%) (Liberia Institute of Statistics and Geo-Information Services, 2017). The motivation for this was to select sites that would better represent a realistic understanding in regard to scaling up the intervention to additional counties in the future. Additionally, Gbarpolu and Bomi counties are adjacent to one another, which helped to ensure the efficient use of programme and evaluation resources.

5.2.2 Sampling frame

Within the selected counties of Gbarpolu and Bomi, the evaluation team used the most recent administrative data (Education Management Information System (EMIS) 2016/17) to identify a sampling frame of schools. The criteria for inclusion were schools with the following characteristics:

- enrolment of at least 20 ECE students between the ages of three and six;
- enrolment of at least 10 over-age students (seven years old and above); and
- schools with at least two ECE classrooms and teachers.

Additionally, prior to random sampling and assignment, we removed schools that appeared to be extreme outliers according to size (number of students or teachers) or share of over-age enrolment, to ensure comparability between treatment and control schools. Using the EMIS data, we identified a population of 105 schools (60 in Bomi, 45 in Gbarpolu) that satisfied the above criteria. These 105 schools comprised the sampling frame for the study. According to EMIS data, 80% of these schools are located in rural areas, while all of the schools in Gbarpolu are rural.

5.2.3 Random selection and assignment of schools, teachers, and students

The aim of the random sampling and assignment procedures was to include schools that are representative of the population and that are as similar as possible across treatment and control conditions. This is particularly important, given that the budget for the intervention and evaluation constrained the evaluation to a sample of 54 schools (27 treatment and 27 control) – a reality that required maximisation of statistical power as far as possible. To this end, we employed a stratified random sampling procedure – stratifying by school size (i.e. number of ECE-age students) and urban/rural status (in Bomi only, as there were no urban schools selected in Gbarpolu) – to ensure a balance between treatment and control groups.

This procedure resulted in a sample of 27 treatment schools and 27 control schools. The sample comprises 30 schools from Bomi (15 treatment, 15 control) and 24 schools from

Gbarpolu (12 treatment, 12 control). This proportion (55% Bomi) is equivalent to the proportion in our population of eligible schools (54% Bomi). Initial balance tests suggest that these treatment and control groups have no statistical differences in regard to the number of students, urban/rural status, student age, share of over-age students, or share of female students.

Within both treatment and control schools, two ECE classrooms and teachers were randomly sampled for inclusion in the evaluation. The sampling frame for this random selection consisted of all nursery and kindergarten classrooms that include students aged three to six. Following the selection of two ECE teachers/classrooms, a random sampling of seven or eight students was taken within each classroom (for a total of 15 per school). Furthermore, during baseline data collection, the research team randomly selected 15 over-age ECE students within each school to track to endline for the purposes of measuring the effect of the ALP intervention on promotion of over-age children. This assessment investigated whether a child had been promoted to primary school. As noted previously, due to funding constraints it was beyond the scope of this study to investigate changes in learning outcomes for over-age children.

Further information on the stratified random sampling procedure, the balance tests between treatment and control schools, the protocol for selecting teachers and classrooms, and the protocol for selecting students is provided in Annex B.

5.2.4 Power calculations

Annex B provides the details of, and rationale for, the study's power calculations. Prior studies of ECE interventions in low- and middle-income countries have found treatment effects of student-centred and play-based instructional interventions above the .30 standard deviation (SD) threshold (Stagnitti *et al*., 2016; Nicolopoulou *et al.*, 2015; Roskos *et al.*, 2010; Biçer , 2017), some of which were similarly conducted over the course of one-year interventions. Using these studies and the student performance data from ELP Phase 1, we estimated minimum detectable effect sizes (MDES) *a priori* to guide school and student selection during the sampling stage. *Post hoc* power calculations using the endline student learning data suggest that the MDES of student-level treatment effects is 0.30 SD for literacy and 0.33 SD for numeracy. This suggests that the evaluation has sufficient statistical power to detect treatment/control differences of 0.33 SD or larger.

5.3. Data collection

5.3.1 Demographic data

At baseline, children's information, including their home address, was collected from the children and the schools. This information was used for following up with children who were not found in the school at endline.

Additionally, data on students' age, sex, and household wealth were collected to account for any potential differences in basic demographic characteristics. Household wealth was measured by asking children whether their households possess a series of assets. These are described in Annex B.

5.3.2 Balance between treatment and control schools

Annex B reports descriptive statistics and shows that the sample is balanced across a range of student- and teacher-level characteristics collected at baseline. The sample is balanced across all student characteristics, but there is some imbalance along teacher gender: 39% of the control group teachers are female, relative to 59% in the control. This difference is statistically significant at the 10% level. By chance one would expect imbalance on some of the variables. However, the fact that we observe balance on most of the other characteristics leads us to conclude that the randomisation worked. We control for teacher gender in all our teacher-level analysis.

5.3.3 Learning assessments

The student assessments collected information on three key constructs:

- **Literacy** – defined by the MoE as 'the knowledge and skills that lay the foundation for reading and writing'. The assessment uses items covering a range of literacy skills across pre-literacy, emerging literacy, and basic literacy levels: letter recognition, letter sounds, expressive vocabulary, listening comprehension, and copying, reading, and writing familiar words.
- **Numeracy** – defined as 'the ability to think using mathematical concepts (Nunes and Bryany, 1996) which provide powerful tools for describing and understanding that world around us' (UNESCO, UNICEF, Brookings Institute, and the World Bank, 2017). The mathematics test uses a range of items to test various mathematics skills across the pre-, emerging, and basic mathematics levels: number recognition, verbal counting, number comparison, addition and subtraction, simple word problems, set production, shape knowledge, measurement vocabulary, and spatial awareness.
- **Executive function** – refers to cognitive skills (such as working memory) that support children's ability to learn, and that coordinate goal-directed behaviour and activities. This module assessed working memory and inhibitory control.

These assessments were administered in English, as this is the language of instruction (alongside children's home languages). Rigorous quantitative analytical methods (item response theory (IRT)),[9] were used to construct two measures of learning: literacy and numeracy. The numeracy index contains items from both the numeracy and executive function constructs. This is because relatively few items on executive function functioned well enough to be included in the final assessments. Since there were few items, and since these items were set within a mathematics context, they were included in the section on numeracy during the assessments.

IRT utilises in-depth analysis of multiple measures of student knowledge to construct scale scores that are representative of a student's capabilities in a subject matter area. This process also has the benefit of scaling student performance such that differences between

[9] Further details are available in Annex C.

high-performing and low-performing students are accurately reflected. Finally, IRT enables analysis that compares students' performance with the benchmarks described in Section 4. IRT allows for the ranking of students according to their ability and places the students and the items onto the same metric. This is a probability model as students are placed on the scale according to the probability of a student answering the corresponding item correctly. The proficiency levels described in the previous section correspond to this metric. This enables us to link learner scores to the benchmarks.

Beyond these three student assessments, Phase 1 also included the administration of a module assessing students' socio-emotional skills. However, Phase 1 analysis revealed low levels of reliability of the data. The team explored options for an alternative socio-emotional instrument for use in Phase 2, but was unable to find any measures of which there was confidence of useable data from the Liberian context. Given this, and given the constraints of the data collection budget, its tangential relationship to the key research questions of interest, and the already high demands of the student engagement, the team decided to forego the administration of any socio-emotional measure for Phase 2.

Due to the impact of COVID-19, we anticipated high levels of student and teacher attrition at endline. Consequently, endline data collection was adapted to include home visits (for children) and phone interviews (for teachers) for participants who were not present at school on the day of assessment.

5.3.4 Classroom observations

We would expect successful implementation and adoption of the intervention to be exhibited through increased usage of the teaching practices found within the ECE curriculum. As such, the impact of the programme is measured not only in the child learning outcomes, but also in the observed practices of teachers within their classroom environments. To this end, we used an adapted version of the MELQO: Measure of Early Learning Environments (MELE) classroom tool to measure teachers' activities and use of the national curriculum within the classroom environment.

The classroom observation protocol involved an enumerator observing and recording classroom activities taking place within the classroom at five-minute intervals over the course of a one-hour period. Within these intervals, observers recorded the subjects or skills being taught by the teacher, the teaching practices being used (by teachers and/or teaching assistants), the materials utilised by the teacher, the forms of engagement by the students, the size of student groups involved in learning activities (whole group, small groups, pairs, or individual students), and the extent of student engagement in the activities. This observation process produced the following variables capturing the experience of teachers and students within the classroom:

- **Student-centred teaching practices.** The number of student-centred teaching practices (e.g. asking/answering open questions, engaging children in discussion, demonstrating with learning objects/materials, singing or telling rhymes, etc.) used across the entire observation period.
- **Skills**. The number of instances across the entire observation period teachers taught gross motor skills (running, stretching, dancing, ball games, chasing/tag, etc.) and

expressive language skills (children's descriptions of objects or pictures, children's telling of stories).

- **Small pair.** The number of instances observed across the entire observation period of students working in pairs or small groups.
- **Engagement**. The amount of student engagement averaged across the entire observation period (e.g. few children engaged, some engaged, most engaged, or all engaged).
- **Student-centred child activities**. The number of student-centred activities that students were engaged in across the entire observation period (free play or open choice, music/movement activities, etc.).
- **Student-centred total**. A composite total of all student-centred teaching practices and learning activities.
- **Teacher materials**. The number of different teaching materials used by the teacher during the observation period.
- **Student materials**. The number of different learning materials used by the students during the observation period.

5.3.5 Surveys of teachers and principals

The evaluation included surveys of teachers and principals to understand staff perspectives of the programme and associated challenges. These surveys included information on:

- basic demographics of the respondent;
- their participation in the training programme and/or different training programmes;
- their current teaching practices and general receptivity to the intervention;
- their attitudes towards child-centred learning;
- their activities during COVID-19 school closures; and
- estimated costs of the programme.

Teachers' attitudes towards child-centred learning was measured using items from Schaefer and Edgarton's (1985) scale on attitudes towards child-rearing, which provides two sub-scales of 'traditionalism' (e.g. 'children should always obey their teacher') and 'progressivism' (e.g. 'children have a right to their own point of view and should be allowed to express it'). Details of each of the above instruments are included in Annex B.

5.3.6 Cost data

In order to calculate the costs of providing the intervention, the research team distinguished between costs incurred in the delivery of the training and mentorship, and costs incurred in the delivery of the intervention in the classroom.

The original evaluation design anticipated that the costs of the training and mentorship would be calculated based on the costs of providing the intervention during the evaluation. However, this posed several challenges in practice as inflation, fuel shortages, and cash shortages in Liberia (as well as the reduction in mentorship visits due to COVID-19) between 2019 and 2021 distorted the costs of the intervention unpredictably during the evaluation. Consequently, the cost of the intervention during the evaluation would not provide a reliable guide for the cost of the intervention in ordinary circumstances.

As a result, the research team worked with We-Care Foundation in September 2021 to estimate the cost of delivering the training and mentorship programme under ordinary conditions. This was done using an ingredients approach, and following the guidelines provided by JPAL (2021). This approach aims to account for the costs incurred by all parties involved in the intervention – such as schools and teachers – and not only the training provider.

The following categories of costs were included:

- programme administration, including the time of the staff hired to plan, deliver, and report on the training and mentorship;
- user training, including the cost of materials provided through the training and mentorship, venue rental, and equipment used in the training;
- user costs, including the costs of transport to and from the training, meals, and lodging for trainees; and
- monitoring costs, such as oversight by senior staff.

The following costs were excluded from the calculation:

- programme administration and overheads, such as head office rental and project accounting;
- the costs of developing the materials;
- the costs of recruiting and training mentors;
- the opportunity costs of participants' time;
- the costs of replacement teachers to cover for trainees during the workshops; and
- averted costs.

Further details of the costing methodology are included in Annex B.

5.3.7 Timeline for data collection

Baseline data were collected in October 2019. While endline data collection was originally anticipated to take place in July 2020, school closures from March 2020 rendered this impossible. Consequently, endline data collection was conducted in February 2021. This was designed to be as soon as possible after schools reopened, as children changed to new classes from December 2020 (i.e. in the new academic year). We anticipated that schools' policies for how to promote children would vary unpredictably between schools. As a result, it was expected that children assessed at baseline would be located in different classrooms and with different teachers when schools reopened. The research team was concerned that this would add additional noise to the data collected at endline, which could increase the amount of time that children were taught by different teachers under different promotion policies.

Consequently, endline data collection was conducted before the final refresher training was concluded. This is for the same reasons as were described in the paragraph above: since some children would be taught by different teachers to those teachers that had been trained in the intervention, and since this would vary unpredictably between schools, the final refresher training would add additional noise to the intervention.

5.3.8 Balance and attrition

We were able to find and assess 79% of students (620 students) assessed at baseline and this rate is balanced across evaluation arms (treatment and control). Only 65% of students (506) were present at school during our endline school visits, but we further tracked and assessed another 114 students at home. We were able to complete surveys with 87% of teachers at endline. We conducted classroom observations for 58% of our original sample of teachers (61% in the treatment group and 54% in the control) as some teachers were no longer teaching ECE in the new academic year. This is reflected in Table 5.

Table 5: Sample size

	Planned units	Baseline (actual)	Endline (actual)
Number of children assessed	810	784	620
Number of teachers surveyed	108	82*	72
Number of classroom observations	108	82	47

** The figure of 108 teachers was based on the assumption of two ECE teachers (and classrooms) per school. However, some schools had fewer classrooms than this.*

Annex B details our tests for whether attrition was random with reference to the characteristics of interest (i.e. gender, age, learning), and whether the samples remained

balanced, after accounting for attrition. In brief, we find no evidence for differential attrition on observable characteristics (i.e. there were no differences between the attriters and non-attriters in terms of gender, age, and student learning at baseline) and the attrition rates were balanced, at 21% in both the treatment and control groups. This is good for the validity of the study, since it means that the treatment did not change the characteristics of the students who ended up attriting. For teachers, there is some evidence of non-random attrition: the teachers who dropped out of the sample at endline were more likely to be male, and exhibited worse teaching practices at baseline: they implemented fewer teacher-centred practices, and students in their classes were less likely to work in pairs or to be engaged in the classroom. However, the coefficients on 'treatment' show that the sample remains balanced along these characteristics. Most of the coefficients of teaching practices are negative, although they are not statistically significant.

In sum, these findings strengthen the internal validity of the study.

6 Findings

6.1. Child assessments

6.1.1 Comparison between treatment and control

RQ1. What impact do professional training for ECE teachers, child-centred and play-based learning, and classes being restricted to age-appropriate children have on the cognitive outcomes of children who are of the appropriate age to attend ECE (i.e. between the ages of three and six, inclusive)? Specifically, what is the impact of the intervention on children's (i) early literacy, (ii) early numeracy, and (iii) executive function in particular?

Table 6 shows the evaluation results for student learning. The coefficient estimates of treatment show the magnitude of the estimated effect of the intervention, in terms of standard deviations. It is clear that there is no positive impact of the intervention on student learning. The coefficient of treatment is negative for literacy and numeracy outcomes, although it is very small and not statistically significant. Note that our results for these two outcomes are relatively precisely estimated,[10] and thus it is unlikely that the null result is due to insufficient statistical power.

Even though the indicators for executive function were included in the numeracy index, for completeness we also report results for these indicators separately in column (3). This index was constructed using principal component analysis, based on the seven indicators for executive function, standardised to have a mean of 0 and a standard deviation of 1 in the control. The impact on executive function is larger, but again it is not statistically significant.

The findings on the impact of the intervention on learning outcomes do not vary based on students' household wealth.

Table 6: Impact on student learning

	Literacy (SD)	Numeracy (SD)	Executive Function (SD)
Treatment	-0.036	-0.018	0.134
	(0.098)	(0.084)	(0.109)
Observations	620	620	620
R-squared	0.559	0.540	0.130
Control mean	0.44	0.59	0.00

SD= Standard Deviation. Standard errors, clustered at the school level, in parentheses. * p<0.05 ** p<0.01 *** p<0.001.

[10] For example, with a standard error of 0.084, we are able to detect an impact of 0.156 SD at a 5% level.

	Literacy (SD)	Numeracy (SD)	Executive Function (SD)
All specifications include controls for baseline literacy and numeracy (including executive function), and district fixed effects. The indicators for literacy and numeracy are constructed using IRT, validated including the baseline data as well. Data are further standardised to have a baseline mean of zero and a standard deviation of one. 'Control mean' shows the mean in the control group at endline.			

In summary, the intervention had no impact on any domain of student learning. As described in Section 3.6, this may be due to the disruptions caused by the pandemic. As a consequence of school closures, only 23% of planned mentorship visits were completed, and schools were closed for 11 of the 16 months of the intervention.

Key point (RQ1):

- The intervention did not have an impact on student learning.

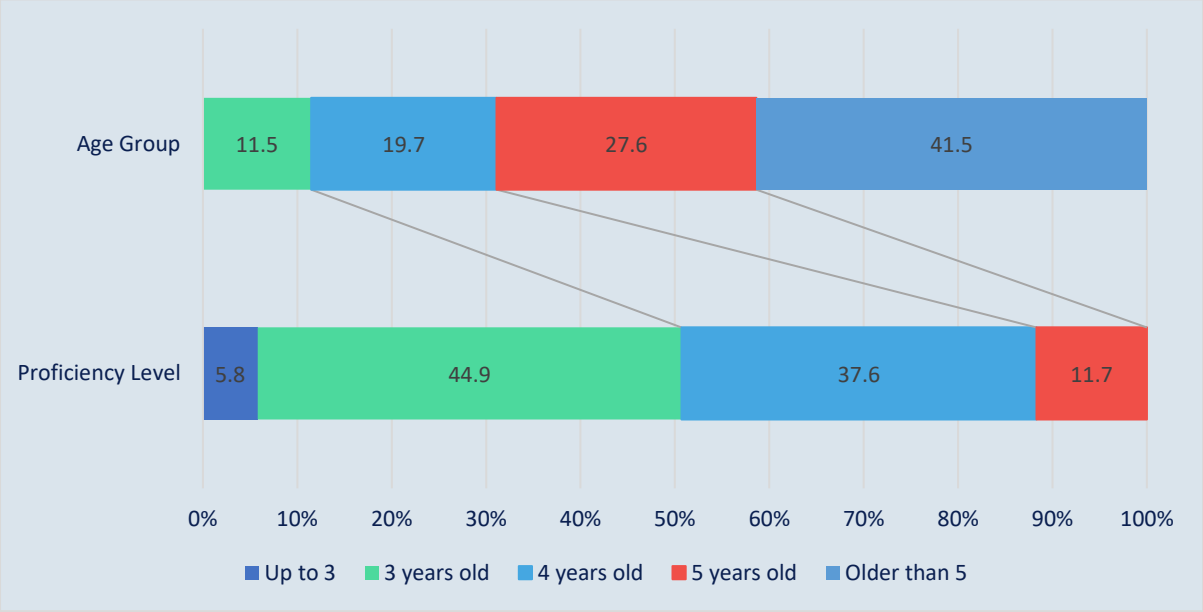
6.1.2 Children's learning outcomes compared to benchmarks

Section 4 described the development of learning benchmarks for each proficiency level. In this section, we analyse children's learning outcomes with reference to these benchmarks. Detailed analysis is available in Annex C.

Language and literacy

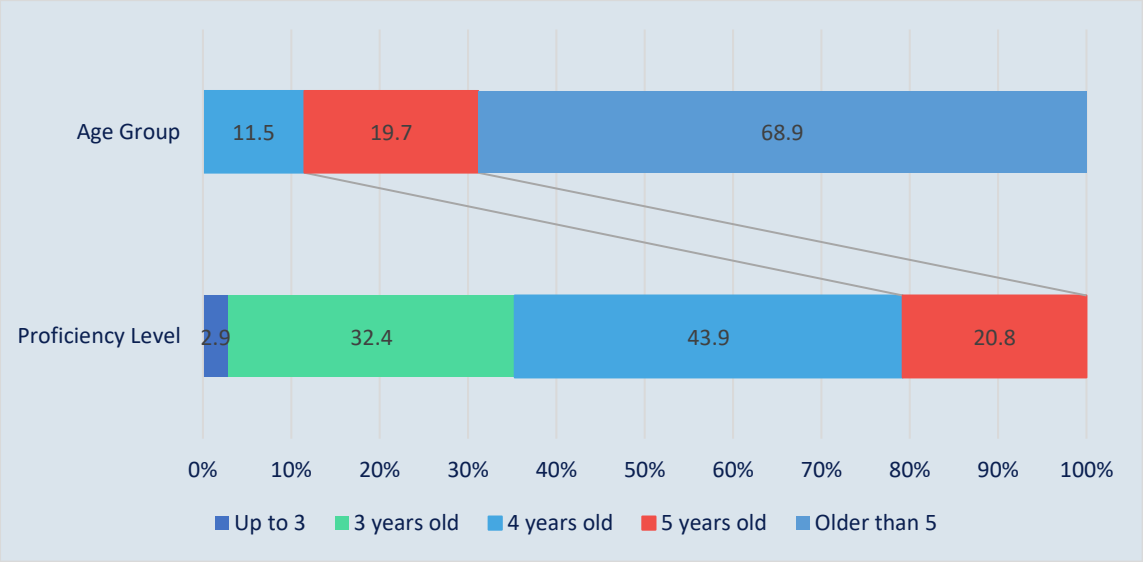
The findings of the language and literacy assessment indicate that the majority of children were not meeting expected language and literacy proficiency levels for their age at baseline or at endline. At baseline, while 68.9% of learners were more than five years old, only 11.7% of learners were achieving at the level expected of children who are five years old. Additionally, while no two-year-olds were assessed in language and literacy, 5.8% of learners were proficient at levels expected for children less than three years old.

Figure 2: Percentage of learners in each age group compared to percentage of learners performing at each language and literacy proficiency level (baseline)



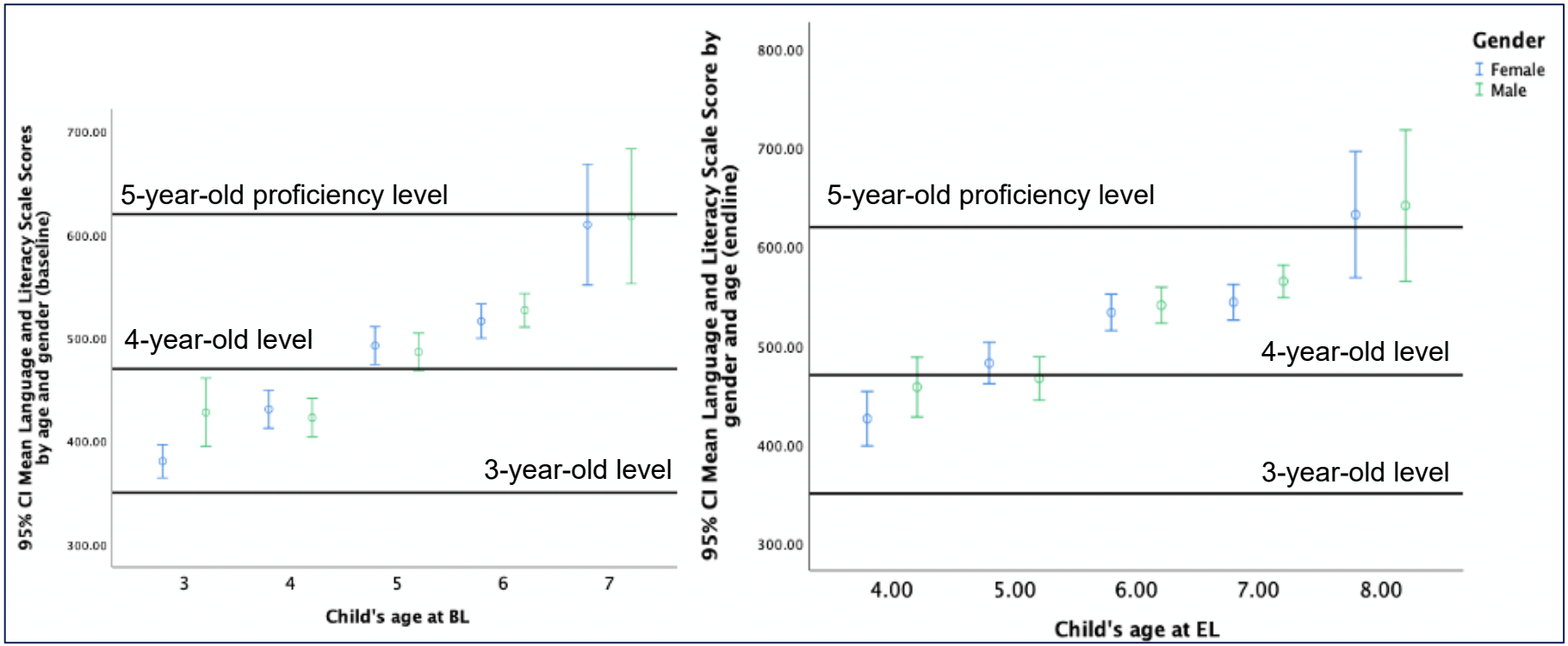
By the endline assessment, the proportion of four- and five-year-old learners achieving what is expected increases. This is mainly explained by the shift in the distribution from learners performing in the three-year-old and four-year-old proficiency bands. While this progress between baseline and endline is positive, the rate of progress is not sufficient to keep up with the rate of expected progress. For example, at baseline, 68.9% of learners were 5 years old or older, and 11.7% of learners were achieving what is expected for a 5-year-old. By endline, 88.6% of learners were 5 or more years old, and 20.8% of learners were achieving what is expected of a 5 year old.

Figure 3: Percentage of learners in each age group compared to percentage of learners performing at each Language and Literacy proficiency level (endline)



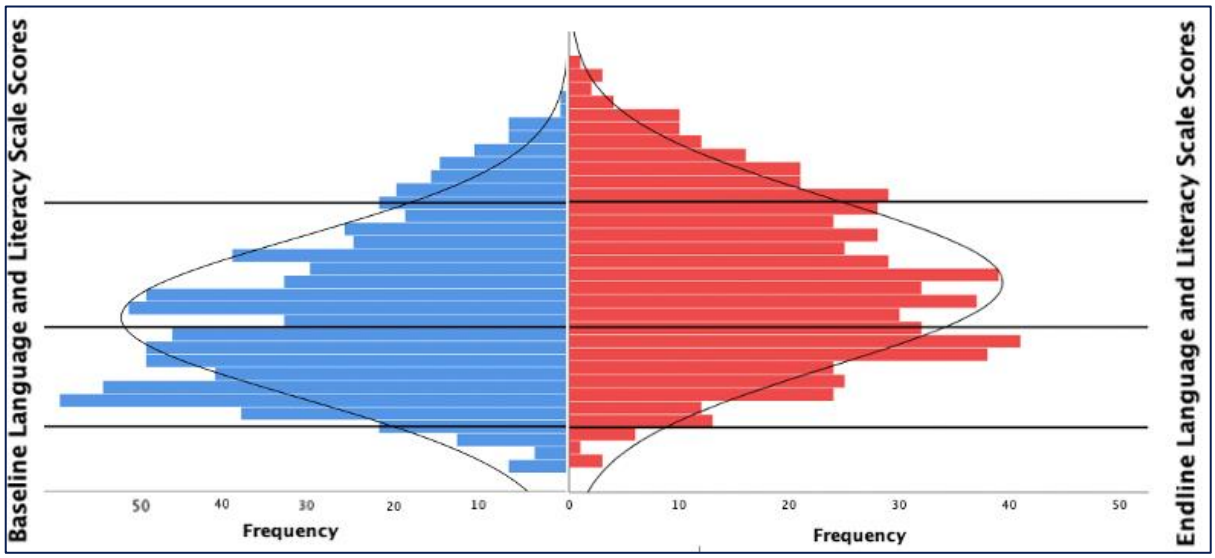
At endline, the average four-year-old is proficient at the level expected for a three-year-old, the average five-, six- and seven-year-old is proficient at the level expected for a four-year-old and the average eight-year-old is proficient at the level expected for a five-year-old. At endline, the average seven-year-old has a lower level of proficiency than the average seven-year-old at baseline. This is because the seven-year-olds at baseline turned eight years old by endline and the six years olds at baseline turned seven years old by endline, however very little learning actually took place during this time relative to expectations..

Figure 4 Mean Language and Literacy Scale Scores by age and gender at baseline and endline



As can be seen in Figure 5, the peak of the distribution shifts from near the cut-off point between three- and four-year-old expectations to the centre of the four-year-old expectation proficiency band. The range of proficiency does not shift a great deal between baseline and endline, indicating that some learners maintain very low proficiency levels as they age.

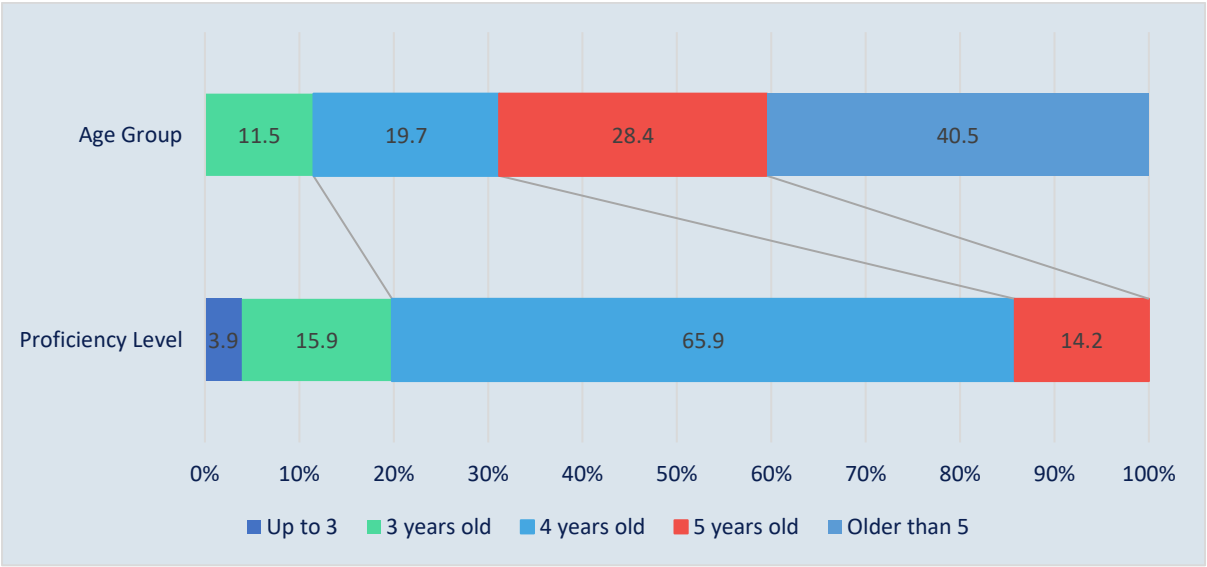
Figure 5: Distribution of baseline and endline language and literacy proficiency



Mathematics

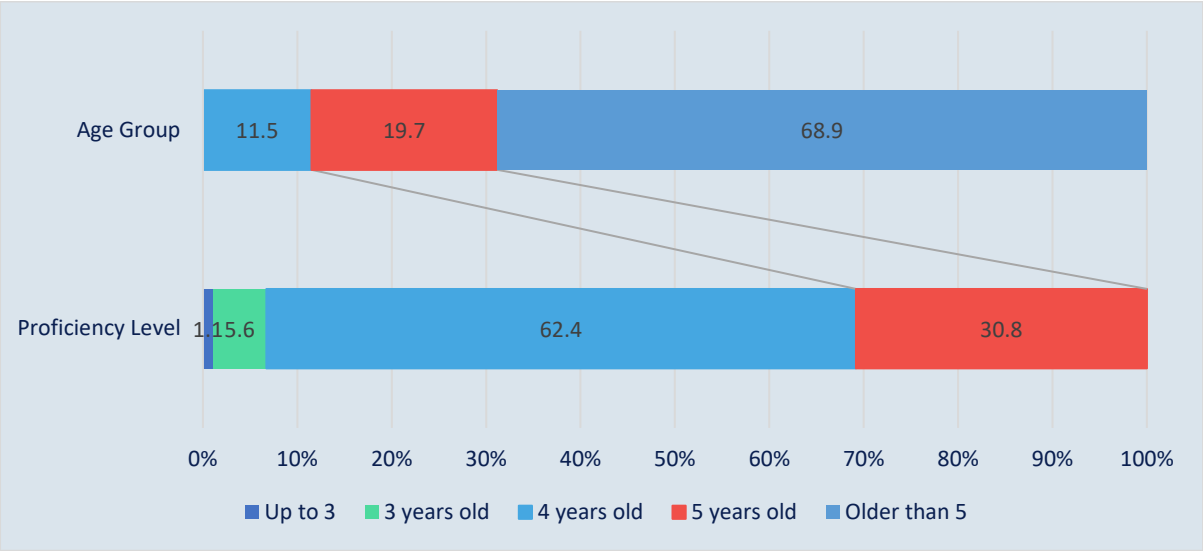
At baseline, approximately 70% of learners were five years old or more. 14% of learners were achieving the proficiency levels expected of a child who is five. Additionally, while there were no two-year-olds in the sample, almost 4% of the learners assessed were proficient in mathematics at the level expected for a two-year-old.

Figure 6: Percentage of learners in each age group compared to percentage of learners performing at each proficiency level for mathematics and executive function (baseline)



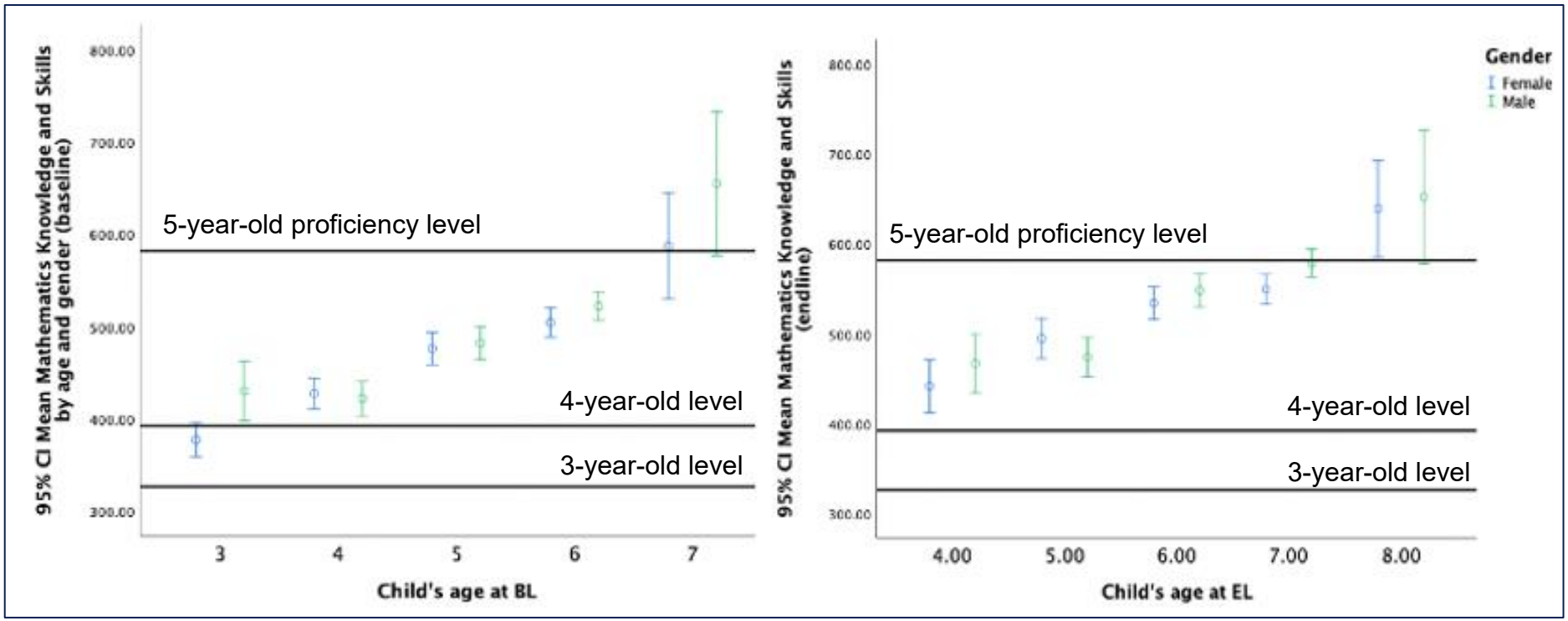
The progress between baseline and endline is positive for mathematics. However, at endline, a large majority of learners are more than five years old and less than nine years old. Approximately a third of learners perform at the level expected for a five-year-old. Additionally, while there are no two- or three-year-olds in the sample at endline, 6.8% of learners are achieving within the range expected of two- and three-year-olds.

Figure 7: Percentage of learners in each age group compared to percentage of learners performing at each proficiency level for mathematics and executive function (endline)



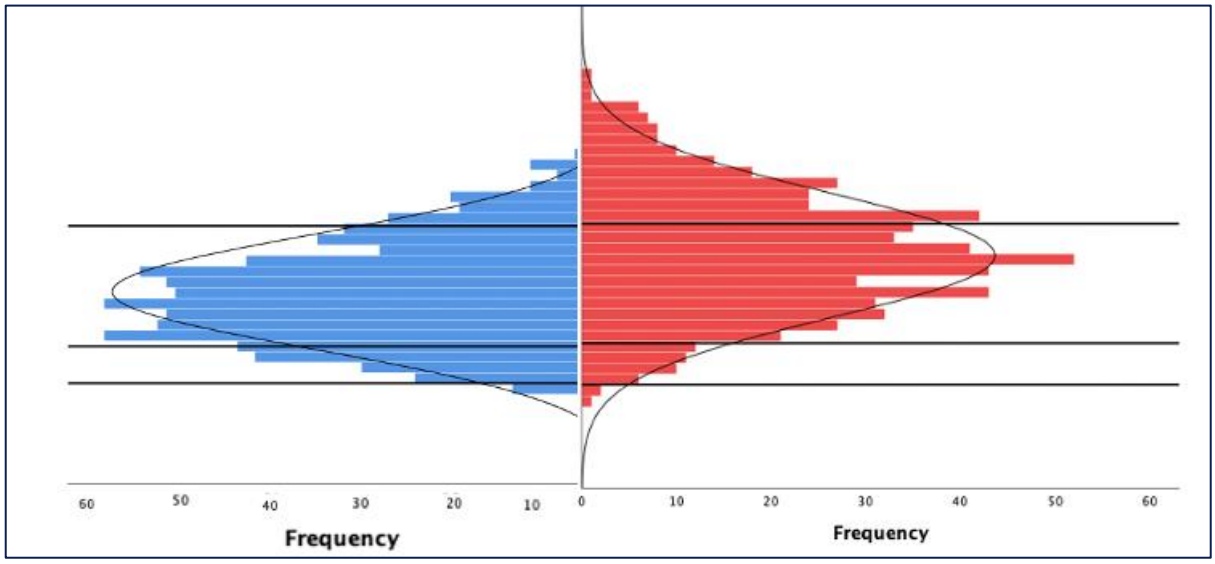
Within the sample, at baseline the average three-year-old girl performed at the level expected of three-year-olds and the average three-year-old boy performed above the expected level. The average four-, five- and six-year-old performed within the range expected for four-year-olds and the average six-year-old performed within the range expected of a five-year-old. At endline, the same pattern is observed, however the children were a year older at endline. Therefore, at endline the average seven-year-old is performing at the level expected of a four-year-old, with the average boy scoring just below the cut-off point for the five-year-old proficiency level). As with literacy, this is because the children aged between baseline and endline. The second implication is that the gap between actual age and expected proficiency levels takes hold around five years of age.

Figure 8: Mean mathematics knowledge and skills scale scores by age and gender at baseline and endline



As can be seen in Figure 9, the range of scores substantially increases between baseline (blue) and endline (red). However, this takes place at the top of the scale and learners with very low proficiency levels remain.

Figure 9: Distribution of baseline and endline mathematics scores



For both language and literacy and mathematics knowledge and skills the proportion of children achieving the highest proficiency level increased between baseline and endline and the proportion of children achieving within the lowest proficiency level decreased between baseline and endline. Similarly, the range of proficiency increased between baseline and endline, with higher scores being achieved by some learners at endline. However, there remain children with proficiency levels expected at the age of two, despite the youngest children in the sample at endline being four years old.

There are no statistically significant differences by gender. At baseline 52.5% of girls met their age-level benchmark for literacy, compared to 52% for boys. Similarly, 63.3% of girls met the age-level numeracy benchmark, compared to 62.8% for boys. Further details are provided in Annex B.

While proficiency increased between baseline and endline, the age of the children in the sample has also increased from three- to seven-year-olds at baseline to four- to eight-year-olds at endline. When observing average scores by age against the age-level expectations in Liberia, the gap between actual proficiency and expected proficiency increased between 2019 and 2021.

Key point (RQ1):

- Children in ECE performed below the benchmarks identified by stakeholders as what would be expected of children of their age for literacy and numeracy.

.

6.1.3 Changes to learning outcomes during COVID-19

Since the schools were closed for the majority of the time between our baseline and endline rounds of data collection – 11 out of the 16 months – we were also interested in understanding how much children learned over that period, and how that compares with how much a child that age typically learns in a year. We examine the changes in overall learning levels between baseline (November 2019) and endline (January 2021) rounds of data collection, roughly one year apart. Our assessment tool, calibrated using IRT, allows us to look at improvements over time.

We find no evidence of substantial learning loss due to school closures. On average, students improved their numeracy and literacy scores by 0.58 and 0.42 standard deviations, respectively, between the baseline and endline rounds of data collection. Given school closures and the limited reach of the TbR programme, this may be explained by children's natural rate of development. To put this in perspective, at baseline each additional year of age was associated with a 0.46 SD increase in student literacy and numeracy scores: i.e. a student who was *x* years old at baseline performed on average 0.46 SD worse compared to a student who was *x*+1 years old. This is a rough measure of how much a student learned in a year prior to COVID-19, given certain assumptions.[11] If we take 0.46 as our measure of pre-pandemic learning over a period of a year, then students' literacy increased by *more* during the pandemic period compared to the pre-pandemic period (0.58 vs 0.46), and their growth in numeracy was only slightly less (0.42 vs 0.46 SD). However, the correlation between baseline learning and age is by no means a perfect counterfactual for how much a preschool child typically learns at school on a given year.

Key points (RQ1):

- Learning levels in ECE did not decrease during school closures, and this may reflect children's natural rate of development.
- The increase in learning outcomes during one year of school closures (i.e. one year without ECE) is approximately equal to the difference at baseline in learning outcomes between children aged one year apart. This raises the possibility that there might be no learning 'value-add' for ECE in the context studied.

[11] For this to be a true measure of how much a child learns in a year, we need to assume that everyone in our baseline sample was in school for a whole year.

6.2. Teaching practice

6.2.1 Uptake of ECE curricula

RQ2. Does the revised ECE programme increase the uptake of the national ECE curriculum?

Data collected from both teachers and principals provide insights into the impact of the programme intervention on the use of the ECE curriculum within the classroom and school.

Table 7 indicates that, at endline, teachers in treatment classrooms report having access to the national ECE curriculum at significantly higher rates (75%) than control teachers (42%) (*p* =.011). [12] Moreover, when looking only at teachers with access to the national curriculum, a larger share of treatment teachers are looking at the materials on a daily basis (64.3%) compared to control teachers (33.3%). However, the difference in regard to the rate of accessing these materials is only statistically significant at p < 0.1 (*p* = .074). Similarly, 52% of control principals at endline report use of the national ECE curriculum in their schools. In treatment schools, 61% of principals report use of the national curriculum and the remainder (39%) use the BRAC teaching materials, which were based on the national curriculum. The reason for the discrepancy with treatment teachers' reported usage of the national curriculum is unclear.

Table 7: How frequently do teachers look at the national curriculum materials (among those with access to the curriculum)?

Control				Treatment		
Frequency	*Count*	*Percentage*		*Frequency*	*Count*	*Percentage*
I never look at them	1	8.3%		**I never look at them**	0	0.0%
Only once	1	8.3%		**Only once**	0	0.0%
Monthly	1	8.3%		**Monthly**	2	7.1%
Weekly	5	41.7%		**Weekly**	8	28.6%
Daily	4	33.3%		**Daily**	18	64.3%
Total	**12**	**100.0%**		**Total**	**28**	**100.0%**

75% of treatment and 67% of control teachers who have access to the ECE curriculum have received training on how to use the materials. However, treatment teachers have more often received training on how to develop and use lesson plans that align with the curriculum (72% vs. 39%; *p* = .01). The primary reason cited by principals for not using the national curriculum is that they do not have access to it (86% of the 21 schools who do not use the curriculum). Otherwise, 5% of schools do not use the national curriculum because they do not feel it is a good fit for their school, and 9% of schools who do not use the national curriculum believe that they use a superior curriculum.

[12] As always for such surveys, it is possible that teachers' reported usage does not reflect their actual usage, but rather reflects the results of a social desirability bias introduced by the intervention.

In summary, the intervention appears to have been successful in increasing the uptake of the national curriculum. This is despite severe disruptions caused by the pandemic, which suggests that relatively little contact is needed to effect this change. However, it is important to note that uptake of the national curriculum alone is not sufficient to improve learning outcomes, as indicated in the previous section.

6.2.2 Teachers' beliefs regarding child-centred learning

As described previously, the national ECE curriculum promotes child-centred pedagogy. We investigate whether participation in the ECE programme affected teachers' beliefs about the suitability of child-centre learning.

Teachers in both treatment and control groups were most likely to support the following statements from Schaefer and Edgarton's (1985) scale on attitudes towards child-rearing (Table 8) (1 = strongly disagree, 5 = strongly agree).

- 'It is more important for a young child to study for the future than enjoy today' (mean = 4.44).
- 'Children should always obey their teacher' (mean = 4.29).
- 'The most important thing to teach children is absolute obedience to whoever is in authority' (mean = 4.21).

While the following statement was also met with higher relative levels of agreement from teachers ('children have the right to their point of view and should be allowed to express it', mean = 4.24), overall, teachers appear to express greater levels of support for more authoritarian perspectives, such as those mentioned above, than those that might be better aligned with a child-centred learning approach (such as 'children learn best by doing things themselves rather than listening to others', mean = 3.13). There was no significant difference between treatment and control groups in this regard.

It is possible that these beliefs are a function of difficult teaching conditions: overall, teachers agreed with the statements 'I don't always have the materials I need to do my job' (mean = 3.84), 'there are too many students in my classroom' (mean = 3.61), and 'it is difficult to teach in this school because the building is in poor condition' (mean = 3.38). While there was a statistically significant difference between treatment and control schools on building conditions, as this is the only statistically significant variable this may be a product of random chance.

We also tested differences in the levels of support expressed for teacher-centred perspectives[13] compared to student-centred perspectives[14]. On average, teachers score

[13] The teacher-centred statements were: children should always obey their teacher; the most important thing to teach children is absolute obedience to whoever is in authority; mean = 4.25

[14] The child-centred statements were: children have a right to their own point of view and should be allowed to express it; I go along with the game when a child is pretending something; children learn best by doing things themselves rather than listening to others; children should be allowed to disagree with adults if they feel their own ideas are better; it's all right for a child to disagree with me; mean = 3.12

1.13 points higher on the teacher-centred perspectives (t = 11.2, p < .01), suggesting less support for student autonomy and exploration.[15]

Table 8: Teacher views on child-centred learning (1 = strongly disagree, 5 = strongly agree)

Category	*Mean*	*Std dev.*	*Control–treatment diff (sig.)*[16]
It is more important for a young child to study for the future than to enjoy today	4.44	0.59	0.17
Children should always obey their teacher	4.29	0.56	-0.04
Children have a right to their own point of view and should be allowed to express it	4.24	0.72	0.11
The most important thing to teach children is absolute obedience to whoever is in authority	4.21	0.78	0.29
I don't always have the materials I need to do my job	3.84	1.06	-0.31
There are too many students in my classroom	3.61	1.32	0.35
It is difficult to teach in this school because the building is in poor condition	3.38	1.33	-0.72*
I go along with the game when a child is pretending something	3.33	1.06	-0.37
Children learn best by doing things themselves rather than listening to others	3.13	1.26	0.44
It is difficult to manage students in my classroom	2.87	1.32	-0.44
Children should be allowed to disagree with adults if they feel their own ideas are better	2.67	1.24	0.16
Most students in this school are not intelligent enough to do well	2.56	1.21	0.23
There is no point in spending a lot of time preparing for a class	2.34	1.1	0.34
It's all right for a child to disagree with me	2.23	0.95	-0.27

In summary, the intervention did not lead teachers to endorse child-centred teaching practices. This may not be surprising as such beliefs about the abilities and appropriate behaviour of children are often held strongly by teachers and parents. This aligns with research elsewhere. Teacher-focused, didactic academic instruction dominates ECE curricula in low-income countries (Baum, 2020), and there is good evidence to suggest that this is the instructional approach of choice for both parents and teachers (Kaul *et al.*, 2015), with a particular emphasis on rote memorisation for the purposes of acquiring traditional numeracy and literacy skills. To this end, pre-primary schools often aim to simply replicate the teaching practices of primary schools, with 'few concessions made to the stage of development of younger students' (Orkin *et al.*, 2012).In some instances, student-centred

[15] Note that even if social desirability bias is making teachers more inclined to agree with statements, it is reasonable to assume that this affects both teacher-centred and child-centred statements similarly, and so conclusions can be reasonably drawn about the significant difference observed.

[16] Positive values in the 'control–treatment diff' column suggest that teachers in control classrooms are more supportive of that particular statement, while negative values represent greater levels of support from treatment teachers. * indicates significance at p < .05

pedagogies are criticised as being 'too playful', with concern that they are not sufficiently focused on academic rigour (Edwards *et al.*, 2019). Given the prevalence and depth of some of these teacher perspectives, it may be unrealistic to expect such beliefs to be changed with a light dosage of teacher training. Overturning the preconceptions of teachers is likely to require much deeper exposure to such teaching practices and ideas.

6.2.3 Classroom observations

We also investigate whether participation in the programme influenced teachers' pedagogical practices in line with the national curriculum. This analysis is based on classroom observations at both baseline and endline. These observations recorded the frequency of instances of student-centred teaching practices, focus on gross motor skills or expressive language, working in pairs, student-centred activities undertaken by children, student engagement, the use of educational materials by the teacher, and the use of educational materials by students. Table 9 examines the impact of the programme on teaching practices, comparing treatment and control teachers.

Table 9: Impact on teaching practice

	Teaching practices					Student engagement	Teaching materials	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Overall	Student-centered teaching practices	Gross motor / expressive language	Pairs	Student-centered student activities	Degree of engagement	Teacher	Student
Treatment	0.040	-0.062	0.099	0.249**	-0.126	-0.060	0.328	0.233
	(1.767)	(0.661)	(0.939)	(0.112)	(1.019)	(0.123)	(0.342)	(0.324)
Observations	49	49	49	49	49	49	49	49
R-squared	0.150	0.093	0.067	0.155	0.082	0.104	0.107	0.156
Control mean	7.71	2.14	1.90	0.10	3.57	3.19	1.95	1.67

Note. Standard errors in parentheses, clustered at a school level. * p<0.10 ** p<0.05 *** p<0.01

We do not detect a statistically significant improvement in student-centred teaching practices, student engagement, or use of teaching resources. Across our eight different measures of teaching practices and learning activities, only one produced a difference between treatment and control groups: after participating in the intervention, teachers are significantly more likely to split students to work in pairs or small groups. However, this treatment effect is only significant at the .05 level, and with our testing of multiple classroom observation outcomes, it is possible that this significant effect is simply a product of random chance. The consistency across the remaining outcomes unfortunately suggests that participation in the intervention does not seem to have produced any difference in teaching practices within the classroom.

In summary, the intervention did not effect a change in teachers' pedagogical practices. As discussed above, this might not be surprising: teachers' practices, often developed over many years of experience, may be resistant to change, especially from relatively light interventions. Nonetheless, it is interesting that uptake of the national curriculum, which focuses on child-centred learning, was not sufficient to promote a child-centred pedagogy. This suggests that, in addition to improving access to the national curriculum, teachers will require additional and potentially intensive support in order to improve their teaching practices.

Key points (RQ2):

- Although school closures resulted in a light intervention, this was sufficient to increase uptake of the national curriculum.
- However, neither the intervention nor the increased uptake of the national curriculum effected a change in teachers' beliefs about child-centred learning, or their teaching practices.

6.3. Impact on over-age children

RQ3. Does the intervention reduce the number of over-age children enrolled in ECE?

RQ4. What proportion of over-age children are promoted to primary school?

At endline, we find that over-age ECE students in the 2019/20 academic year were promoted to primary grades or ALPs at significantly higher rates in treatment schools as compared to control schools. In total, 502 over-age ECE students were surveyed at baseline. At endline, the research team was able to establish contact with 380 (75.7%) of these students. Of the over-age students enrolled in control schools at baseline, 36.5% were promoted to primary grades or ALPs by endline (i.e. in the new academic year). In comparison, the promotion rate for over-age children in treatment schools was nearly 25 percentage points higher (61.1%). This difference is statistically significant (Pearson χ^2 (1) = 22.4319; $p < 0.001$), suggesting that the primary-level intervention did have a positive effect on the promotion of over-age students.

Data collected from teacher surveys on enrolments of over-age children at baseline suggest that current ECE classrooms across treatment (30.3%) and control (33.7%) classrooms had an equivalent share of over-age children ($p = .58$) prior to the start of the intervention. At endline (the start of the 20/21 school year), 95.7% of teachers in the control group report that their classroom has over-aged children, compared to 58.6% in treatment schools, a statistically significant difference (Pearson χ^2 (1) = 8.43; $p = 0.004$).

This suggests that the intervention was successful in reducing the number of over-age children enrolled in ECE. Whether schools will continue to enrol children of primary school age directly into primary school remains to be seen, however. Nonetheless, it is notable that a significant effect was achieved despite a disrupted (and therefore light-touch) intervention.

Key point (RQ3 and RQ4):

- Despite disruptions to the delivery of the intervention, the intervention was successful in reducing the number of over-age children enrolled in ECE.

6.4. Teachers' and principals' perspectives

6.4.1 Receptivity to further training

RQ7. Are principals and teachers receptive to receiving additional training and support?

Whereas the ECE curricular intervention under consideration was focused on providing materials and training to teachers for the purpose of strengthening their teaching practice, we asked teachers (both control and treatment) how open they would be to receiving additional training and support.

Responses were recorded across the range of individual teaching activities and subjects. Table 10 reports the results, sorted from most to least favourable (as reported by teachers). According to these results, teachers appear to favour additional support in content areas (numeracy/literacy) as opposed to pedagogy (free play/movement/gross motor/play-based). There are no differences in support for any specific teaching practices or content areas between treatment and control groups.[17]

Table 10: What areas of additional support do teachers favour most/least? (5 = strongly agree; 1 = strongly disagree)

Category	Mean	Std dev.
Maths/numeracy	3.84	0.417
Teaching one-on-one	3.80	0.483
Language/literacy	3.79	0.530
Expressive language	3.75	0.667
Fine motor skills	3.71	0.706
Play-based, child-centred learning	3.62	0.822
Using worksheets	3.61	0.755
Free play/open choice	3.61	0.679
Teaching in small groups	3.57	0.759
Gross motor activities	3.50	0.853
Music/movement	3.43	0.850

[17] Using Mann-Witney U comparison of means and independent samples t-tests.

Category	Mean	Std dev.
Teaching whole class/same time	3.21	1.140

Principals in treatment schools were asked about their receptiveness to additional training and support, specifically on the national ECE curriculum – 100% of these principals said they would be 'very interested' in receiving this type of support. Both principals and teachers in the treatment group were also very optimistic that other schools would want to receive the same training – 96.9% of teachers, 95.8% of principals – and either agreed or strongly agreed.

In summary, there is a clear demand from teachers and principals for further training. In particular, teachers' report being interested in more training on content rather than pedagogy. It is notable that the intervention, which focused on child-centred learning, may have focused more on pedagogy than content, and thus may not have met this need directly.

Key point (RQ7):

- Teachers and principals are receptive to further training, and may be specifically interested in support on content rather than child-centred pedagogy.

6.4.2 Perspectives on over-age children

RQ8. Are principals and teachers receptive to promoting over-age children to primary school?

Teachers and principals exhibit different perspectives on the extent to which over-age children pose challenges to the learning environment of ECE-age children. Principals appear to see the number of over-age children in ECE as a larger problem for pre-primary learning environments than teachers themselves (54% vs. 33%). Teachers are much more likely to suggest that over-age students pose no challenge whatsoever to their delivery of a quality ECE experience (54% vs. 31%) (see Figures 10 and 11).

Figure 10: Teachers: 'Does the number of over-age children make it difficult for you to teach effectively?'

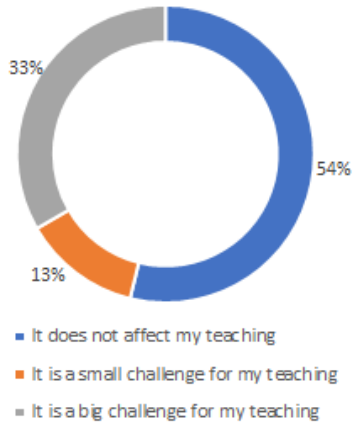
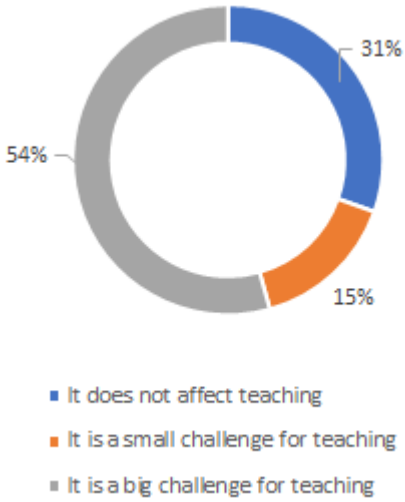
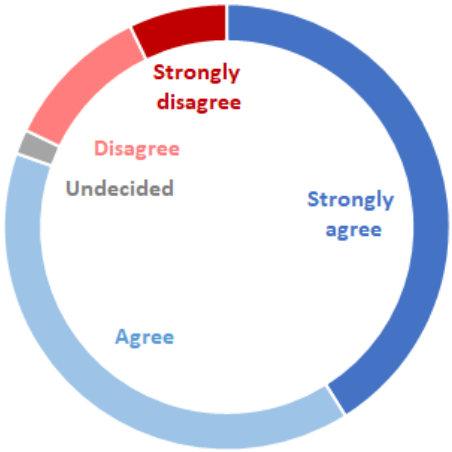


Figure 11: Principals: 'Does the number of over-age children make it difficult for pre-primary teachers to teach effectively?'

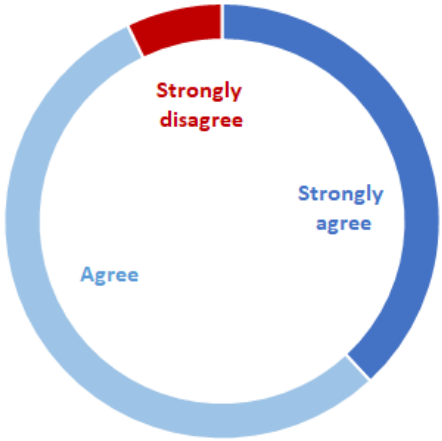


Principals also have a slightly more favourable view than teachers of the value of over-age promotion for the benefit of ECE-age students (Figure 12). As a whole, both teachers and principals appear to view promotion as less beneficial for the over-age children than the ECE-age children, with the views of principals and teachers being nearly identical on this point (Figure 13). There are no differences in treatment vs. control teachers or principals in their views on the benefits of promotion for pre-primary students (Mann-Whitney U test; p = 0.35 & p = 0.34). Interestingly, however, teachers who participated in the intervention are significantly less likely than control teachers to expect over-age children to benefit from promotion (p = .019), although there is no difference between treatment and control principals on this same question.

Figure 12: 'Promoting over-age children would be beneficial for the learning experience of pre-primary students'

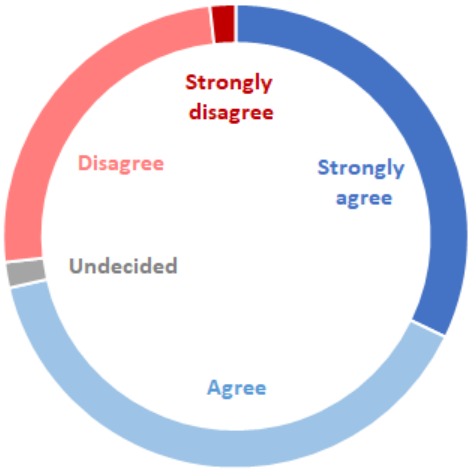


Teachers

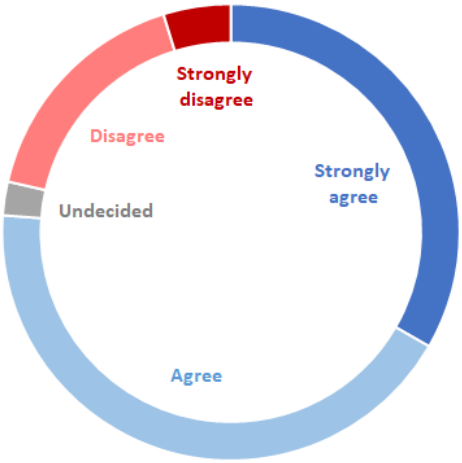


Principals

Figure 13: 'Promoting over-age children would be beneficial for the learning experience of those over-age students'



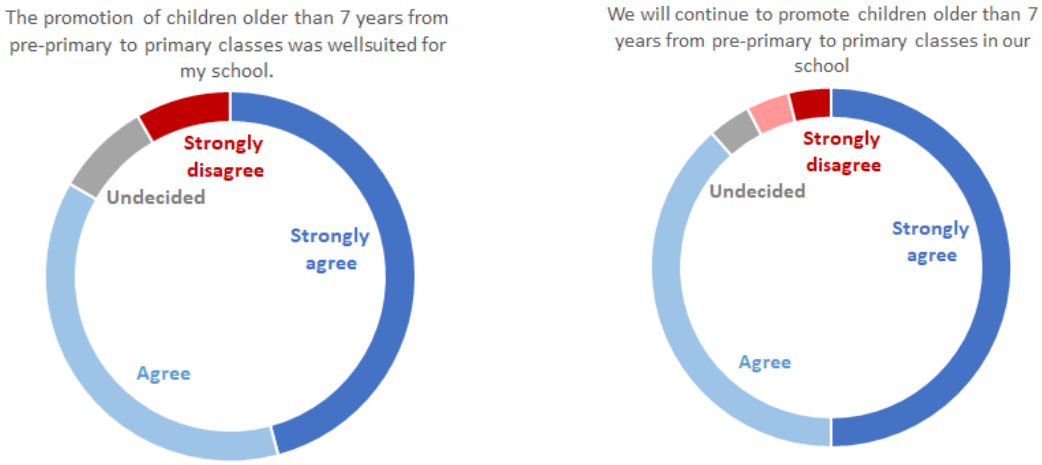
Teachers



Principals

Lastly, principals in the treatment schools largely responded that the promotion of over-age ECE students was well-suited for their schools and that they will continue to promote these over-age students to primary-level classes (Figure 14).

Figure 14: Principals' attitudes towards promotion of over-age children in intervention



In summary, while teachers and principals in both treatment and control schools are receptive to promoting over-age children to primary school, principals appear to see over-age children as a larger challenge for ECE classrooms than teachers themselves. Notably, promotion of over-age students is seen to be slightly more beneficial for those ECE-age students remaining in pre-primary than for the over-age children promoted; however, overall, a majority of principals and teachers view over-age promotion as being beneficial for both over-age and ECE-age children. This is interesting for two reasons. First, taken at face value, this is inconsistent with the view that over-age children are enrolled in ECE because they are 'not ready' for primary school. However, principals and teachers may believe that this concern is ameliorated if primary school teachers are equipped to support over-age children. Alternatively, principals may have other reasons for enrolling over-age children in ECE that override the interests of the child. Second, an assumption in the intervention rationale and design was that the presence of over-age children in ECE hinders the learning of children in ECE. This assumption does not appear to be shared by ECE teachers. This may reflect either an issue in the programme design itself, or the need to demonstrate to teachers the relevance of the intervention.

Key point (RQ8):

- Teachers and principals (in both treatment and control schools) are receptive to promoting over-age children. Promotion is seen as slightly more beneficial to children remaining in ECE, but still beneficial for over-age children.

6.4.3 Self-efficacy in ECE

RQ9. Through this intervention, do principals and teachers feel equipped to use the national ECE curriculum?

Teachers were also queried regarding their confidence in teaching the subjects and using

the pedagogies prioritised in the national ECE curriculum. There is some overlap in the level of teachers' confidence and their desire for more training. For instance, teachers report having the most confidence in their ability to teach numeracy; however, this is also the subject in which they most strongly desire additional training/support.

Moreover, teachers do not feel more confidence in, or a greater desire to improve, their teaching in the areas most closely related to the national ECE curriculum, such as play-based, child-centred, free play, and open choice pedagogies. Additionally, we find no differences in teachers' confidence in teaching these subjects or applying these pedagogies when comparing treatment teachers to control teachers.[18]

Table 11: How confident do teachers feel in their ability to effectively teach the following topics and pedagogies?[19]

Category	*Mean*	*Std dev.*
Maths/numeracy	3.7	0.57
Teaching one-on-one	3.62	0.776
Expressive language	3.59	0.804
Language/literacy	3.55	0.807
Free play/open choice	3.5	0.853
Teaching in small groups	3.46	0.873
Gross motor activities	3.46	0.894
Play-based, child-centred learning	3.43	0.931
Fine motor skills	3.36	1.017
Music/movement	3.18	1.064
Teaching whole class/same time	2.98	1.104
Using worksheets	2.96	1.144

In addition, we investigated principals' confidence in using their current curriculum. There initially appears to be a large difference between treatment and control schools in regard to principals' belief that 'the curriculum we currently use for pre-primary classes is a good fit for our school' – as indicated in Table 12, 88.9% of principals in the treatment group agree or strongly agree with this statement, compared to only 70.3% in the control group. Nonetheless, this is only significant at $p < 0.1$ ($p = 0.079$).

[18] Using Mann-Whitney U tests of means with Bonferroni adjustments for multiple comparisons.
[19] 1 = not confident at all; 2 = a little confident; 3 = somewhat confident; 4 = very confident.

Table 12: Principals: 'The curriculum we currently use for pre-primary classes is a good fit for our school'

The curriculum we currently use for pre-primary classes is a good fit for our school	**Control (%)**	**Treatment (%)**
Strongly agree	37.0%	59.3%
Agree	33.3%	29.6%
Disagree	22.2%	3.7%
Strongly disagree	7.4%	7.4%
Total	**100.0%**	**100.0%**

Table 13 reflects principals' agreement with the statement 'I feel that my school is equipped to use the current curriculum we are using'. There appears to be a large difference between treatment and control schools in this regard – 77.7% of principals in the treatment group agree or strongly agree, compared to 59.2% in the control group. However, this different is not statistically significant (p = 0.25).

Table 13: Principals: 'I feel that my school is equipped to use the current curriculum we are using'

I feel that my school is equipped to use the current curriculum we are using	**Control (%)**	**Treatment (%)**
Strongly agree	29.6%	40.7%
Agree	29.6%	37.0%
Disagree	29.6%	11.1%
Strongly disagree	11.1%	11.1%
Total	**100.0%**	**100.0%**

In summary, the intervention did not lead to teachers being more confident in delivering child-centred pedagogy. However, the intervention may have increased principals' confidence in the curriculum used by their school.

Key point (RQ9):

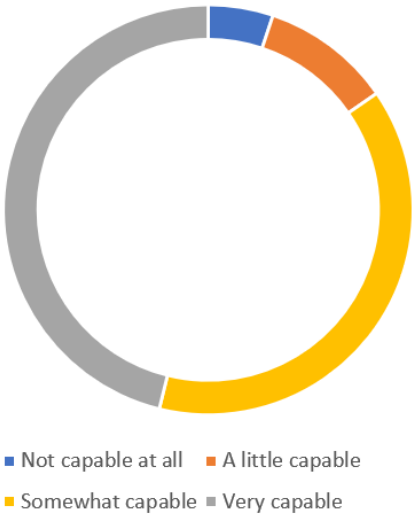
- The intervention increased principals' confidence in the use of the curriculum in their school, but it did not increase teachers' confidence in the use of child-centred pedagogies.

6.4.4 Self-efficacy in supporting over-age children

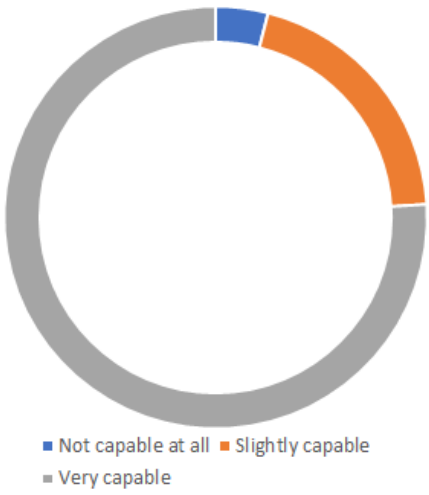
RQ10. Through these interventions, do principals and teachers feel equipped to support over-age children?

Most teachers and principals regard themselves as 'very' or 'somewhat' capable of meeting the needs of over-age learners in ECE classes (Figure 15). Upon testing for differences in the self-efficacy of teachers and principals as regards meeting the needs of over-age ECE students, we find no significant differences between treatment and control schools. On this 1–4 scale of self-efficacy (1 = not capable at all; 2 = slightly capable; 3 = somewhat capable; 4 = very capable), teachers and principals in treatment and control schools report nearly identical levels of confidence (mean treatment = 3.24; mean control = 3.27).

Figure 15: 'To what extent do you feel capable of meeting the learning needs of children older than seven years who are enrolled in your pre-primary classes?'



Teachers



Principals

Key point (RQ10)

- The intervention did not increase teachers' and principals' confidence in their ability to support over-age children. Principals are much more confident than teachers in their schools' abilities in this regard.

6.5. Cost of the intervention

RQ5. What is the cost of training teachers in the new programme and providing them with the necessary resources (e.g. school visits, school/classroom volunteers)?

RQ6. Will the programme entail any additional reoccurring costs, such as the hiring of new staff (e.g. volunteers used in the intervention, additional teachers for the movement of over-age students to primary), the development of materials, and the increased engagement with the community/parent–teacher association?

In order to calculate the costs of the intervention, we consider (a) the costs of delivering the training and mentoring to teachers, and (b) the costs of delivering the intervention in the classroom. Finally, we contextualise these costs with reference to the costs of delivering ECE calculated in ELP Phase 1.

6.5.1 Costs of delivering training and mentoring

Tables 14 to 16 present the costs of the training and mentoring. As described in Section 5.3.6, this reflects the cost of the intervention to both the service provider and the transport and lodging costs of teachers travelling from out of town. This assumes:

- two refresher workshops;
- 100 teachers recruited to the programme;
- a retention rate of 85% in the initial workshop and 85% in the refresher workshop and mentorship visits – and thus, an output of 73 teachers trained;
- eight mentorship visits per teacher; and
- that 90% of teachers attending the training would need to travel a long distance to the training site, and would need to lodge in town. The average cost of transport for teachers travelling from out of town is assumed to be US$ 50 per round trip. The average cost of lodging is assumed to be US$ 15 per day and assumes shared accommodation.[20]

To calculate the average cost per child, we assume an STR of 30:1, as per ELP Phase 1. While it is hoped that the training will improve teachers' practices throughout the remainder of their career, we limit our calculation to include only those children reached in the three years that follow the intervention. This is likely a conservative estimate, as teachers surveyed in the control group of the Centre for Global Development's study on partnership schools in Liberia reported having 15 years of teaching experience on average (Romero *et al.*, 2017). However, this estimate allows for the possibility of teacher drop-out, or the provision of further training.

[20] This is based on the average cost of travel to workshop venues during the evaluation, and the average cost of accommodation in the towns in which the training was held.

Table 14 presents the total cost of the training and mentoring, as well as the cost of each component. In sum, the total cost of delivering the intervention would be US$ 94,670 to train 73 teachers. The average cost per teacher would be US$ 1,297. The average cost per child reached would be US$ 14.41. The costs of the intervention are roughly equally split between the initial workshop, the two refresher workshops, and the mentorship programme.

Table 14: Total cost of intervention by each component

Component	Cost	Percentage of total
Initial workshop	US$ 30,799	32.53%
Refresher workshops	US$ 31,795	33.58%
Mentorship	US$ 32,076	33.88%
Total	US$ 94,670	
Total cost per teacher (based on 73 teachers receiving a full package of training)	US$ 1,297	
Total cost per child (calculated based on 30 children reached a year by a teacher over a three-year period)	US$ 14.41	

Table 15 presents a comparison between staff versus non-staff costs. The primary driver of cost is non-staff costs (covered below), rather than staff time.

Table 15: Comparison of staff and non-staff costs

Type of cost	Cost	Percentage of total
Total staff costs	US$ 28,333	29.93%
Total non-staff costs	US$ 66,337	70.07%

Table 16 presents the total non-staff costs of the training and mentorship, divided by category. Categories that account for more than 10% of the total cost are highlighted.

Table 16: Non-staff costs of workshop by category of expenditure

Categories	Total cost	Percentage of total
Materials provided to trainees	US$ 5,759	6.08%
Training materials (fixed)	US$ 255	0.27%
Training materials (variable)	US$ 1,803	1.90%

Categories	Total cost	Percentage of total
Trainee transport (total)	US$ 11,350	11.99%
Of which, trainee transport (short distance)	US$ 900	*0.95%*
Of which, trainee transport (long distance)	US$ 10,450	*11.04%*
Trainer transport (total)	US$ 9,995	10.56%
Of which, for workshops	US$ 1,870	*1.98%*
Of which, for mentorship	US$ 8,125	*8.58%*
Refreshments	US$ 8,430	8.90%
Venue rental	US$ 2,700	2.85%
Lodging for trainees from far areas	US$ 9,735	10.28%
Lodging for training staff	US$ 16,310	17.23%
Of which, for workshops	US$ 7,560	*7.99%*
Of which, for mentorship	US$ 8,750	*9.24%*

The largest driver of cost is travel and lodging for training staff, which together account for 27.79% of the total cost. This is primarily driven by travel and lodging during mentorship visits, which accounts for 17.82% of total costs. The second largest contributor to costs is travel and lodging for trainees travelling from out of town, which together account for 21.32% of the total cost of the intervention. In summary, the costs of the intervention are driven up by the travel time between schools and central training venues, which is a function of both distance and the quality of travel infrastructure. While this may be a concern for value for money, as travel costs do not relate directly to the intervention, it is unclear how this cost can be reduced while still reaching schools outside of urban centres.

'Fixed' training materials refers to equipment provided by or used at the venue as part of the training (projector and flipboards). Variable training materials refers to stationery used during the workshop. Materials provided to trainees include activity books, handouts, and teacher planners (and assumes that three of each type would be provided to each teacher, once in the initial workshop and again in the refresher workshops). Together, these materials account for a relatively small proportion of the total cost (8.25%). Refreshments, costed at US$ 10 per participant per day, also account for a small but significant proportion of total costs (8.9%).

Tables 17 and 18 explore the effect of different variations to the model on total cost and cost per teacher respectively. These concern retention (assuming either full retention, or a

decrease of only 10% at each stage), the number of refresher workshops (reduced to one), and the number of mentorship visits (reduced to six per teacher).

Table 17: Variations of model – total cost

Variation	Total cost	Percentage of original
Retention		
Full retention	US$ 115,439	121.94%
10% lost in initial workshop and again before refresher	US$ 100,108	105.74%
Refresher workshops		
One workshop	US$ 78,829	83.27%
Mentorship visits		
Six visits per teacher	US$ 87,253	92.17%

Table 18: Variations of model – cost per teacher

Variation	Cost per teacher	Percentage of original
Retention		
Full retention	US$ 1,154	89.02%
10% lost in initial and again before refresher	US$ 1,236	95.30%
Refresher workshops		
One workshop	US$ 1,080	83.27%
Mentorship visits		
Six visits per teacher	US$ 1,195	92.17%

While increasing the retention rate for teachers increases the overall cost of the intervention, it reduces the cost per teacher. If the intervention were to have full retention – which, although unlikely, illustrates the maximum by which this variable could be improved – the reduction in per teacher cost would be large (-10.98%). However, the reduction in costs is much more modest (-4.7%) if retention rates are only improved by 85% to 90%, which may be the most realistic assumption. The most significant gains may also be made by reducing the number of refresher workshops from two to one (-16.73%). A smaller reduction may be achieved by reducing the number of mentorship visits from eight to six per teacher (-7.83%). However, as demonstrated by the findings of this evaluation, the dosage of mentorship visits

is likely to be very important for the effectiveness of the intervention, and thus reducing the number of refresher workshops and mentorship visits may be ill-advised. These variations could be the subject of future evaluations.

6.5.2 Costs of delivering intervention in classroom

As stated previously, the intervention was developed in order to entail no additional costs to schools. It was anticipated that all materials needed to deliver the intervention would be provided in the training workshops and mentorship. Nonetheless, principals of treatment schools were asked about any additional costs associated with administering the intervention. Of the 27 treatment schools, eight principals reported incurring additional costs.

Key points (RQ5):

- In its original design, the intervention costs US$ 1,297 per teacher.
- This is primarily driven by travel and lodging for teachers to travel from their schools to the central training venue, and for mentors to travel from the central training venue to each school.
- The cost of the intervention could be reduced significantly by reducing the number of refresher workshops from two to one (leading to a reduction in cost by 16.73%). However, this risks reducing the effectiveness of the intervention.

We report these costs and average them across schools to estimate the average marginal cost of programme participation for schools (Table 19). Of the eight schools incurring additional costs, two hired additional staff for the purposes of running the programme. Beyond the staffing costs incurred by these schools, other schools incurred one-time costs for expenses such as transportation (four schools); mats (two schools); chairs, tables, and benches (two schools); teaching materials (one school); school infrastructure (one school); teacher incentives (one school); and stationery (one school). Non-staff costs varied between US$ 6 and US$ 503, with a mean of cost of US$ 113 between the eight schools reporting additional costs, and a mean cost of US$ 33 between all 27 treatment schools.

The two schools incurring staff costs spent US$ 4,068 and $1,079 respectively, which are significant. The first school hired two additional full-time teachers and the second school hired one full-time teaching assistant. Nonetheless, because the large majority of schools (25 of 27) did not incur any additional costs to participate in the intervention, the median is $0.00.

Table 19: List of all school staff and non-staff costs for delivering the intervention

School	Non-staff costs (US$)	Staff costs (US$)	Total costs (US$)
School A	115.18	4,058.00	4,173.18
School B	104.71	1,079.00	1,183.71
School C	29.32	0.00	29.32
School D	502.62	0.00	502.62
School E	6.02	0.00	6.02

School	Non-staff costs (US$)	Staff costs (US$)	Total costs (US$)
School F	94.24	0.00	94.24
School G	48.17	0.00	48.17
School H	6.02	0.00	6.02

Key points (RQ6):

- The vast majority of schools (70%) did not incur additional costs to deliver the intervention.
- Of those that did incur costs, the mean non-staff cost was US$ 113 per school.
- 93% of treatment schools did not incur additional staff costs.

6.5.3 Putting the cost analysis into context

In order to put the costs of the intervention into perspective, we compare this analysis to the cost estimates in the ESA (2016) and ELP Phase 1 (2018). In interpreting this analysis, it should be kept in mind that this comparison is not straightforward. While the costs calculated in the ESA and in ELP Phase 1 calculated the cost of providing a full ECE to a child, the costs of the intervention calculated in this report concern only the costs of a single training and mentorship programme. Thus, the costs of the intervention should not be understood as an alternative to the cost of full provision, but rather than an addition.

The ESA (2016) provides an estimated cost per child of US$ 24 per year, based on the current standards of provision, which assumes an STR of 53:1. ELP Phase 1 (2018) estimated the likely per child cost required to meet the standards of provision envisaged by the MoE, and assumed an STR of 30:1. The costs associated with two scenarios in ELP Phase 1 are listed in Table 20:

Table 20: Per child cost estimates from ELP Phase 1

Scenario	Annual cost per child to the school
S1: Qualified teachers, STR 30:1, but with student stationery and food paid for by families	US$ 67.41

Scenario	Annual cost per child to the school
S2: Qualified teachers, STR: 30:1, but with student stationery and food provided by the school[21]	US$ 222.78

In order to compare these figures to the intervention costs, we assume that schools will not incur additional staff expenses, as per the intervention design (and as per 93% of treatment schools in the evaluation). We combine the training and mentorship cost per child (US$ 14.41) with the mean non-staff costs per child (US $0.55).[22] This results in a per child cost of US$ 14.96.

While this initially appears to be a significant additional expense compared to the current provision of US$ 24 per child estimated by the ESA (2016), it is important to note that the ESA assumes a significantly larger STR (53:1) than the intervention (30:1). If the same STR (of 53:1) is used for the intervention, the cost per child decreases to US$ 8.71 per child. Nonetheless, this would reflect a significant addition to current expenditure.

The additional expense of the intervention is smaller when compared to S1 (US$ 67.41 per child) – but, as before, it remains significant. The additional cost of the intervention is marginal compared to S2 (US$ 222.78 per child), which arguably reflects the best estimate of provision as it includes expenses that are otherwise borne by the family.

Key point:

- The cost of the intervention is significant compared to what the state provides for ECE, but much less significant compared when compared to what ECE would cost if it was provided to the standards envisaged by ECE policy.

6.6. Quality of intervention

We rely on teacher reports to gauge the likely quality of the training delivered by BRAC. Unsurprisingly, teachers in the treatment group (84%) are much more likely to have participated in at least one ECE training session than teachers' in the control group (32%). Nonetheless, it is concerning that 16% of teachers selected into the treatment group at baseline did not participate in any ECE training. The reasons for this are unclear. BRAC reported (i) instances in which long-term illnesses affected participation, and (ii) instances in which schools reallocated a teacher who had been in ECE at baseline to another class.

The dosage of training received from BRAC, as reported by teachers, is 5.4 days on average (excluding mentorship visits) – this is slightly lower than the maximum five-day training and two-day refresher anticipated in the intervention design. Teachers in the treatment group received on average more observation visits than teachers in the control

[21] Providing children with food at school is a significant proportion of the cost of ECE provision (i.e. US$ 128 per child).
[22] This assumes an average of two classes per school, each with approximately 30 children.

group (4.2 vs. 2.4 visits).[23] This aligns with the expected number of mentorship visits due to COVID-19 (i.e. three visits); the difference between treatment and control is not statistically significant (p = 0.12). Unfortunately, fewer teachers in the treatment group reported that the observers of their classroom met with them after the lesson to discuss their teaching (55.6% in treatment, 75% in control). While this difference is not statistically significant (p = 0.2), it is nonetheless concerning – it is possible that BRAC visited teachers to observe their classes, but it did not always debrief with these teachers after each class.

Teachers' perceptions of the focus of BRAC's training varies considerably. The most frequently reported topics were classroom management (reported by 60.8% of schools) and developing lesson plans using the curriculum (56.5%). A minority of schools that received BRAC's training reported a focus on play-based learning (30.4%), numeracy (21.7%), and literacy (13%). It is worth noting, however, that 14 months had passed since BRAC's workshops and the endline data collection, and it is possible that this affected teachers' recall.

Most teachers either agreed or strongly agreed (88.6%) that BRAC had delivered the programme well. Unfortunately, a notable minority (8.6%) strongly disagreed with this statement, suggesting that some schools may have been underserved by BRAC. Notwithstanding this, the vast majority of teachers agreed or strongly agreed that the training was well-suited for their school (97.1%), and that they would continue to use the lesson plans and activities recommended by BRAC (94.2%).

Key points:

- The dosage of mentorship visits was significantly lower than planned, due to the COVID-19 pandemic. While most teachers endorsed the quality of BRAC's programme, a significant minority disagreed.

6.7. Impact of COVID-19 (Additional Research Questions)

6.7.1 Student re-enrolment

ARQ2. What has been the effect of COVID-19 on student re-enrolment?

To accurately assess the impact of COVID-19 on student re-enrolment after school closures, it would be necessary to compare enrolment data prior to school closures and after schools reopened, as well as the normal rate of drop-out between academic years. As this question was added to the study after the onset of the pandemic, such data were not collected. In lieu, data were collected from two sources: records of student absenteeism at endline for students included in the study at baseline, and principals' self-reports of changes in enrolment.

At endline, 65% of students enrolled in the study at baseline were present at school on the day of endline data collection. Girls were seven percentage points less likely than boys to be

[23] These visits may be from the school principal or VPI, or other organisations delivering training.

present. In the survey of principals, 53% of principals reported that student enrolment had increased, while 43% reported a decrease and the remainder reported that it had stayed the same.

Of those principals who reported a decrease, the vast majority (78%) believed this was because families could not afford school expenses. A majority (52%) of principals who reported a decrease in enrolment believed that this may have been because families had moved to another area. This may explain the variation in principals' perceptions of whether enrolment had increased or decreased; it is possible, for example, that families moved to pursue economic opportunities and this led to a decrease in enrolment in less economically viable areas and an increase in enrolment in areas with greater economic opportunity. We did not collect data on principals' perceptions of why enrolment had increased.

However, there are two very significant caveats to these findings. First, endline data collection was undertaken in February 2021. It is possible that many families may have not yet re-enrolled, but planned to do so later in the year: this may be especially so if families had moved. Second, these data do not consider the relative sizes of increases or decreases in student enrolment. It is possible, for example, that those schools that reported a decrease in enrolment experienced a large change, while schools that reported an increase experienced a comparatively small change. With these caveats, it may be safest to consider the answer to this research question as inconclusive.

Nonetheless, these data are similar to the data observed on student re-enrolment after the Ebola epidemic in 2014. In 2015, a survey by the World Bank found that only three-quarters of families of primary school-aged children had reported that their children had returned to school. A survey in 2016, however, found an increase in enrolment from pre-epidemic levels. The authors suggest that this may be because schools were discouraged from charging any fees or additional expenses, thus broadening access.

The Education Sector Emergency Response Plan, discussed in Section 1.3, noted the importance of school hygiene procedures in order to encourage families to re-enrol their children. In our survey of principals, 74% reported that the school had introduced new procedures since reopening (while, notably, 26% reported that there had been no change). Of those principals that reported new procedures, these primarily concerned hand hygiene (97%) and the wearing of masks (48%). Very few principals reported screening children or staff for COVID-19 symptoms (3%), limiting certain activities (5%), and limiting school sizes (8%). Only 21% of principals reported requiring staff and children to maintain a physical distance from each other.

Key point (ARQ2):

- As many as 35% of students in ECE may have been absent from school one to three months after schools reopened. However, more students may have re-enrolled after endline data collection.

6.7.2 Teachers' activities and retention

ARQ3. How has COVID-19 affected ECE teachers' ability to continue to provide play-based, child-centred ECE in schools after the end of lockdown? What has the effect been on teacher retention, especially for 'volunteer' teachers?

While the original research question focused on teachers' activities after schools reopened, we were also able to collect data on teachers' activities during lockdown. We divide this section into activities during lockdown, and activities subsequent to schools reopening.

During lockdown

The majority (61%) of teachers were unaware of any distance learning activities offered to either children or families by their school, government, or any other organisation during school closures. Of those teachers that were aware of such programmes, the most frequent responses to what these entailed were home visits (55%), TV or radio programmes (27%), and printed materials (22%). Of these teachers, the majority (50%) believed that fewer than 25% of their students were likely to have participated in these activities, while 73% believed that fewer than half would have participated.

Nonetheless, the majority (64%) of teachers reported having contact with children in their class during school closures. Of these, 92% reported having contact with at least some children in their class on at least a weekly basis. The most frequent reason given for this contact was providing educational information (89%), followed by checking how the child was doing during school closures (28%).

The vast majority (88%) of teachers did not undertake any professional training during school closures. Of the seven respondents that reported undertaking training, only one reported that this training included how to support children during closures.

Teachers' responses reported considerable variation in the payment of teachers' salaries during school closures. Approximately 38% of teachers reported not receiving a salary, while 36% received a full salary, and the remainder received a partial salary. The majority (54%) reported that these salaries were paid 'somewhat less frequently than normal', with a significant minority (14%) reporting that they had been paid 'much less frequently than normal'.

After lockdown

The vast majority of teachers sampled at baseline were confirmed to still be teaching at endline (79%), either at the same school (72%) or at a different school (7.3%). In contrast, 12.2% of teachers reported having stopped teaching. Of the 10 teachers that were no longer teaching, the most common reasons for leaving the profession were the need to care for family or children (three teachers) and concerns relating to COVID-19 (two teachers). The remainder either had to attend to their farms, had retired, were no longer interested in teaching, or had health concerns unrelated to COVID-19. Unexpectedly, all respondents who had stopped teaching were contracted teachers. Of the volunteer teachers in our

sample, 73% were still teaching in the same school and 27% were untraceable (and thus we do not know if they were teaching elsewhere or had stopped teaching).

The majority of teachers (66%) believed that how they taught in the upcoming year would be very different from how they had taught in the previous year. The most common reasons given for this were that they had instructions to teach differently (38%) and that children needed to catch up what they had missed in the previous year (35%). Significant proportions of teachers reported that they had learned new ways of teaching (24%), or that they thought children would need different support (22%). Only 8% of teachers reported that their teaching would need to change significantly due to difficult conditions. While these responses do not directly relate to the intervention being evaluated, they nonetheless provide an important insight into the future of teaching and learning in Liberia post-pandemic.

Key points (ARQ3):

- Most teachers were unaware of distance learning programmes offered during school closures, but still maintained contact with their classes.
- The payment of teacher salaries was severely disrupted during the pandemic.
- Teacher retention appears to be low (between 73% and 79%), but teachers may have returned to their posts after endline data collection.

6.8. Summary of key findings

In summary, the ECE programme increased uptake of the national ECE curriculum among teachers and principals (RQ2). While this may have increased teachers' and principals' beliefs in the appropriateness of the curriculum for their school, it did not lead to an increase in teachers' confidence in delivering play-based activities, as prescribed by the curriculum (RQ9). The intervention also appears not to have had a measurable effect on teachers' beliefs about child-centred learning, or their teaching practices. Overall, the intervention did not lead to an improvement in early literacy, early numeracy, or executive function when compared to the control schools (RQ1). Nonetheless, teachers and principals in the treatment schools intended to continue using the ECE curriculum materials. Both teachers and principals were receptive to receiving additional training and support, and believed that other schools would be interested in receiving the intervention (RQ7).

The intervention was successful in encouraging schools to promote over-age students (RQ3), and in reducing the number of over-age children enrolled in ECE (RQ4). However, the intervention did not increase principals' and teachers' confidence in their ability to meet the needs of over-age children (RQ10). Teachers and principals reported intending to continue promoting over-age children from ECE (RQ8). However, the majority of teachers in both treatment and control schools believed that the presence of over-age children did not pose any challenge whatsoever to their teaching of ECE children.

The vast majority of schools did not incur additional costs in order to deliver the intervention (RQ6). We estimate the cost of delivering the intervention to 27 treatment schools to be US$ 1,297 per teacher, or US$ 14.91 per child if delivered according to the original design of one initial workshop, two refresher workshops, and eight mentorship visits. This would be a significant addition to the current cost of ECE provision (US$ 24 per child per year), but a relatively small proportion of the per child cost estimated in ELP Phase 1 to be required for quality ECE (US$ 222 per child per year).

Due to the timing of endline data collection, this study found inconclusive evidence on students' re-enrolment rates (ARQ2). While only 65% of children enrolled in the study at baseline were present at endline, anecdotal evidence suggests that many children may have re-enrolled later in the year. Moreover, while approximately half of principals reported that enrolment in ECE had decreased in their school, approximately half reported that enrolment had increased.

At endline, 79% of teachers sampled at baseline were still confirmed to be teaching. While none of the teachers who were confirmed to be no longer teaching were 'volunteer teachers', 27% of volunteer teachers were not present at endline and could not be traced (ARQ3). Although the majority of teachers (66%) believed that their teaching would change significantly in the upcoming year, only 8% of teachers reported that this was due to new difficulties in teaching conditions. For the remainder, the most common responses were new guidelines for teaching or needing to focus on remedial classes. This suggests that the pandemic has not had a long-term negative impact on teaching conditions for most teachers.

7 Discussion

7.1. Impact of COVID-19 on the effects of the intervention

It appears likely that COVID-19 had a severe impact on the efficacy of the intervention. In the main, the dosage was reduced considerably, due to school closures, from eight to 12 mentorship visits to two to three mentorship visits. This reduced dosage is considerably less than the recommended practice of one mentorship visit per teacher each month that is cited in work that has synthesised global evidence on teacher training (World Bank, 2021, p. 2). As the purpose of the intervention was to change teaching practices from teacher-centred to child-centred, it may be considered unsurprising that such a fundamental change could not be effected with a low dosage of the intervention. It is likely that the pandemic may have had other deleterious effects, such as high levels of teacher stress and low motivation (exacerbated by loss of income), and few opportunities to practise new skills (due to school closures).

Moreover, COVID-19 significantly affected data collection. Due to the interruption of schooling during the 2019/20 academic year, endline observation of teachers occurred at the beginning of the subsequent academic year (2020/21). As a result, classroom observations took place for teachers who had been trained during the intervention but who were not necessarily still actively participating in the intervention. Granted, an intervention of this kind would hope to impact teacher practices beyond the minimal expectations of their required curricula (e.g. ideally teachers would see the value of student-centred and play-based learning and integrate these practices more permanently into their regular lesson planning, regardless of whether the school/principal expects these teaching approaches to be used). Notwithstanding, it is still possible that we may have observed more treatment vs. control differences had the intervention and data collection not been interrupted by COVID-19.

7.2. Impact of COVID-19 on teaching and learning

COVID-19 appears to have had a very significant effect on the provision of teaching. In addition to school closures, teachers estimated that very few of the children in their classes are likely to have participated in distance learning activities. Notwithstanding, the majority of teachers (64%) reported maintaining contact with children in their class, primarily to provide educational information (89%). However, the intensity and effectiveness of this support is not known.

Despite these difficulties, learning outcomes improved during school closures. While this may be expected as part of the natural development of young children, it may also raise important questions about the ‘value-add’ of the current standard of provision to children’s learning. Due to this study’s design, we are unable to answer this question.

7.3. Effectiveness of the MoE's COVID-19 response

A separate report, summarised in Section 1.4, has investigated the MoE's response to COVID-19. However, the results from the endline data collection also provide partial insights into the effectiveness of this response.

As described in Section 1.4, the MoE's primary response to promoting learning during COVID-19 has been the TbR initiative. Unfortunately, responses from teachers suggest very limited uptake of this programme, as fewer than 11% of teachers were aware of any radio- or TV-based initiative. Moreover, of those teachers who were aware of any efforts to continue teaching and learning during lockdown, most believed that only a minority of their children would participate in these activities. Key stakeholders in the education sector recognised that this may be a possibility: the Emergency Response Plan report, for instance, acknowledged limited access to radio as a key constraint, and stakeholders participating in the EIE TWG were aware that families with young children would struggle to participate in distance learning.

In reopening schools, the MoE undertook a concerted campaign to ensure that schools introduced hygiene measures. This appears to have been moderately successful, as 74% of principals reported introducing new measures. This has primarily focused on hand sanitising and the wearing of masks. However, more extensive measures anticipated by the MoE have had limited uptake: only 20% of schools practised social distancing, and only one school in the sample screened students and staff for COVID-19 symptoms.

7.4. Implications and recommendations

The disruptions from COVID-19 make interpreting the findings of this study complex. At a minimum, the study suggests that schools require relatively little support to increase their uptake of the national curriculum. However, uptake of the national curriculum does not necessarily lead to a change in teachers' beliefs about student-centred learning or their teaching practices. Similarly, schools appear to require relatively little support to reduce the enrolment of over-age children in ECE – chiefly, this may be due to an increased awareness of the benefit of age-appropriate enrolment for children who would otherwise be over-age.

However, neither the ECE programme nor the promotion of over-age children improved the learning outcomes of children enrolled in ECE. The low dosage of the intervention due to the pandemic may be the most likely reason for this: the reduction of mentorship visits from eight to 12 to only two to three per teacher was very substantial. The pandemic may have hindered the intervention in other respects too, such as affecting teacher retention and reducing the amount of time teachers had to practise their new skills. It remains a possibility that the intervention was ineffective due to the quality of training, or because the training did not focus sufficiently on the desired skills and behaviours. However, principals' and teachers' positive feedback on the quality of the programme, and their belief that other schools would want to receive the same programme, offers some evidence against this. Overall, there is a clear demand from schools for more training and support on ECE. As the costs of such a training programme are significant compared to current expenditure on ECE, it is important to ensure that such programmes demonstrate value for money in evaluations, prior to being scaled up.

7.5. Recommendations

Intervention

1. In addition to **improving access to the new ECE curriculum for teachers, complementary interventions will be needed to change teaching practices**. The uptake of the national curriculum by teachers and principals is unlikely to be sufficient to increase the use of child-centred pedagogies and to improve learning outcomes, without accompanying interventions such as training and support.

2. Any teaching training focused on changing teaching practices and teachers' attitudes towards child-centred pedagogies is likely to require a relatively **high dosage** to be effective. The eventual dosage used in this evaluation, which was limited by COVID-19, was likely too low to be effective, although further research would be needed in order to establish the efficacy of any higher dosage.

Emergency response

3. The MoE was correct to identify the need to ensure the continuity of salary payments as a priority. However, the success in achieving this continuity appears to have been only partial. We recommend the **MoE prioritise making improvements in payroll administration in order to increase resilience to shocks** in the future.

4. Despite being aware of the limitations of TbR during the Ebola epidemic, the EIE TWG saw no effective alternatives for responding to the COVID-19 pandemic at scale, due to limited transport and communications infrastructure. **The MoE's ability to respond to future shocks is likely to be similarly constrained unless this infrastructure is improved.** To achieve this, development partners of the MoE may wish to prioritise funds for this purpose.

5. In addition to improving transport and communications infrastructure, **the MoE may benefit from enhancing its capacity to disseminate information during school closures effectively**. A lack of awareness of government activities will otherwise remain a barrier to the uptake of available services.

Further research

6. Despite limited dosage and major interruptions, there were still some indications of initial successes in the uptake of elements of the intervention. Hence an **additional study is recommended in order to assess the intervention without the interruption of the pandemic**, before taking any decision on scaling up the full intervention. Such a study may include an additional component on social behaviour change to target teachers' attitudes towards child-centred learning.

7. Given the success of the intervention in reducing over-age enrolment despite the interruption of the pandemic, the additional study may consider testing a **shorter intervention aimed specifically at reducing over-age children in ECE**. This study should include an assessment of the impact of such an intervention on the learning outcomes of over-age children.

8. Since teachers' usage of the national curriculum did not correlate to a change in teaching practices, the additional study may consider the **inclusion of qualitative research to understand how teachers interpret and use the national curriculum**, and how teachers engage with training programmes that seek to change their attitudes towards child-centred pedagogy.

References

Baum, D. (2020) 'Non-state actors in early childhood education: Implications for education equity and quality'. Paper prepared for the Global Education Monitoring Report 2019. GEM Report Fellowship. UNESCO.

Biçer, N. (2017) 'The Influence of Student-Centered Methods in Turkish Language Instruction on Academic Success: A Meta-Analysis Study', *Universal Journal of Educational Research* 5(4), pp. 687–697.

Edwards, D. B., Okitsu, T., and Mwanza, P. (2019) 'Low-fee private schools, the state, and globalization: A market analysis within the political sociology of education and development', *Education Policy Analysis Archives* 27(133), pp. 1–42.

Himelein, K. (2015) 'The socio-economic impacts of Ebola in Liberia: results from a high frequency cell phone survey', World Bank, www.worldbank.org/content/dam/Worldbank/document/Poverty documents/Socio-Economic Impacts of Ebola in Liberia, April 15 (final).pdf [Accessed 25 October 2021]

Ilieva, I.P., Hook, C.J., Farah, M.J. (2015) 'Prescription Stimulants' Effects on Healthy Inhibitory Control, Working Memory, and Episodic Memory: A Meta-analysis', *J Cogn Neurosci.* 27(6), pp. 1–21.

JPAL (2012) *Costing Guidelines*. www.povertyactionlab.org/sites/default/files/research-resources/costing-guidelines.pdf [Accessed 25 October 2021]

Kaul, V., Chaudhary, A. B., and Sharma, S. (2015) 'Quality and diversity in early childhood education: A view from Andhra Pradesh, Assam and Rajasthan', Centre for Early Childhood Education and Development.

Kim, K., Robinson, N., Härmä, J., Jeffery, D., Rose, P., and Woldehanna, T. (2022) 'Misalignment of policy priorities and financing for early childhood education: Evidence from Ethiopia, Liberia, and Mainland Tanzania', *International Journal of Educational Research* 111.

Liberia Institute of Statistics and Geo-Information Services (2017) *Household income and expenditure Survey 2016.* Liberia Institute of Statistics and Geo-Information Services, Monrovia, Liberia, http://microdata.worldbank.org/index.php/catalog/2986/download/41608 [Accessed 25 October 2021]

MoE (2016a). *Liberia Education Sector Analysis*. https://documents1.worldbank.org/curated/en/481011575583469840/pdf/Liberia-Education-Sector-Analysis.pdf [Accessed 25 October 2021]

MoE (2016b) *Getting to Best Education Sector Plan 2017–2021*. https://www.globalpartnership.org/sites/default/files/getting_to_best_education_sector_plan_2017-2021._liberia.pdf [Accessed 11 December 2021]

MoE (2020) *Liberia COVID-19 Education Emergency Response Plan.* http://moe-liberia.org/wp-content/uploads/2020/05/Liberia-COVID-19-Education-Response-Plan_24-May_2020.pdf [Accessed 25 October 2021]

Neuman, M. J., and Powers, S. (2021) 'Political Prioritization of Early Childhood Education in Low-and Middle-Income Countries', *World Bank Policy Research Working Paper 9549*, World Bank.

Nicolopoulou, A., Cortina, K. S., Ilgaz, H., Cates, C. B., and de Sá, A. B. (2015) 'Using a narrative-and play-based activity to promote low-income preschoolers' oral language, emergent literacy, and social competence', *Early Childhood Research Quarterly* 31, pp. 147–162.

Nunes, T. and Bryany, P. (1996) *Children Doing Mathematics*. Blackwell, Oxford.

OPM (2019) 'ELP Systems Research: Liberia Diagnostic Report', OPM, Oxford.

OPM (2021) 'ELP Systems Research COVID-19 Response: Liberia', OPM, Oxford.

Orkin, K., Yadete, W. A., and Woodhead, M. (2012) 'Delivering quality early learning in low-resource settings: Progress and challenges in Ethiopia', *Working Papers in Early Childhood Development No. 59.* Bernard van Leer Foundation, The Hague, Netherlands.

Romero, M., Sandefur, J., and Sandholtz, W. (2017) 'Can Outsourcing Improve Liberia's Schools? Preliminary Results from Year One of a Three-Year Randomized Evaluation of Partnership Schools for Liberia', *CGD Working Paper 462*, Center for Global Development, Washington, DC.

Roskos, K. A., Christie, J. F., Widman, S., and Holding, A. (2010) 'Three decades in: Priming for meta-analysis in play-literacy research', *Journal of Early Childhood Literacy* 10(1), 55-96.

Schaefer, E & Edgarton, M. (1985) Parental and child correlates of parental modernity. In *Parental Belief Systems: the Psychological Consequences for Children* (ed. I.E. Siegel), 287-319.

Stagnitti, K., Bailey, A., Hudspeth Stevenson, E., Reynolds, E., & Kidd, E. (2016). An investigation into the effect of play-based instruction on the development of play skills and oral language. *Journal of Early Childhood Research*, *14*(4), 389-406.

UNESCO (2012). Education For All Global Monitoring Report. Paris: UNESCO.

UNESCO, UNICEF, Brookings Institute and the World Bank. (2017). *Overview MELQO: Measuring Early Learning Quality and Outcomes.* Paris: UNESCO.

USAID (2021). *Language of Instruction Country Profile: Liberia*. Washington, DC: USAID. Available: [Accessed 25 October 2021]·

World Bank (2021). *Technical Guidance Note: Structuring Effective 1-1 Support*. Washington, DC: World Bank.

Annex A Terms of reference

ELP SYSTEMS RESEARCH PROGRAM – LIBERIA PHASE 2

A. PROJECT BACKGROUND AND OBJECTIVES

The Early Learning Partnership (ELP) is a multi-donor trust fund, managed by the World Bank, which works with countries to improve early learning opportunities and outcomes for young children, through both research and operational support.

With the support of the UK Department for International Development (DFID), ELP has launched the ELP Systems Research program. The program's objectives are (1) to provide policymakers in a set of focus countries with actionable information to help guide the delivery of quality, equitable early learning at scale; and (2) to build the international evidence base in the emerging field of systems research in early childhood education. The program is supporting one team of consultants for each of the following focus countries/regions: Ethiopia; Jamaica; Liberia; the province of Punjab, Pakistan; and Tanzania.

The focus of the program is the political economy and governance of early learning service provision. The program seeks to generate knowledge on what works to remove institutional and systemic barriers to the provision of early childhood education at scale, and on what reforms could create the right institutional incentives to improve ECE provision. The program is guided by a framework adapted by the research teams from Pritchett (2015).[24] The framework understands a system as a set of accountability relationships between the state, specialized agencies, frontline service providers, and parents/citizens and students.

Within this overall framework, four cross-cutting research themes have been developed:

- Quality at scale: What works to improve quality in different early learning environments (e.g. home, school, community-based)? How can these strategies be implemented at scale?
- Equity: What strategies should countries employ to reach children most in need (including inequalities by disability, socioeconomic status, geography, language, ethnicity, and gender), both to promote equity in access and equity in outcomes?
- Cost-effectiveness: What are the costs of different services, policies, and system-level reforms? Which country initiatives have potential to deliver learning outcomes at scale cost-effectively? How can countries drive toward the most cost-effective ECE service provision?
- Non-state sector: What is the current level of coverage by non-state providers, and how do governments engage with them? What are the most effective ways to regulate and finance the non-state sector? What drives parental demand for ECE services from non-state providers?

[24] Pritchett, Lant. 2015. "Creating Education Systems Coherent for Learning Outcomes: Making the Transition from Schooling to Learning." RISE Working Paper 15/005.

The program is being conducted in two phases:

- Phase 1 (2017-18) supported an early learning system diagnostic in each focus country. Depending on the country, Phase 1 activities included a combination of situation analysis, quality and outcome measurement, cost analysis, actor mapping, and analysis of non-state provision.
- Phase 2 (2019-20) will fund research that supports and evaluates system improvements in the focus countries. These Terms of Reference are for Phase 2 activities in Liberia.

B. SCOPE OF WORK

In Phase 2 of the ELP Systems Research Program, the consultants will focus on evaluation of one or more system-level interventions or reforms in Liberia. The consultants should be guided by precisely defined key research question(s) which:

- Reflect priorities of the government of Liberia, i.e. are questions that government agencies and other stakeholders responsible for early learning service delivery want answered.
- Build on the findings of the country diagnostic conducted in Phase 1 and one or more of the tools adapted during Phase 1.
- Concern changes in one or more of the accountability 'design elements' (delegation, finance/resources, information, motivation) explored in the recently developed ELP Systems Research framework adapted from Pritchett (2015); rather than only the "proximate determinants of learning" such as materials, additional teachers, or pedagogies.
- May involve, where possible, attribution of causal impact through experimental or quasi-experimental methods. However, proposals may make the case that process evaluations or other non-impact-evaluation strategies will be more feasible and/or informative.
- May include iterative/rapid-cycle process evaluation or technical assistance or capacity building to the government, provided there is a strong research rationale and linkage to the key research questions.

The specific intervention and research questions to be evaluated, as agreed upon by the consultants and ELP, are summarized in section (E). Modifications to these interventions/research questions may be made by mutual agreement between the consultants and ELP.

C. DELIVERABLES/SPECIFIC OUTPUTS EXPECTED FROM CONSULTANT

The key deliverables for the consultancy are as follows:

- Detailed work plan including (1) a mapping from a systems framework to specific research questions/hypotheses to specific research activities/tools; (2) plans for fieldwork including sampling procedures and sample size calculations as appropriate; (3) detailed (month by month) timeline of activities including key interactions with country stakeholders.

- Signed contract with service provider to implement the intervention to be evaluated, as described in section (E).

- Government partnership case study: Consultants will produce a 2-3 page case study of a significant aspect of their work with Liberian government. The topic must be agreed upon with the ELP team in advance, but potential topics include adaptation of tools (e.g. MELQO) for use by the government, implications of and/or efforts to address capacity limitations in the government, and key government decisions informed by the research. The case study should describe key roles, challenges, outcomes, and lessons learned. The case study may pertain to the research team's Phase 1 work if a suitable topic exists.
- Final evaluation report(s), in-country presentation, and raw data. Final report(s) should be written for an audience of Liberian stakeholders and include an executive summary. Detailed or highly technical analysis may be included as an appendix. The final results should be presented in country to the government, World Bank team, and other key stakeholders.
- Publications: Consultants must lead – defined as one or more team members being lead author responsible for publication – at least two publications to be submitted to peer-reviewed journals and/or working paper series. One of the two required publications may be specific to Liberia, e.g. a writeup of the evaluation results; at least one must be a cross-country paper in collaboration with one or more other country research teams. One of the papers may be substantially or wholly based on data collected in Phase 1. It is anticipated that Publication 1 will be a paper on financing of ECE across focus countries, and that Publication 2 will be an academic paper on the evaluation results, but the topics may be modified by agreement between the consultants and the World Bank. In addition, consultants are expected to participate – defined as one or more team members being a co-author and providing data and analysis from Liberia as needed – in three publications led by other ELP Systems teams.

Acceptance Criteria for Deliverables

- Contract with service provider: contract signed and scope of work integrated into research questions as discussed in section (E).
- Government partnership case study: previously agreed-upon topic; key roles, challenges, outcomes, and lessons learned written in a manner to be informative outside the immediate context of the case study.
- Evaluation reports, in-country presentation, and raw data: Technical quality of the research, defined as suitability for publication in a peer-reviewed journal; responsiveness to the country priority issues identified through Phase 1 an the proposal development process described in section E; evidence of completion of in-country presentation (e.g. video, attendance by ELP team member); completeness and deidentification of raw data.
- Publications: each paper submitted to a journal or working paper series, with one or more team members as lead author.

D. SPECIFIC INPUTS TO BE PRESENTED BY THE CLIENT

ELP team members are available throughout the work period to provide advice to the consultants, particularly on managing relationships with the Ministry of Education and other stakeholders. The ELP team is available to assist with the organization of the final in-country workshop.

ELP will organize up to two workshops for all country research teams during Phase 2. The workshops will provide an opportunity for collaboration on cross-country work, and external advisors who are specialists in specific aspects of research may conduct training sessions based on demand from country teams. Advisors may be made available for additional

training and technical support beyond the workshop based on demand from the consultants and the other country teams.

E. SPECIAL TERMS & CONDITIONS / SPECIFIC CRITERIA

In response to the Phase 1 findings, and with the explicit aim of addressing the system constraints related to finance and training as well as rewards/incentives, Phase 2 will support the design, pilot, and evaluation of two interventions:
(i) A pre-primary program to improve the quality of ECE instruction in order to better prepare children for starting school;
(ii) A primary school program to assist teachers in primary schools to use a 'Teaching and the Right Level' (TARL) approach in order to support the learning of overage children who had been moved from the pre-school to primary school.

This approach will address the twin problems of low quality and high overage enrolment. The consultants' work will include:
- Subcontracting a service provider (SP) to conduct small-scale pilots of the interventions and inform roll-out to approximately 27 schools (and 54 classrooms – 27 primary and 27 pre-primary) during the 2019-20 academic year.
- Performing a rigorous evaluation of the larger scale pilot to determine whether the pre-primary intervention has the potential to work at scale. Note that consultants will only be responsible for conducting a rigorous evaluation of outcomes on pre-primary aged children (up to age 6).

Based on discussions with the Government of Liberia and the World Bank, the following steps are expected, which may be revised by mutual agreement between the World Bank and the consultants:

Step 1: Consultants design a ToR and selection criteria for the implementing service provider (SP) and conduct a procurement process for the intervention. The SP would need to conduct a review of available materials and engage the MOE actively to understand how they can be used to inform both interventions. This involves a close look at the curricula, building on the in-depth review in Phase 1, an overview of existing textbooks, lesson plans, teacher guides, etc. available for ECE programs, primary school programs, and TARL programs. SPs should propose how much innovation is required for each intervention. SPs should start identifying the sample of schools where the approach might be tested and an approach to quickly establish how the program can be adapted before the larger-scale roll out in September 2019.
Step 2: Select the SP, jointly with the MOE and the World Bank.
Step 3: The Liberian academic calendar is structured around 6 periods, with breaks in between. The 5th period of the year starts on April 15th and ends on May 31st 2019. The 6th period starts on June 3rd and ends on July 19th. SPs should start testing materials and training approaches in May, iteratively revising through collecting views from teachers, head teachers, students, and parents until mid-June 2019.
Step 4: From June 2019 to September 2019, the SP will design the full suite of materials necessary for the larger scale pilots of both interventions.
Step 5: Roll-out interventions from September 2019 and carry out baseline assessments of learning in English Language and Mathematics at the pre-primary level, for the ECE intervention.
Step 6: Conduct assessments of implementation fidelity, with particular focus on the ECE intervention, from September 2019 and June 2020, and focus on engaging the MOE closely to ensure longer-term sustainability and attracting additional funding for the intervention. Document learning in interim reports.
Step 7: Conduct endline evaluation of the ECE intervention in June 2020 and submit final reports and deliverables by August 2020.

Annex B Detailed methodology

Sampling

Power calculations

The number of treatment (27) and control (27) schools was determined *a priori* through power and sample size calculations, estimating the number of schools required to measure likely significant effects, given the existing resources for the evaluation. And, as described in the subsequent section, the evaluation design involves the assessment of 15 students within each of these 54 schools (7 seven or eight in each of two ECE classrooms). However, the number of students per school could decrease for two reasons: (i) the sampled schools have fewer ECE students than in previous years (according to the EMIS data), or (ii) the study experiences student attrition between baseline and endline. As such, we considered the possibility of losing up to five students per school, for a total of 10.

The number of schools and students, combined with two other parameters – the intracluster correlation coefficient[25] (ICC) and the amount of explained variation in the statistical models (R^2) – will determine the study's MDES. The MDES varies according to our estimates of the evaluation's parameters and it is not possible to perfectly predict the values for these parameters.

Given our best (but also conservative) estimates for each of these parameters (ICC = 0.45, R^2 = .75, # students = 20, # schools = 54), our expected MDES was 0.32 SD, which is in line with prior studies of ECE interventions in low- and middle-income countries. The other lines of Table 21 show possible changes in the MDES, given variations to the ICC, R^2, or number of students surveyed, which fall outside the control of the evaluators. The evaluation team took steps to minimise attrition of participants, including the collection of students' home information for the purposes of following up with students who are not found in the school at endline. We expect a maximum attrition rate of 33%, which is captured by the decrease from 15 to 10 students. The biggest risk to the MDES would be a small amount of explained variation (R^2) in our statistical models. However, we have used a conservative estimate for the possible ICC coefficient (ρ = .45), benchmarked against prior evaluations in similar contexts, including one prior study on education in Liberia. If we happen to find lower between-school variation in student outcomes (e.g. down to an ICC of .25), the result would be an increase in statistical power, and thus a reduction of the MDES. Our estimate for the squared correlation coefficient is moderate in size (R^2 = .75). However, we feel this is a justifiable estimate for two reasons: (i) our final models will include a baseline measure of the outcome variable; and (ii) the stratified random sampling procedure is expected to increase the explanatory power of our regression models. If the R^2 in our final models turns out to be lower than expected (for example R^2 = .50), the result will be an increase in the MDES.

[25] The ICC represents the proportion of variation in the outcome that exists between schools, as opposed to between students within a school.

Table 21: Power calculations: estimated *a priori*

ICC	R^2	Number of students per school	Number of schools	MDES
0.45	0.75	10	54	0.32 SD
0.45	0.75	15	54	0.30 SD
0.25	0.75	15	54	0.26 SD
0.45	0.50	15	54	0.40 SD
0.25	0.50	15	54	0.32 SD
0.45	0.50	10	54	0.41 SD

Table 22: Power calculations: observed *post hoc*

Subject	ICC	R^2	Number of students per school	Number of schools	MDES
Literacy	0.22	0.56	14.5	54	0.30 SD
Numeracy	0.28	0.54	14.5	54	0.33 SD

Random selection and assignment of schools

We employed a stratified random sampling procedure – stratifying by school size (i.e. number of ECE-age students) and urban/rural status (in Bomi only, as there are no urban schools in Gbarpolu) – to ensure a balance between treatment and control groups. These two criteria were used to create nine strata of urban and rural schools of similar sizes in Bomi and Gbarpolu. Within each stratum we randomly selected:

- three schools for assignment to treatment status;
- three schools for assignment to control status;
- one school to serve as a treatment replacement (in the case that a treatment school in that stratum does not participate); and
- one school to serve as a control replacement (in the case that a control school in that stratum does not participate).

This procedure resulted in a sample of 27 treatment schools and 27 control schools, with nine potential treatment replacements and nine potential control replacements. Table 23 indicates the initial balance tests, which suggested that the treatment and control groups

have no statistical differences in the number of students, urban/rural status, student age, share of over-age students, or share of female students.

Table 23: Treatment/control balance tests

Variable	Difference	t	*P*
Number of students	3.29	0.77	0.45
Number of ECE-age students	6.74	1.22	0.23
Number of over-age students	-2.18	-0.35	0.73
Rural	.037	0.41	0.68
Female	-.014	-0.74	0.46

Random selection of teachers/classrooms and students

The random selection of teachers was conducted by enumerators upon their initial visit to the school, by using the following protocol:

- If the school's ECE classes are not divided by grade level (KG I, KG II, nursery, etc.) or student age, two teachers were selected alphabetically by surname.
- If the school's ECE classes are divided by grade level, one KG teacher and one nursery teacher were selected alphabetically by surname.
- If classrooms are divided by age groups, one teacher of five- to six-year-olds and one teacher of three- to four-year-olds was selected alphabetically by surname.

Following the selection of two ECE teachers/classrooms, a random sampling of seven or eight students was taken within each classroom (for a total of 15 per school). Within each classroom, the lead enumerator wrote a list of children aged three to six using one of the following methods:

- a class list with recorded ages/birthdates, with the teacher confirming who is present; or
- a class list without ages/birthdates, with the teacher validating the ages of each child and who is present; or
- by asking all children aged three to six to come forwards to record their name and age.

In the first class, eight students from the list of three- to six-year-olds were randomly selected for inclusion. In the second class, seven students from the list of three- to six-year-olds were randomly selected for inclusion. When there was an insufficient number of children between the ages of three and six in a class, we randomly sampled seven-year-olds (using the same method) to fill in the gap (and then eight-year-olds, etc.).

Data collection

Learning outcomes

In ELP Phase 1, assessment of the achievement of preschool children was conducted by using an adapted version of the Measuring Early Learning Quality and Outcomes (MELQO) direct child assessment. The adapted assessment focuses on domains relating specifically to primary school readiness: literacy and language, mathematics, and executive function. The process of defining the constructs and contextualising the instruments during Phase 1 involved a review of national curricula and assessments, discussions with key stakeholders in Liberia's ECE sector, a study of classroom recordings made in early learning environments, and pre-testing of assessments in schools.

The instruments developed for Phase 1 were used for Phase 2, with some slight improvements based on the field experience from piloting the instruments in Phase 1, and given the specific research needs of Phase 2. These improvements are listed for each construct below:

- **Literacy** ─ For Phase 2, we added easier items to the literacy section from Phase 1 to minimise the instances of floor effects. To this end, we added an additional item assessing print awareness, and we introduced easier reading and writing words.
- **Numeracy** ─ For Phase 2, we added in some items of varying difficulty levels to ensure that the item–person map resembles a normal distribution, without much clustering, floor, or ceiling effects. To this end, we added some additional items on measurement vocabulary and we made changes to items assessing additional and subtraction skills.
- **Executive function** ─ Two items were used to assess executive function during Phase 1: forwards digit span, where children are asked to repeat a sequence of numbers spoken by the examiner, and backwards digit span, in which children are asked to repeat the numbers in reverse order. However, these items did not perform very well in Phase 1, as these items proved to be very difficult for a majority of the children. For Phase 2, the team implemented a different executive function module. This module assessed working memory and inhibitory control, as in Phase 1, but with a total of three revised items. Two of these items are from IDELA and have been designed specifically for ECE-aged students.

Additional data collection instruments

The following list provides an overview of each additional data collection instrument, and what was measured.

- **Classroom observation (baseline and endline)**
    - The classroom observation instrument is adapted from the MELQO: MELE tool, and includes the following:
        - evidence of student assessment, participatory learning, activity-based learning, play-based learning, and a flexibility approach in learning; and

        - the proportion of time that students are engaged in the classroom.

- **Teacher survey (baseline and endline)**
    - Basic demographics: age, gender, education, training, experience, experience teaching at schools.
    - Implementation (i.e. did the teacher participate in the different components of the programme?): five-day training, quarterly refresher training, visits by mentors/coaches.
    - Use of the programme and national curriculum.
    - Interaction with parents.
    - Number of hours spent teaching in a week.
    - Class size, and age range in class.
    - Teacher motivation.
- **Document inspection (endline)**
    - Availability of resources provided by the programme: lesson plan book, class register, evaluation register, ECE reading books, and stationery.
    - Evidence of regular assessment of student performance in the evaluation register.
    - The print-richness of the class environment.
- **Principal interviews (endline)**
    - Reception of the ECE programme.
    - Reception of the promotion of over-age children.
    - Estimate of the recurring costs of the programme.
- **Service provider interviews and document review (endline)**
    - Barriers to the scaling up of the programme.
    - Review of the financial and human resources cost of the programme.

Household assets

Household wealth was measured by asking children whether their households possess a series of assets, including the following: radio, mobile phone, refrigerator, television, sofa, table, watch, bed, cupboard, clothes iron, fan, car, and generator. This set of household assets was selected based on the items' reliability (Cronbach's alpha = .70) and their relationship with household income/consumption (Pearson's correlation $\rho = 0.47$) in Bomi and Gbarpolu in Liberia's 2015 National Income and Expenditure Survey. Students identified the presence of these assets in their homes after being shown pictures of each item.

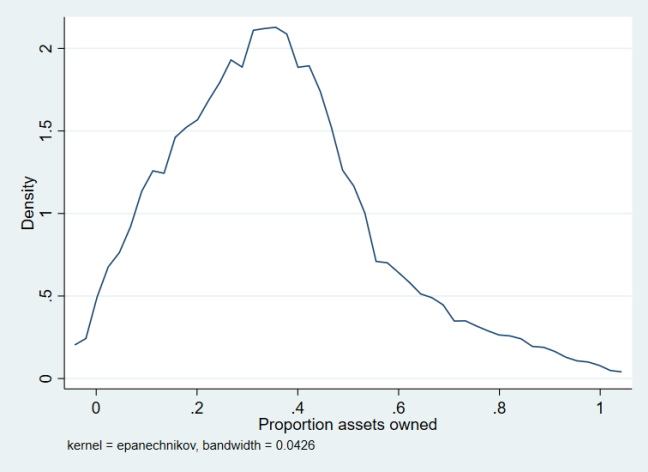
Student demographic characteristics are described in Table 24; as shown here, the average age of students included in the study is 4.99; 52% of the participants are female; and, on average, households possess 35% of the items in the asset index (Figure 16).

Table 24: Student demographic characteristics

Variable	N	Mean	SD	Min	Max
Age	787	4.99	1.07	3	7
Female	787	0.52	0.49	0	1

Household assets (proportion owned)	620	0.35	0.20	0	1

Figure 16: Proportion of household assets owned



Analysis

Balance and attrition

A complete summary of balance characteristics is attached in an Excel file separately to this report. Table 25 further examines the characteristics of attriters vs non-attriters, and tests whether these vary by treatment, using the following estimating equation:

$$x_{i,s} = \beta_0 + \beta_1 Attrite_i + \beta_2 Treat_s + \beta_3 (Attrite \times Treat)_{i,s} + \epsilon_i$$

Where $x_{i,s}$ is the characterstic of interest (gender, age, or learning), attrite is a dummy variable equal to one if the student could not be assessed at endline, and Treat is the treatment dummy. The coefficient β_1 tests whether attrition was random (i.e. the characteristics of the attriters are different to the non-attriters in the control). The coefficient β_2 tests for baseline balance for the non-attriters, and β_3 tests whether the composition of attriters is any different in the treatment vs control group. We see from Table 25 that attrition was non-random, and the sample of non-attriters is balanced on observable characteristics. This further strengthens the internal validity of the study.

Table 25: Attrition analysis (student level)

	(1)	(2)	(3)	(4)
	Female	Age	Numeracy (SD)	Literacy (SD)
Attrition	0.042	0.023	0.010	-0.034
	(0.077)	(0.147)	(0.144)	(0.150)

Treatment	0.006	0.064	0.051	0.053
	(0.045)	(0.140)	(0.129)	(0.146)
Attrition x treatment	0.010	-0.182	-0.180	-0.083
	(0.098)	(0.179)	(0.195)	(0.188)
Observations	787	787	787	787
R-squared	0.002	0.002	0.003	0.002
Non-attriters mean	0.51	5.01	0.02	0.02

Note. Standard errors in parentheses * p<0.05 ** p<0.01 *** p<0.001. The attrition rates are 21% in both the treatment and control arms.

Table 26 repeats the analysis above, but for teacher-level characteristics. There is some evidence of non-random attrition: the teachers who dropped out of the sample at endline were more likely to be male, and exhibited worse teaching practices at baseline: they implemented fewer teacher-centred practices, and students in their class were less likely to work in pairs or to be engaged in the classroom. However, the coefficients on 'treatment' show that the sample remains balanced along these characteristics. Most of the coefficients of teaching practices are negative, although they are not statistically significant.

Table 26: Attrition analysis (teacher level)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Female	**Student-centred practices**	**Expressive language**	**Working in pairs**	**Child engagement**	**Teaching materials for teacher**	**Teaching materials for student**
Attrition	-0.198	-0.857	-1.040	-0.286**	-0.414*	-0.365	0.103
	(0.181)	(0.541)	(1.359)	(0.105)	(0.206)	(0.395)	(0.442)
Treatment	0.131	-0.131	1.012	-0.143	-0.280	-0.214	0.583
	(0.160)	(0.578)	(1.681)	(0.125)	(0.175)	(0.390)	(0.417)
Attrition x treatment	0.147	-0.202	-2.290	0.143	0.540	0.770	0.194
	(0.247)	(0.618)	(2.002)	(0.125)	(0.300)	(0.527)	(0.612)
Observations	85	85	85	85	85	85	85
R-squared	0.060	0.103	0.066	0.126	0.057	0.029	0.066
Non-attriters mean	0.55	1.45	6.67	0.20	2.47	2.02	2.29

Note. Standard errors in parentheses ="* p<0.05 ** p<0.01 *** p<0.001". Attrition rates are 24% in the treatment group and 30% in the control.

Intervention effects on children's learning outcomes

The treatment effects on student learning – literacy and numeracy – were tested using the following estimating equation:

$$y_{i,s} = \beta_0 + \beta_1 Treat_s + \mathrm{B}X_i' + \gamma_g + \epsilon_{i,s},$$

where $y_{i,s}$ is the outcome of interest (literacy or numeracy) for student *i* in school *s*, X_i is a vector of student-level characteristics, γ_g refers to strata fixed effects, and $\epsilon_{i,s}$ is the error term, clustered at a school level. The student-level controls are: baseline literacy and numeracy scores, age, gender, and the index for asset ownership.

Comparison of learning outcomes disaggregated by gender

Figures 17 to 20 compare girls and boys at baseline and endline for literacy (Figures 17 and 18) and numeracy (Figures 19 and 20). The y-axis reflects the proportion of children performing at the expected level.

Figure 17: Comparison of literacy performance by gender at baseline

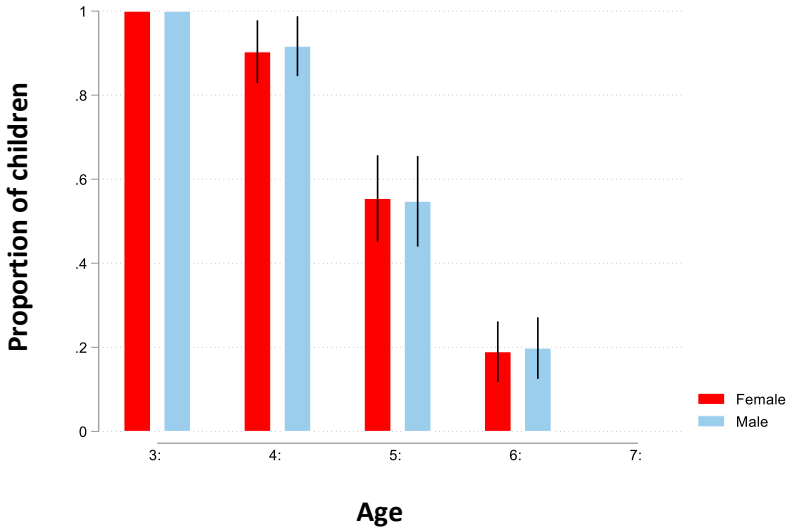


Figure 18: Comparison of literacy performance by gender at endline

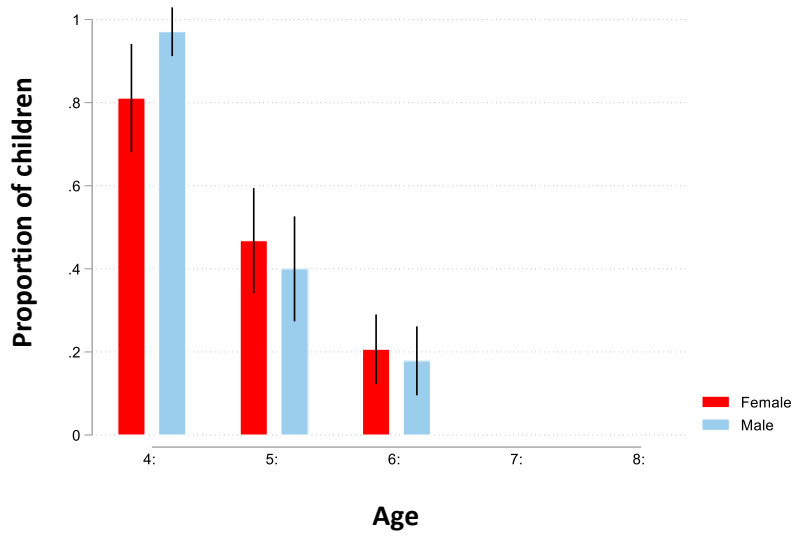


Figure 19: Comparison of numeracy performance by gender at baseline

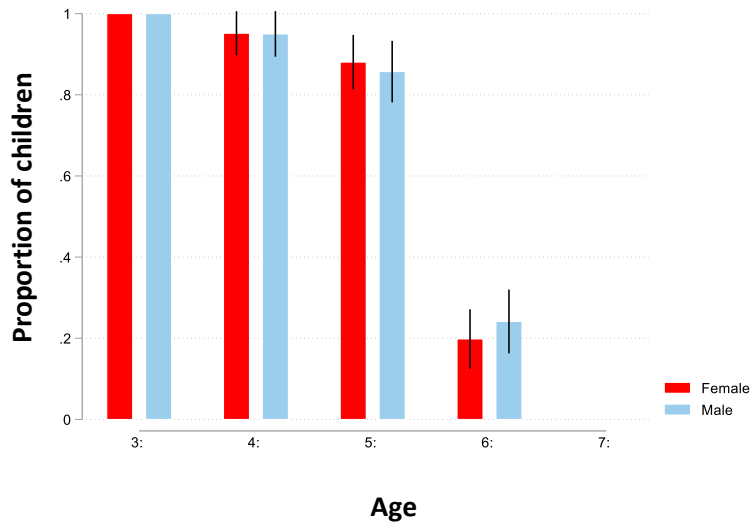
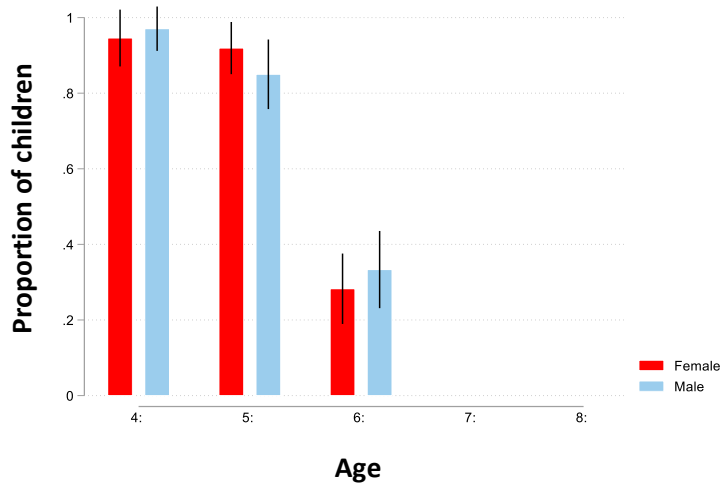


Figure 20: Comparison of numeracy performance by gender at endline



Changing in learning outcomes during COVID-19

Table 27 shows learning levels by age of the child at baseline and round of data collection. The column 'difference' captures the growth in learning between baseline and endline for each age. We focus on children aged five and six, who constitute the majority of the sample. A five-year-old, for example, performed 0.37 SD worse than a six-year-old in the numeracy test (0.02 - 0.39) at baseline. But these same five-year-olds improved their numeracy scores by 0.65 over the period of data collection.

Table 27: Student learning by round of data collection and student age

	Age at baseline				
	3	**4**	**5**	**6**	**7**
Panel A. Numeracy					
Baseline	-0.79	-0.55	0.02	0.39	1.54
Endline	-0.24	0.08	0.68	0.92	1.79
Difference	0.55	0.63	0.65	0.54	0.25
Panel B. Literacy					
Baseline	-0.82	-0.58	0.07	0.40	1.35
Endline	-0.43	-0.09	0.56	0.74	1.59
Difference	0.39	0.49	0.49	0.34	0.24
Obs	71	122	176	232	19

Finally, we test whether the learning trends were different for students who were present at or enrolled in school during endline data collection. There is no evidence that children who were still present at endline data collection learned more during the year. Their numeracy test score improved by 0.59 SD, compared to an improvement of 0.61 SD for those no longer at the school. In contrast, those who were still enrolled but who were absent at data collection improved their numeracy score by only 0.44 SD. But this difference is not statistically significant.

Classroom observations

The estimating strategy for classroom observations is:

$$y_{t,s} = \beta_0 + \beta_1 Treat_s + \mathrm{B}X'_t + \gamma_g + \epsilon_{t,s},$$

Where $y_{t,s}$ is the outcome for teacher t in school s. The teacher-level controls included are gender and the baseline measure for the same outcome.

Table 28 reflects the data on classroom observations at endline. These variables measure the number of times that each particular practice or materials was used over the course of the one-hour observation period: for example, teachers used on average 2.1 student-centred teaching practices (with a minimum of zero and a maximum of seven) over the course of the entire classroom observation period.

Table 28: Classroom observation variables at endline

Variable	N	Mean	SD	Min	Max
Student-centred teaching practices	49	2.10	2.14	0	7
Skills	49	2.10	2.95	0	10
Small pair	49	0.24	0.48	0	2
Engagement	49	3.11	0.43	2	4
Student-centred child activities	49	3.71	3.37	0	14
Student-centred total	49	8.16	6.48	0	21
Teacher materials	49	2.16	1.23	0	7
Student materials	49	1.91	1.23	0	6

Intervention – dosage

Figure 21 and Figure 22 report the distribution of mentorship visits between ECE and ALP teachers, respectively. On average, each ECE teacher received 2.78 visits, and each ALP teacher received 2.22 visits.

Figure 21: Distribution of mentorship visits received by teachers in ECE

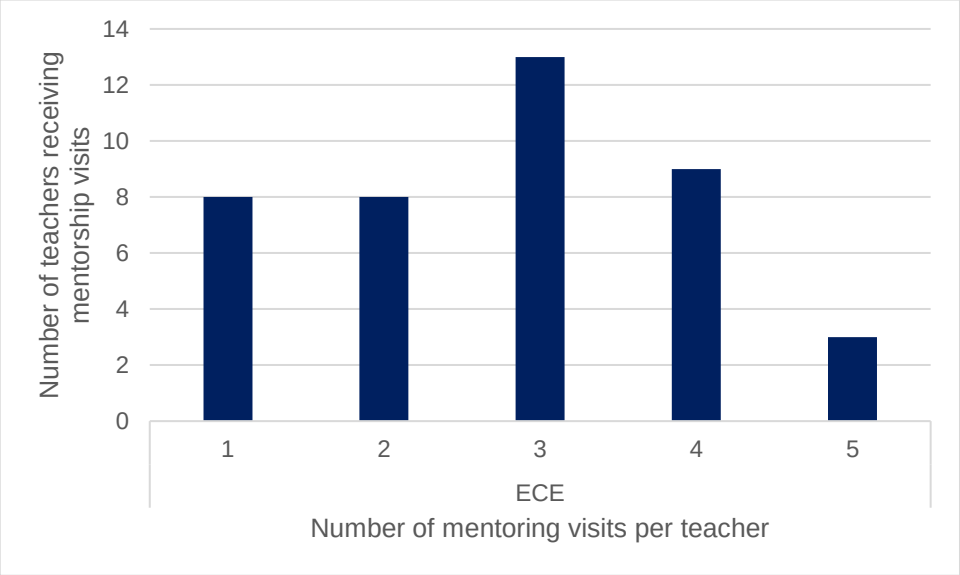
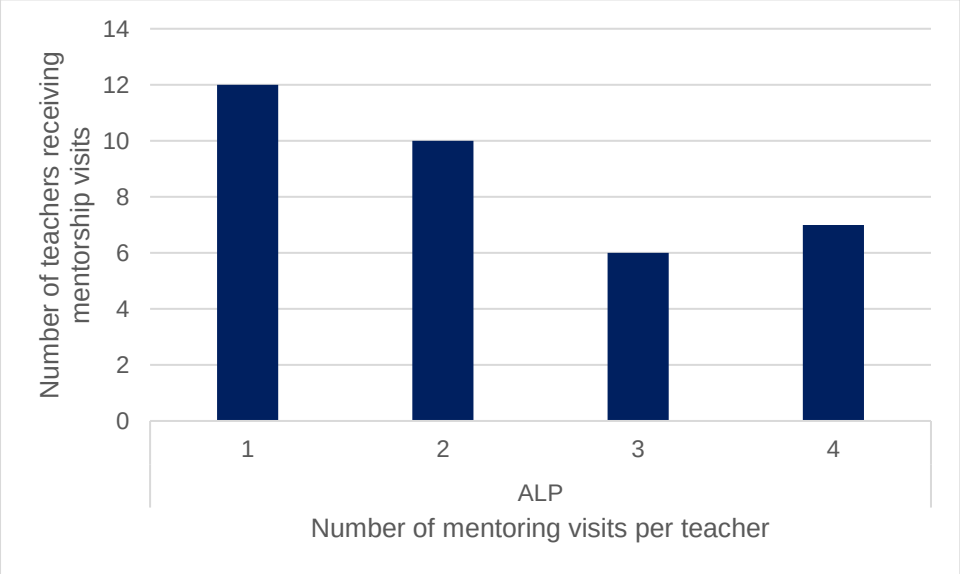


Figure 22: Distribution of mentorship visits received by teachers delivering ALP



A range of factors, including teacher absence, teachers' long-term leave due to health conditions, teacher strikes, and challenges retaining volunteer teachers, affected dosage. Given the rural location of many schools and the requirement for mentors to travel frequently, if a teacher was absent or a school closed early on a particular day, the mentoring visit was not rescheduled. The most significant impact on dosage, however, was the timing of mandated suspensions of class due to the COVID-19 pandemic.

Annex C Learning measurement specifics

Methodology

Clarifying constructs

The first step in undertaking a learning measurement exercise is to define exactly what is to be measured. Defining exactly what it is that we expect children to know and to be able to do is at the heart of the measurement of learning. The constructs measured through ELP Liberia were language and literacy, and mathematics knowledge and skills (including executive function).

Children in intervention and non-intervention schools were administered the same assessments: a language and literacy assessment and a mathematics knowledge and skills assessment. The content of the assessments is outlined in Table 29.

Table 29: Language and literacy and mathematics knowledge and skills content domains

Cognitive domain	Content domain
Language and literacy	Letter names
	Letter sounds
	Oral vocabulary
	Listening comprehension
	Writing
	Reading
Mathematics knowledge and skills	Numbers and quantities
	Measurement vocabulary
	Operations and relationships
	Spatial vocabulary
	Working memory
	Inhibitory control

Targeting

The research team avoided the flawed assumption that tested children can read or write already. Test items were designed to measure skill levels below, at, and above the skills assumed to have been reached given the age of the child. The major weakness in data measuring literacy and numeracy in low- and middle-income countries is that assessments measure skills at levels that are too high for most of the children taking the tests. This was observed in the case of Mali, where over 80% of students in Grade 2 could not read a single word in four national languages, while in the Nigerian state of Sokoto 81% of the students could not read full words (UNESCO, 2012). These floor effects are not problematic if the goal of the assessment is to establish national learning levels. For evaluation purposes it would not be possible to detect growth in literacy or numeracy unless those children previously out of range of the tests come into the ability level measured by the test. For those children who do not come into range, but whose literacy or numeracy levels do improve, the impact of the intervention will be underestimated.

Therefore, the ELP assessment tools were designed to ensure item difficulty matched students' ability. The level of the assessment and students' abilities were assessed during the piloting, and were adjusted based on the results.

Psychometric analysis

The use of Rasch modelling (IRT) increases the amount of analysis required as more sophisticated techniques are used, rather than adding up a total number of items correct in the test and converting the number into a percentage score. In this regard the evaluation team followed these steps: the first step was to test the psychometric properties of the items to ensure they were useful measures of what students know. The second step was to remove any items that did not perform well and that would bias the results if they were counted in the analysis. In a third step the team ranked the items according to difficulty. This was done using fit-for-purpose software. The software then also ranked students according to their ability and placed the students and the items onto the same metric. This is a probability model as students are placed on the scale according to the probability of a student answering the corresponding item correctly.

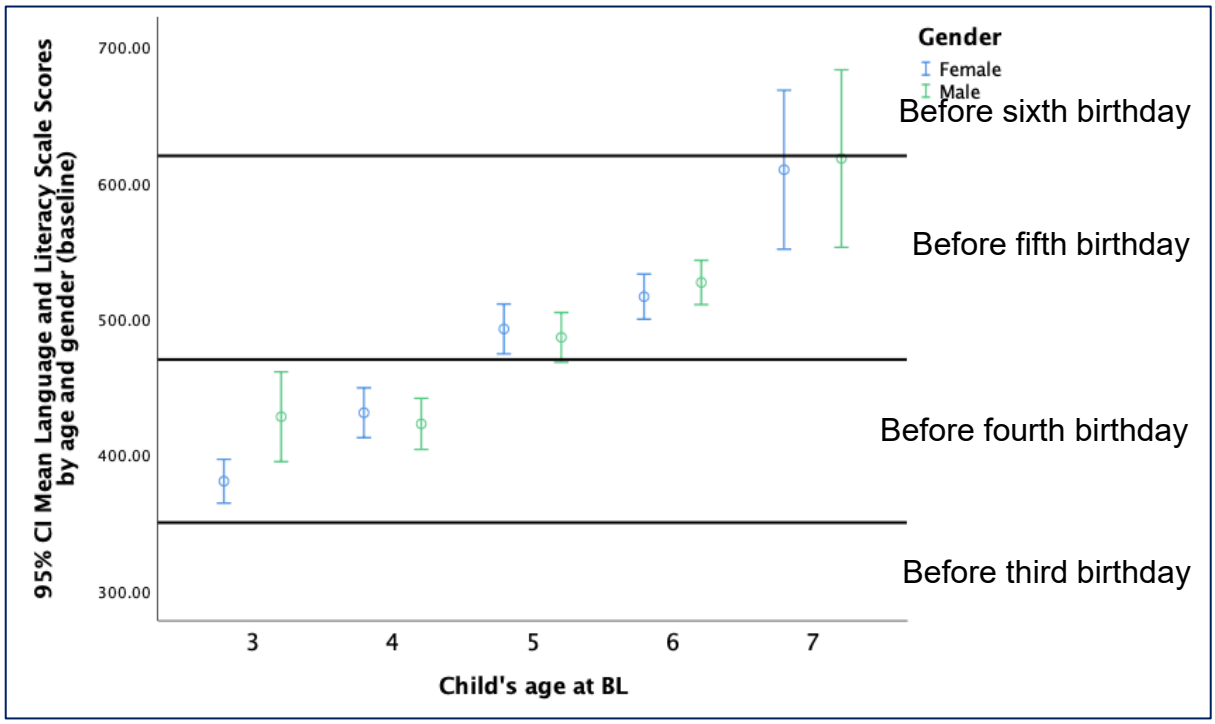
This was done using a mean of 500 and a standard deviation of 100.

The analysis produces two variables for each learning construct, per student. The first is a scale score. The scale score is a precise measure of where, along the achievement scale, the student sits. The scale score is useful for fine-grained analysis like multi-level modelling, regressions, and correlations. The second variable is which proficiency band the student falls within. This variable is useful for describing what learning students have achieved and how what children know and can do has changed, over the course of time.

Additional analysis

Figure 23 presents the average scale scores in language and literacy for children by age and gender at baseline. At baseline, the average three-year-old performs within the range expected of a child before their fourth birthday and, on average, boys have a higher level of proficiency at this age. This gender gap is not observed between average scores for four-year-old boys and girls; however, the average four-year-old is not much more proficient in language and literacy than the average three-year-old. The average four-year-old also falls within the range expected of a child before their fourth birthday. The average five- and six-year-old is more proficient in language and literacy than the average three- and four-year-old; however, the average five- and six-year-old performs within the range expected of a child before their fifth birthday. The average seven-year-old is proficient at approximately the level expected for a child turning five.

Figure 23: Mean language and literacy scale scores by age and gender (baseline)



As can be seen in Figure 24, the majority of learners know and can do what is expected before a child's fourth birthday. The peak of the distribution sits between the fourth and fifth birthday proficiency levels. By endline (Figure 25), a year later, the proportion of learners achieving what is expected before a child's sixth and fifth birthday increases. This is mainly explained by the shift in the distribution from learners performing in the 'before a child's fourth birthday' proficiency band to the 'before a child's fifth birthday' proficiency band.

Figure 24: Distribution of baseline language and literacy proficiency

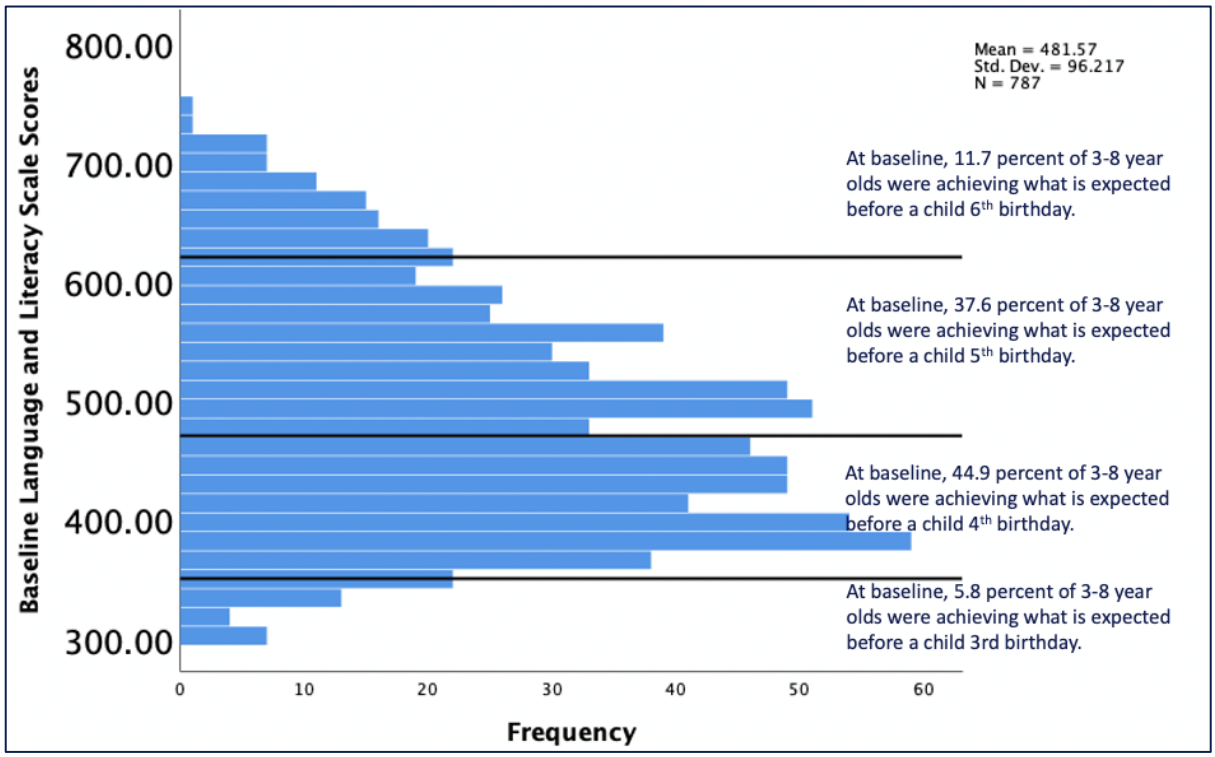
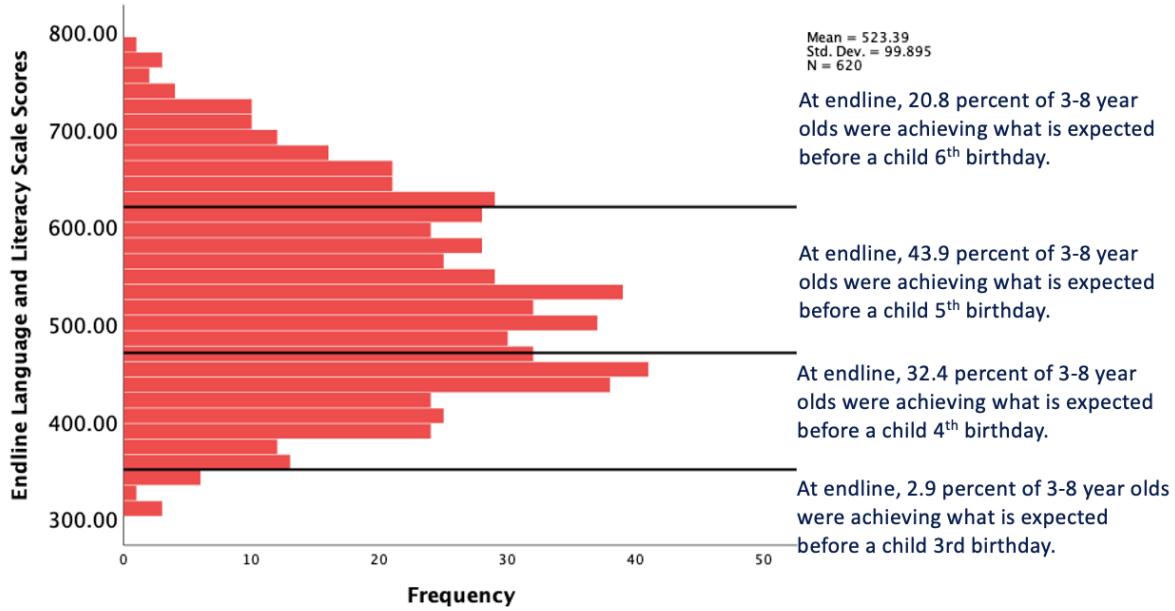


Figure 25: Distribution of endline language and literacy proficiency



As can be seen in Figure 25, the peak of the distribution shifts from near the cut-off point between three- and four-year-old expectations to the centre of the four-year-old expectation proficiency band. The range of proficiency does not shift a great deal between baseline and endline, indicating that some learners maintain very low proficiency levels as they age.

The final two figures below show the distribution of mathematics scale scores at baseline and endline.

Mathematics knowledge and skills

Figure 26: Distribution of baseline mathematics scale scores

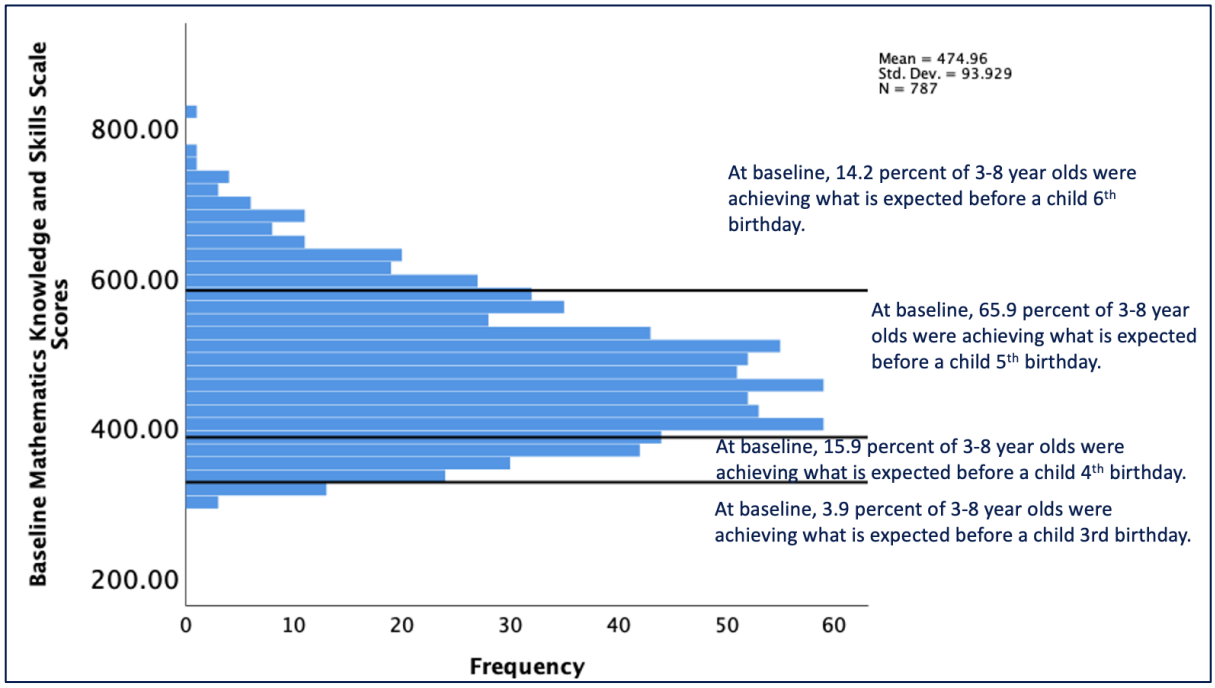


Figure 27: Distribution of endline mathematics scale scores