EDI Global

Independent Evaluation of the Improving early childhood care and education in rural Uganda program

On behalf of Lively Minds

Endline Report

June 2023

Contact Persons:

Mr Respichius Mitti, Regional Director; **d.mitti@edi-global.com**

Dr Johanna Choumert Nkolo, Director of Research; **j.choumert.nkolo@edi-global.com**

Ms Jyothsna Yasarapu, Senior Research Officer**, j.yasarapu@edi-global.com**

Mr Bence Kocsis, Data Innovation Lead, **b.kocsis@edi-global.com**

Ms Visheshika Baheti, Assistant Research Officer, **v.baheti@edi-global.com**

Uganda	**Tanzania**	**United Kingdom**	**Kenya**
EDI Global (Uganda) Limited	EDI Global (Tanzania) Limited	EDI Global (UK) Limited	P.O. Box 40612,
P.O. Box 3213,	PO Box 393, Bukoba, Kagera,	2-4 Packhorse Rd,	Nairobi,
Kampala,	Tanzania	Gerrards Cross,	Westlands District
Uganda	T: +255 (0)783 135 299	SL9 7QE, UK	T: +254 (0) 715 510175
T: +256 (0)414 233 204/114		T: +44 (0)1494 435112	

Contents

LIST OF TABLES AND FIGURES

ABBREVIATIONS AND ACRONYMS

CEO	Chief Executive Officer
CPR Index	Child-Parent Relationship Index
DiD	Difference-in-Differences
ECCE	Early childhood Care and Education
ECD	Early Childhood Development
FGD	Focus Group Discussion
HEQ Index	Home Environment Quality Index
IDELA	International Development and Early Learning Assessment
KII	Key Informant Interview
LM	Lively Minds
LMICs	Low- and Middle-Income Countries
LMTRP	Lively Minds Together Radio Programme
MWB Index	Mother's Wellbeing Index
PPI	Poverty-Probability Index
PS	Play Scheme
TOT	Training-of-Trainers
USAID	United States Agency for International Development
VHT	Village Health Team

ACKNOWLEDGEMENTS

EXECUTIVE SUMMARY

The following report presents the findings from the independent evaluation of the "Improving early childhood care and education in rural Uganda programme". This programme is testing the feasibility of the District of Mayuge adopting and implementing the Lively Minds programme through the VHT system, with technical support from Lively Minds. It aims to improve childcare, education, health, and wellbeing outcomes for caregivers and young children in resource-poor communities. The programme seeks to achieve this goal by providing additional Early Childhood Care and Education (ECCE) services to low-income communities through the existing networks of Village Health Teams (VHTs). In each community, mothers organise community play schemes, and they receive ongoing support from both the VHTs in their community, as well as in a form of regular parenting workshops throughout a two-year period.

The study was conducted in 24 villages of Mayuge district in Uganda, in partnership with the district administration. The primary aim of the evaluation was to estimate the impact of the programme on the treatment villages. The study specifically focused on evaluating changes in various aspects, including participating children's school readiness, their supportive home environment, wellbeing of their caregivers, child-parent relationships and hygienic behaviours.

The endline report provides details on the study's objectives, methodology, descriptive statistics of the sample of respondents, as well as the analysis and learnings from both the quantitative and qualitative components of the evaluation.

Quantitative data was collected from a total of 422 pairs of mothers and children in treatment and control communities across four sub-counties. Two rounds of data collection were conducted, with Baseline round taking place in May 2022, and the Endline round taking place in January 2023. Additionally, a series of focus group discussions (FGDs) and key informant interviews (KIIs) were conducted to gain a comprehensive understanding of the programme's impacts.

Overall, **the results of the quantitative analysis indicated that the intervention had only limited impact on children's school readiness**, as well as on the secondary indicators such as home environment quality, mothers' wellbeing or child-parent relationship. However, it's important to note **that the lack of significant findings is due to limiting factors in the quasi-experimental design** including imbalance on key outcomes at baseline between treatment and control groups; differential attrition (different types of children and mothers dropping out from the endline study between treatment and control groups); challenges with statistical matching; and the lack of robust point estimates. As a result, the quantitative survey failed to produce reliable assessments of the impact of the Mayuge LM programme.

Conversely, the qualitative study found that most programme participants, from various communities as well as across both implementers and beneficiaries, expressed that the programme had brought about meaningful change in their communities. They noted a greater sense of community as well as a transformative shift in adults' mindset around how to support young children's development. Respondents highlighted the increased involvement of fathers in their children's education and noted a reduction in cases of domestic violence, as reported by community and sub-county officials. Additionally, the qualitative accounts suggest that beneficiary children demonstrated improvements in behavioural, interpersonal, and cognitive skills as a direct result of participating in the programme. Children exhibited better communication abilities, conflict resolution skills, and emotional management capabilities. Mothers consistently emphasised how the programme had strengthened the bond between themselves and their children, leading to improved communication and stronger relationships within families. The programme also fostered unity and mutual support within communities, promoting child health and education.

On the implementer side, VHTs have reported personal growth, improved communication skills, and increased respect from local leaders. In each community, they felt that mothers have been sufficiently empowered to organize community play schemes, and ongoing training is given to maintain high standards. Respondents from the qualitative study felt that the programme has been instrumental in addressing critical gaps in child support and education in their communities, filling both and government functionaries with a sense of optimism that the programme has managed to garner the level of community buy-in, thereby ensuring its long-term sustainability.

1. INTRODUCTION

1.1. THE NEED FOR IMPROVED ECCE IN UGANDA

Nearly three in four pre-school aged children in low- and middle-income countries lack access to the most basic nurturing care.[1] Access to early childhood education has been slow and inequitable, both across and within countries. Worldwide, vulnerable children are disproportionately excluded from quality pre-primary education – even though it can have the greatest impact on them. Lack of affordable childcare is also a key barrier for parents, compounding multi-generational socioeconomic inequalities within countries. Children

---

[1] Bauld, A. (2022). A Look at the Lack of Basic Nurturing Care Worldwide for Preschoolers. Harvard Graduate School of Education. https://www.gse.harvard.edu/news/uk/22/05/look-lack-basic-nurturing-care-worldwide-preschoolers

who do not receive quality early childhood education are at greater risk of poor health, poverty, early marriage, and parenthood.[2]

While returns on investing in early childhood programs are extremely high, particularly for the poor and disadvantaged, in Uganda, only about 13% of children between the ages of three and five are enrolled in pre-primary schools.[3] With no state-funded opportunities, pre-primary education tends to be offered by fee-paying services, that often lack of appropriate learning materials and involve poor quality education for children.

A large proportion of children without these opportunities are located in hard-to-reach, resource-poor rural villages. Rural parents often hold misconceptions about the value of early childhood care and education (ECCE) due to their childhood experiences[4], as well as about their own ability to provide ECCE to their children themselves. Additionally, the government's focus on formal schooling structures can reinforce the false belief that quality ECCE can only be provided by professionals, and parents cannot take action themselves.

1.2. BRIEF OF THE MAYUGE LM PROGRAMME

In order to address the above challenges associated with poor early childhood care, Lively Minds (LM) has designed a holistic, and low-cost intervention that aims to expand and improve ECCE in rural areas. The core programme involves training mothers to run educational Play Schemes for their communities' preschoolers, and to provide nurturing care at home, using local resources. The intervention has been adapted to be delivered through government, so that it can be scaled and sustained while remaining low-cost. The intervention seeks to enhance the cognitive abilities, school readiness, health and physical development of preschool children through interactive games. The mothers participating in the PS programme are also offered the opportunity of focused Parenting Workshops, during which they can learn about useful new concepts around how to better nurture the health and development of their young children at home for little or no cost. The Play Schemes and Parenting Workshops are complemented by the Lively Minds Together Radio Programme (LMTRP), which adapts the content of the Parenting Workshop for radio and aims to educate audiences through engaging discussions, stories, and cost-free activities and games that provide the parents with ways to support their family's health, well-being, and their children's development.

---

[2] Global Partnership Strategy (GPS) for Early Childhood. (2022) United Nations Educational, Scientific and Cultural Organization, France.

[3] UNICEF, Save The Children and Ministry of Education and Sports (MoES). (2020) Planning for increased access to early childhood care and education: Thematic Studies

[4] Rural children lag in early childhood educational skills (2005). PRB. Retrieved March 13, 2023, from https://www.prb.org/resources/rural-children-lag-in-early-childhood-educational-skills/

The programme in Uganda aims to enhance the school readiness of 39,000 children, while also improving their health, social-emotional development, and cognitive skills. Additionally, the programme aims to improve parenting skills, self-esteem, and support networks for 9,750 mothers. The ultimate goal of the intervention is to strengthen the capacity of 60 local government officials and ensure that over 650 Village Health Team (VHT) officials are equipped to provide high-quality ECCE services independently, without ongoing support from LM.

The intervention has already been evaluated in similar settings in Ghana, where the programme has demonstrated promising results in terms of enhancing children's school readiness in areas such as cognition, executive function, early numeracy, and motor skills. Additionally, the programme has also shown positive impacts on children's health and nutritional status, socio-emotional skills, and parenting practices.[5] In 2019, LM received a grant from the United States Agency for International Development (USAID)'s Development Innovation Ventures programme to test and position the programme for scale-up in Uganda over a three-year period. As a part of this initiative, LM has partnered with Mayuge District Local Government in Uganda. Mayuge District, with technical support from LM, is implementing the programme in all twelve sub-counties. The intention is to develop a sustainable model of early childhood intervention that can be successfully incorporated into existing government systems. The programme is made up of three phases:

✓ In the "Setup Phase", the programme is established in all villages in a sub-county using a 'train the trainers' cascade model. LM staff train sub-county teams to deliver training to VHTs. VHTs cascade training to 30-40 mothers in their communities during eight two-hour workshops, where they learn how to run the Play Schemes. The Play Schemes are then launched. The first Parenting Workshops also take place during this phase (one term in each per sub-county).

✓ In the "Embed Phase", LM staff provide technical support to sub-county staff, who are responsible for monitoring and quality assuring the programme across the villages in the sub-county. LM staff hold monthly capacity-building workshops with and coach sub-county officials in the field, who in turn provide monthly top-up trainings for VHTs and supervise Play Schemes and Parenting Workshops. Each sub-county attends five terms in this phase.

---

[5] Amadu, S., Attanasio, O.P., Caeyers, B., Cattan, S., Sosa, L.C., Krutikova, S., Leighton, P., Masselus, L. and Yakubu, M., 2018. *Improving early childhood development in rural Ghana through scalable low-cost community-run play schemes: Baseline report* (No. R144). IFS Report.

✓ In the "Sustain Phase", the oversight and management of the programme will be taken over fully by the existing governance structures in the district, who will ensure the continued funding and success of the programme in the long term. Sub-counties enter this phase in their sixth term in the programme, and remain in it indefinitely. The first sub-counties will enter this phase in January 2024 and this report does not attempt to evaluate this phase .

1.3. PRESENTATION OF THE STUDY AND OUTLINE OF THE REPORT

The endline report is organised as follows: Section 2 provides a brief introduction and overview of the programme, Sections 3-6 describe the research design, key outcomes, and data collection procedures, Section 7 presents the study context while Sections 8-11 present the analytical strategy and results.

2. THE MAYUGE LM PROGRAMME

2.1. PROGRAMME COMPONENTS

The Mayuge LM programme adopts a training-of-trainers (ToT) model to deliver the programme. The programme is made up of the following components.

✓ **District onboarding and engagement:** LM has entered a partnership with Mayuge District Local Government in Uganda to support them to embed the programme. LM provides two years of technical and financial support to district and sub-county staff to help the district and each sub-county set-up and embed the programme with the expectation that the programme will be incorporated in the district's existing systems and activities in the sustain phase.

✓ **Training of VHTs:** The programme is administered at community level through local VHT members in each of the communities that the programme runs in. In each village, 3 VHTs are selected by the sub-county and each attend a series of 5 days of training sessions centred around topics on programme implementation and the value of ECCE. Any VHTs who do not complete this training (being late to training, poor attendance, etc.) are not appointed as Trainers. Once a month, VHTs attend a top-up workshop organised and facilitated by the sub-county team. At these workshops they discuss the Play Schemes and prepare to deliver the monthly parenting workshops.

✓ **Training of Mothers:** After the programme is announced in the community, VHTs recruit approximately 30-40 women (with or without young children). The VHTs

provide eight training workshops to the mothers (each around 2 hours) that cover various ECCE topics, including the importance of education, how to engage children through play, child-friendly teaching techniques, and hygiene. The training also provides a comprehensive overview of the programme and equips the mothers with the skills to facilitate PS for small groups of five to seven mothers going forward. Any participants who do not show commitment to the programme (being late to training, poor attendance, etc.) are dropped from the programme.

✓ **Play Schemes:** Once trained, the mothers run Play Schemes for all pre-school children in their community. The Play Scheme takes place on 3 afternoons each week for 2 hours. Mothers are divided into 3 groups, and each group is assigned their afternoon to teach. Up to 60 children are enrolled in each session. At the Play Scheme the children participate in structured group play – rotating between 6 different play stations, each facilitated by a different mother. The stations include activities such as matching, building, reading books, understanding shapes and counting, as well as outdoor play.

✓ **Parenting workshops:** Once a month, the VHTs provide group parenting workshops for the mothers on topics such as child health, learning, responsive caregiving and maternal wellbeing so that they can use local materials to follow better pro-learning practices at home within their community.

✓ **Radio programme**: During COVID, Lively Minds adapted their curriculum so it could be delivered through radio. Once normal activities resumed it was decided to maintain the radio component alongside the original programme. A small team of Sub-County officials are trained to record scripted radio episodes, on play or parenting topics, and these are broadcast in weekly episodes in local language on Busoga One, a local radio station with a reach of approximately 400,000 Community members are encouraged to listen to the radio show during regular programme activities.

✓ **On-going support:** Sub County teams visit each Play Scheme at least once a term to supervise each village and to provide improvement suggestions.

2.2. PROGRAMME THEORY OF CHANGE

The impacts on children's cognitive development are expected to manifest through two channels displayed in Figure 1.

Figure 1: Programme Theory of Change
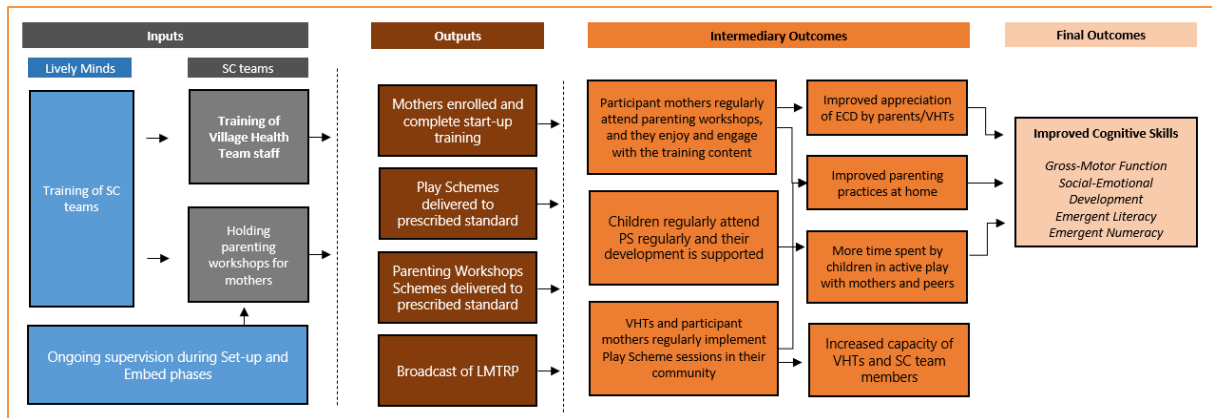
Source: EDI Global, based on discussions with LM

Firstly, the cognitive development of children is expected to be influenced by the pro-learning opportunities offered by the community-led play sessions, organised by mothers and VHTs. During these sessions, children are engaged through games and activities that are specifically designed to promote a wide range of competencies for children.

Secondly, the positive parenting practices promoted by the Mayuge LM programme in participating households are also expected to impact children's cognitive development. These parenting practices are expected to ensure that parents spend more quality time with children at home, improving not only children's school readiness, but mothers' own relationship with their child as well. Mothers are also trained on best practices around promoting healthy hygienic behaviours at home as well as how to provide more nutritious meals to their children.

3. EVALUATION DESIGN

3.1. RESEARCH QUESTIONS

In line with the Theory of Change described in Figure 1, the evaluation aimed to answer the following research questions to verify how the programme may have brought about change in the study communities and the district as a whole:

- ✓ To what degree have project outcomes (as outlined in the Theory of Change) been achieved?
- ✓ Were there any unexpected outcomes?

- ✓ Who has benefited (women, men, girls and boys) and in what ways?
- ✓ Are the changes likely to be sustainable in the long term?
- ✓ To what extent have frontline workers (Mothers, VHTs) committed to sustaining the project? What are the barriers and what are the facilitators?
- ✓ How effective has the ToT model (with VHTs as trainers) been in bringing about lasting change?
- ✓ What must be in place in order for the intervention to be replicable and scalable?

3.2. EVALUATION METHOD

In order to answer the research questions of the section above, LM has collaborated with EDI Global to conduct a mixed-methods impact evaluation of the programme. The study intervention was rolled out in the district in staggered phases in four of the twelve sub-counties of Mayuge district[6].

Accordingly, one major component of the evaluation is a quasi-experimental evaluation, using two rounds of quantitative data from May 2022 and January 2023 from a set of 24 villages, across two treatment and two control sub-counties. The quasi-experimental evaluation was primarily designed to measure changes in children's school readiness as well as mothers' wellbeing and hygienic practices over time by comparing data from the intervention group with a comparison group from the sub-counties that would only receive the intervention after 2023 January (hence functioning as a "control" group in the first phase).

Additionally, in order to gain more detailed understanding of the impact and experience of the programme, the evaluation included a qualitative component made up of a series of in-depth interviews and focus group discussions (FGDs) with a variety of stakeholders.

4. IMPLEMENTATION OF QUANTITATIVE SURVEY

4.1. TIMELINE

---

[6] From the 12 rural sub-counties of Mayuge District, an initial pilot was conducted in 2 sub-counties between 2016 and 2018, leaving 10 sub-counties to be part of the project funded by USAID. The district implementation commenced in October 2019 and a mixed-methods evaluation was originally scheduled to begin in early 2020 and to track 4 treatment and 4 control cohorts of sub-counties. Baseline data collection started in February 2020 in the first 2 treated sub-counties by another evaluation partner by a separate evaluation partner commissioned by LM. However, the COVID-19 pandemic and the lockdowns in response meant that it was not possible for the programme to run in the aforementioned 2 sub-counties. Once the programme resumed after COVID, 4 sub-counties (the original 2 from the pilot and the 2 sub-counties established in October 2019) could no longer form part of the evaluation led by EDI Global. It was therefore decided together with LM to restrict the evaluation to 2 treatment and 2 control sub-counties.

The timeline of the quantitative and qualitative evaluations is shown in Figure 2 below. Overall, the baseline data collection started on 10ᵗʰ May 2022 and ended on 17ᵗʰ May 2022. The endline data collection started on 30ᵗʰ January 2023 and ended on 6ᵗʰ February 2023. In the treatment communities, the recruitment of participant mothers took place between 9ᵗʰ and 12ᵗʰ of May 2022, while the recruitment of the participant mothers in control villages took place between 27ᵗʰ January and 6ᵗʰ of February 2023. In both cases during baseline and endline, schedule of the data collection team followed closely the order in which VHTs held recruitments in the study villages, ensuring that they would only approach a given cluster for survey after recruitment has been completed.

The intervention with the participant mothers in treatment communities started on 14ᵗʰ of May 2022. The data collection schedule ensured that all surveys were completed by this day in the treatment arm, with the remaining data collection only taking place in the control communities from 14ᵗʰ of May onwards.

Figure 2: Timeline of data collection and intervention schedule
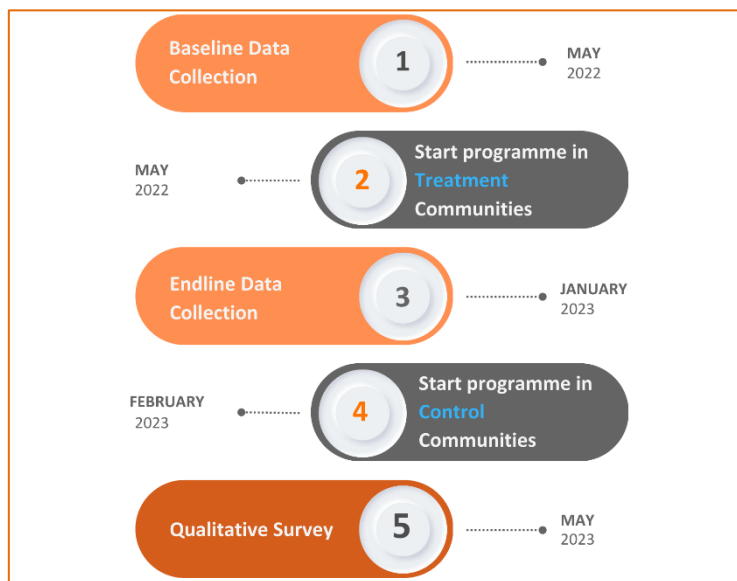
4.2. SAMPLING STRATEGY AND SAMPLE SIZE

As a part of the quantitative study, the two sub-counties that were scheduled to receive the intervention in May 2022 were designated as the treatment group, and 2 out of 2 sub-counties due to receive the intervention the following year were assigned to the control group[7].

---

[7] A 3ʳᵈ Sub-Country, Jaguzi sub-county, was excluded from the study due to its unusual position as a subcounty with many island-based communities.

Within the two treatment and two control sub-counties, EDI Global randomly selected a series of twelve-twelve villages in each study arm, from a list of all villages within the sub-counties (Table 7 in Appendix A).

To minimise possible selection bias in recruiting survey respondents, the sampling strategy used in the control villages was designed to mirror the recruitment procedures used in the treatment villages as closely as possible. Specifically, before the day of the survey, a community meeting was organised for eligible mothers in each control village with the assistance of the local VHTs. During these meetings, EDI Global staff presented the intervention and asked interested mothers to sign up for the programme, which would begin in one year's time, similar to recruitment in the treatment villages. This approach is expected to result in mothers surveyed in the control group being similar in their likelihood of participating in the programme.

During endline data collection, the mother-child pairs who participated in the baseline were re-contacted and re-interviewed. In cases where the mother-child pair could not be tracked or refused to continue participating in the study, they were replaced by another mother-child pair from the same village.

Figure 3: Geographical distribution of study sub-counties within Mayuge District
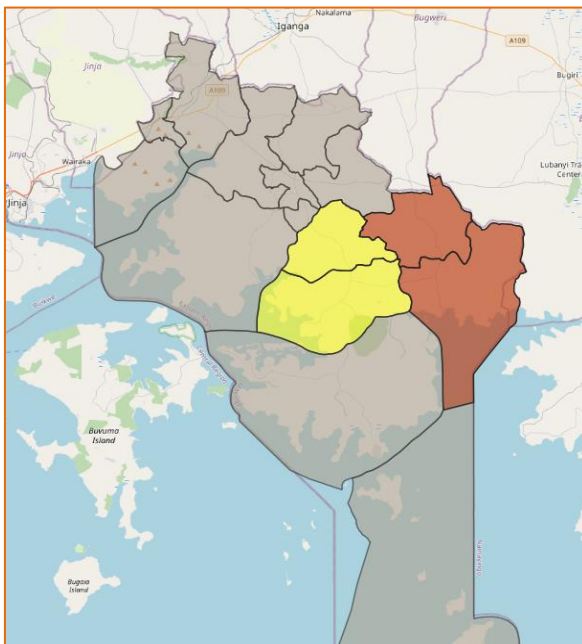
NOTE: Control sub-counties are marked in yellow, while Treatment sub-counties are marked in red

4.3. DATA COLLECTION PROTOCOLS

Overall, EDI Global collected data from a sample of 422 mother-child pairs,[8] from across a sample of 24 villages (twelve treatment and twelve control). The data was collected in two rounds, with the baseline data collection taking place in May of 2022 and the endline data collection taking place in February 2023, each over a two-week period.

The quantitative data collection was conducted through two sets of questionnaires – a mother questionnaire, as well as a child questionnaire that included a series of school-readiness measures using the International Development and Early Learning Assessment (IDELA) survey tool developed by Save The Children.[9] Selected respondents were called over the phone ahead of the field team's visit for data collection, and appointed a slot for interviewing at a central location within the community. Mothers would be present as the IDELA tests were administered to the children, as the mother interview was also conducted shortly afterwards.

The IDELA tool is designed to measure children's school readiness across four different domains: fine motor skills, socio-emotional development, emergent numeracy, and emergent literacy. Each sub-domain includes a series of questions and tasks for the child. The score for each item is calculated as the ratio of actual score over the total attainable score, taking a value between 0 (scoring none on all questionnaire items) and 1 (scoring highly on all questionnaire items) .

The mother questionnaire asked mother respondents to provide information on their household characteristics, hygienic practices at home, awareness of the LMTRP (at baseline), pro-learning activities done with the child, as well as on their subjective assessments of self-worth and relationship with their child. The mother tool has been developed by LM and refined for the current evaluation with EDI Global's support.

The same IDELA survey tool was used during both baseline and endline data collection rounds. For the endline round, the mother questionnaire has been refined through the following changes:

✓ Dropping questions around LMTRP listenership
   o These questions were added to the baseline tool in order to get early data on respondents' views of the programme – however, they were not of primary interest to the data collection.

✓ Adding additional socio-economic indicators

---

[8] Participant mothers would only be selected for the survey if their child (aged three to six) was also available for survey. Mothers did not necessarily need to be biological mothers of the target child, but they had to be caregivers and live in the same household.

[9] For more information, please see the resources provided by the IDELA Network at https://idela-network.org/

o In order to improve the statistical matching of respondents, new questions on the households' socioeconomic status and school enrolment of the target child were added to the endline tool.

✓ Adding measures of programme compliance (treatment geographies)
   o To get a better sense of impacts on the treated mothers and children, the endline tool would also ask mothers how frequently they have been attending the PS activities in the treatment villages.

For further information on the indicators marked questionnaire please see Appendix B

4.4. ATTRITION

Out of the 422 mother-child pairs interviewed at the baseline, EDI Global teams were able to interview 346 mother-child pairs at the endline as well (81% of the baseline sample). As shown in Table 1 below, the attrition rates are comparable between the treatment and control groups.

Reasons for attrition include relocation to a new location (~56% of the attrition cases), target child unavailability (~19%) and respondents being too busy to participate (~21%).

Table 1: Attrition by treatment status at endline

Reason for attrition	Overall sample*	Treatment	Control
Family moved to new location	42 (55%)	23 (54%)	20 (58%)
Target child not available for interview	14 (18%)	9 (21%)	5 (14%)
Mother not available for interview	16 (21%)	7 (17%)	9 (14%)
Mother not interested/refused	1 (1%)	1 (2%)	0
Other reasons	2 (2%)	2 (4%)	0
Total (% of endline sample)	**76 (18%)**	**42 (19.3%)**	**34 (16.5%)**
*Number of cases (% of total)			

4.5. ADDITIONAL SAMPLE

To deal with attrition in the sample and maintain statistical power, we replaced the mother-child pairs who could not be surveyed in endline with other eligible mother-child pairs in the same village. To do this, field enumerators used a randomised list of mothers who signed up for the intervention and interviewed them if they, together with their three to six year old child were living in the community. Altogether 76 replacement mothers were selected for the endline sample, ensuring that the total number of respondents across the two study arms remained the same (Table 11 in Appendix A).

The characteristics of the dropped out and replacement samples discussed further in section 8.3.

5. IMPLEMENTATION OF QUALITATIVE STUDY

5.1. OVERVIEW OF QUALITATIVE SURVEYS

The qualitative component of the evaluation was designed to complement the findings of the quantitative survey and took place between 8th and 26th of May 2023. The qualitative study was composed of the following interviews:

✓ **Focus Group Discussions (FGDs)**
The study included a series of ten FGDs – made up of seven sessions with beneficiary mothers and three sessions with VHTs, all of whom participated in the intervention. During these 120-minute sessions, the qualitative interviewers (one acting as notetaker and the other as facilitator) guided the discussion by ensuring that the participants broadly stuck to the topics of interest, remained respectful of each other, and that no one person dominated the ongoing conversations.

✓ **Key Informant Interviews (KIIs)**

During these one-on-one conversations, the interviewers created a private survey environment that allowed for in-depth discussions and further probing into the respondents' answers.

Altogether, EDI Global conducted KIIs with ten respondents. These were interviews of about 60 minutes in length, aiming to dive deeper into some of the key research questions around implementation experience and perceived impacts.

The respondent selection for the KIIs was purposive, with the aim of ensuring that the interviews took place with the stakeholders who were expected to be most knowledgeable about how the programme may have been adapted in the study communities, as well as across the key government functionaries. The list of ten key respondents to contact from the District and Sub-county teams, as well as LM staff are shown in Table 2 below.

Table 2: KII respondent profiles

Role	Areas that were covered in KII
Chief Executive Officer (CEO)	View of implementation and PS quality

Uganda LM Senior Management	Changes perceived in communities & sub-counties
District Lively Minds Coordinator (LMC); **Sub-County LMC (Control Sub-Counties);** **Sub-County chief (Control Sub-Counties)**	Changes perceived in communities & sub-counties; what may have caused improvements in the control communities[10] in the period, what is the value-added of the programme, implementation experience[11]
Sub-County LMC (Treatment Sub-Counties)	Changes perceived in communities & sub-counties team; what is the value- added of programme, implementation experience

6. OUTCOME MEASURES (QUANTITATIVE SURVEY)

Given the large number of research questions specified in section 3.1, an additional concern revolved around testing multiple hypotheses. The evaluation was looking at the effect of the intervention on cognitive development scores across four different IDELA sub-domains, the impact on a series of parenting practices, as well as agreement with multiple wellbeing and child-parent relationship statements. The large number of potential outcome variables means that we will potentially detect an effect even when there isn't one, purely by chance.[12] We account for this possibility by aggregating the individual variables into indices where necessary and define them as the main success indicators for the evaluation.

Accordingly, the following outcome variables will be highlighted as primary indicators of the programme's success in Mayuge district.

6.1. PRIMARY OUTCOMES

---

[10] Prior to conducting the qualitative study, the preliminary findings already hinted at the likelihood of detecting only a limited number of significant effects across various outcomes. Taking into account the underlying assumptions of the impact models, one additional objective of the qualitative work was to explore into the factors that may have influenced treatment and control villages in distinct ways. Understanding some of the potential confounding effects helped the study better contextualise the quantitative findings. For additional discussion, please see section 11.

[11] The reference period for each KII has been adjusted based on how long they have been involved with the implementation of the community-level intervention. In treatment sub-counties, this took place after May 2022, while in control sub-counties, the implementation only started in February 2023. The District LMC has been associated with the programme even longer, as some out-of-study sub-counties have already been running the programme since before the evaluation.

[12] Anderson, Michael L. "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association,* 103.484 (2008): 1481-1495.

The main interest for the evaluation is the impact of the programme on children's school readiness. This is because as the programme's Theory of Change (Figure 1) indicates, any improvements around children's learning environments primarily serve as a means to help children in resource-poor setting achieve better cognitive and developmental outcomes.

Accordingly, the main outcome variables will include the following variables:

✓ **Total IDELA score:** Composite score of all IDELA sub-domains; measuring children's overall school-readiness through a multi-dimensional index;

✓ **IDELA sub-domain scores:** Additionally, outcomes will be separately estimated for each IDELA sub-domain below (Table 3):

- Gross and Fine Motor Development

- Socio-Emotional Development

- Emergent Literacy and Language

- Emergent Numeracy.

The rationale for estimating outcomes separately for sub-domain scores is that the sessions of the PSs, as well as improved learning environments at home may be more suited to improve children's development in certain specific areas better than in others. Table 3 below outlines the four different sub-domains of child development that the survey tool was measuring, along with each corresponding competency.

Table 3: IDELA sub-domain scores[13]

Gross and Fine Motor Development	Emergent Literacy and Language	Emergent Numeracy	Social-emotional Development
Hopping on one foot	Print awareness	Measurement and comparison	Peer relationships
Copying a shape	Expressive vocabulary	Classification/ sorting	Emotional awareness & regulation
Drawing a human figure	Letter identification	Number identification	Empathy
Folding paper	Emergent writing	Shape identification	Self-awareness
	Phonemic awareness (First sound discrimination)	One-to-one correspondence	Conflict resolution

---

[13] Guide to create summary IDELA scores are available by Save The Children at**:** https://idela-network.org/. Note that two optional domains of "executive function" and "approaches to learning" items are not included in the total IDELA score for this study.

	Listening comprehension	Simple operations (addition & subtraction)	
		Simple problem solving (puzzle)	

6.2. SECONDARY OUTCOMES

The most important secondary outcomes include the measures of positive learning environment at home, which are envisioned to contribute to children's school-readiness as suggested by the Theory of Change in Figure 1. More specifically, the indicators include:

- ✓ **Mother's wellbeing index (MWB Index):** Composite score of the subjective assessment of all statements relating to mothers' self-worth. Through more active engagement with their children, mothers' subjective wellbeing is also envisioned to improve through the positive psychological impacts of stronger relationships and active play with the child.

- ✓ **Child-Parent Relationship index (CPR Index):** Composite score of the subjective assessment of all statements relating to mothers' relationship with their child; Similarly, as above, focused ECD activities are expected to building stronger bonds between mothers and children.

- ✓ **Handwashing index:** Composite score of all the positive hygienic behaviours related to handwashing that the respondent mother reports following at home.

- ✓ **Home Environment Quality index (HEQ Index):** Composite score of all the pro-learning activities that children are engaged at home by an adult or older sibling (including playing, drawing, telling stories, reading, singing or outing together). Through the parenting workshops, mothers are expected to value the importance of active engagement with their children more and practice more pro-learning activities with the child at home. Additionally, they are also expected to share their learnings with other household members (father, grandmother, etc.) who may also contribute to these activities at home.

6.3. CONTROL VARIABLES

In order to disentangle the effects of the programme on the primary and secondary outcomes, the following statistical controls have been included in the estimation models.

- ✓ **Child age:** Children's age directly influences many of the school-readiness measures, including their Gross and Fine-motor skills, Emergent Numeracy,

Language as well as Socio-Emotional skills. Additionally, as children grow older their parents may interact with them differently at home depending on how older children's relationship evolves with them.

- ✓ **Child gender**: Male and female children may be affected differentially by the LM programme in terms of their cognitive development. Additionally, mothers may interact differently at home with their daughters compared to their sons at home, in the context of the study villages.[14]

- ✓ **Mother age:** Including mother's age may influence how receptive participant mothers are to the intervention. Older mothers may have more experience with children and implement suggested activities more comfortably, while on the other hand they may also have more rigidly held beliefs on ideal parenting practices.[15]

- ✓ **Poverty Probability Index (PPI):** Poverty is a well-established risk factor for negative child outcomes, including malnutrition, poor health, and limited educational opportunities.[16] By controlling for poverty, we can better isolate the effects of the programme on children's school readiness and their learning environment.

- ✓ **Mother's education:** More educated mothers may have an advantage in conducting certain pro-learning activities with their children such as reading stories. Additionally, mother's education level is often a strong predictor of a child's cognitive and socio-economic outcomes.[17] By including maternal education as a control variable in a regression model, we can account for the confounding effects of maternal education on children's development.

A more detailed description on the construction of the above variables is outlined in Appendix B.

---

[14] Croke, K., & Atun, R. (2019). The long run impact of early childhood deworming on numeracy and literacy: Evidence from Uganda. *PLOS Neglected Tropical Diseases*, *13*(1), e0007085.

[15] Abimpaye, M., Dusabe, C., Nzabonimpa, J. P., Ashford, R., & Pisani, L. (2020). Improving parenting practices and development for young children in Rwanda: Results from a randomized control trial. *International Journal of Behavioral Development*, *44*(3), 205-215.

[16] Tusting LS, Willey BA, Lucas H, et al. Socioeconomic development as an intervention against malaria: a systematic review and meta-analysis. Lancet. 2013;382(9896):963-972. doi:10.1016/S0140-6736(13)60851-X.

[17] Mahoney, G., Boyce, G., Fewell, R. R., Spiker, D., & Wheeden, C. A. (1998). The relationship of parent-child interaction to the effectiveness of early intervention services for at-risk children and children with disabilities. *Topics in early childhood special education*, *18*(1), 5-17.

7. STUDY CONTEXT

7.1. SOCIODEMOGRAPHIC AND ECONOMIC CHARACTERISTICS OF THE SAMPLE

In this section, we describe and compare the household characteristics of our sample with the national and regional averages from the Uganda Demographic and Health Survey (2016) and other studies of similar context.

At baseline, the mother questionnaire collected a handful of socio-demographic indicators on the household, these are presented in Table 8 in Appendix C. The indicators collected at baseline primarily captured the demographic characteristics of the child or the mother herself. When it comes to mothers' educational attainment less than half of the mothers in both study groups had finished primary school (35%) with less than 5% holding an O-level certificate or above by May 2022. This female educational level in the baseline cohort slightly lower than the average in rural Uganda[18], but remained balanced across treatment and control samples, as well as in the replacement sample at endline. Additionally, about one-fifth (20.6%) of the respondent mothers were heads of their own households, which is a notably lower rate than the average, in rural Uganda (30.6%[19]). The average age for respondent mothers was 33.8 years, with more than 81% of the respondents being younger than 40. The median number of household members in the sample was seven and the median number of children the target mother had was three.

As a part of the baseline survey, no additional information was collected on the household's economic status (wealth or income). However, for the endline study additional variables on the socio-economic standing of the household were collected in order to construct a poverty index using ten easily collectable indicators from Uganda's 2012/13 National Household Survey.[20] In the study communities, it was found that, 51% of the households used burnt stabilised bricks and 32% used unburnt bricks with mud for the construction of the wall of the dwelling. For the construction of the roof of their dwelling, 89% used iron sheets and 11% used thatch/grass. Most households used firewood for cooking (98%) and the remaining used charcoal. 72% of the households had earth/rammed earth as their flooring, and the remaining had cement screed (23%), concrete (3%) or brick/stone (1%). On average, each household had 1.7 mobile phones.

Altogether, over 92% of the sampled mothers had two or more children at the time of the baseline data collection. As for the child respondents themselves, most sampled children fall towards the higher end of the target age bracket of three to six years for LM's intervention, with an average age of 4.93 across the sample. Over 37% of child-

---

[18] Uganda Bureau of Statistics (UBOS) and ICF. 2018. Uganda Demographic and Health Survey 2016. Kampala, Uganda and Rockville, Maryland, USA: UBOS and ICF.
[19] *Ibid*.
[20] Schreiner, M. (2011). Simple Poverty Scorecard Poverty-Assessment Tool Uganda. *Retrieved, January*, *15*, 2019.

respondents are six years old, which means that by the time of the endline survey, they could be less suited to measurement of cognitive skills using the IDELA tool, which has a target age group of three to six.

7.2. CHILD DEVELOPMENT

When it comes to measures of children's cognitive development, the quantitative survey only collected information on children's school-readiness, as measured by the IDELA tool. While there are currently no nationally representative surveys available using this tool, the scores of the sampled children are comparable to other studies of low-income settings as reported by Save the Children (2018)[21] across the world. As shown in Table 4 below, the domain of "Emergent Literacy and Language" seems to be lower in our study than for children in other low-income settings across the world.

Table 4: Comparison of child development scores

	Save The Children (2018) (Average for six-year-olds)	Study Sample (Mean child age: 5.7 years)
Total IDELA score	60%	53%
Sub-Domain score: Emergent Numeracy	59%	55%
Sub-Domain score: Emergent Literacy and Language	51%	34%
Sub-Domain score: Gross and Fine motor skills	73%	74%
Sub-Domain score: Social and Emotional Learning	52%	50%

7.3. MEASURES OF MOTHER WELLBEING AND PRO-LEARNING PRACTICES AT HOME

During the caregiver interviews, each mother was also asked to describe the types of activities that have been used at home to engage their three to six-year-old child and which family member has been involved in these activities. According to the mothers' responses, singing and outdoor activities are the most common ways that children are engaged at home, with 74% and 75% of children across the sample having participated in these activities at least once in the three days prior to the survey (Table 5). However, reading and drawing with adults in the household are less common activities, with only around 50% of children in the sample having participated in these activities with an adult in their home during the preceding three days.

Table 5: Overview of positive child engagement activities at home

---

[21]Save the Children. (2018). Beyond access: exploring equity in early childhood learning and development. Available at https://idela-network.org/wp-content/uploads/2018/06/IDELA-Report-2018-4WEB.pdf

Activity with the child (any adult in the household)	Frequency of engaging the child in the last three days	Percentage of children	Standard error	Sample size
Reading a book with the child	Not at all	49.53	2.43	422
	1-2 times	31.52	2.26	
	3-4 times	15.4	1.76	
	5 or more times	3.55	0.9	
Telling a story to the child	Not at all	44.79	2.42	422
	1-2 times	36.97	2.35	
	3-4 times	16.11	1.79	
	5 or more times	2.13	0.7	
Singing with the child	Not at all	26.3	2.14	422
	1-2 times	38.63	2.37	
	3-4 times	29.62	2.22	
	5 or more times	5.45	1.11	
Going to some outside place with the child (e.g., market, shop, for a walk)	Not at all	25.83	2.13	422
	1-2 times	41	2.39	
	3-4 times	26.3	2.14	
	5 or more times	6.87	1.23	
Playing with the child	Not at all	45.5	2.42	422
	1-2 times	32.46	2.28	
	3-4 times	18.72	1.9	
	5 or more times	3.32	0.87	
Drawing together with the child	Not at all	53.79	2.43	422
	1-2 times	30.33	2.24	
	3-4 times	13.74	1.68	
	5 or more times	2.13	0.7	

Additionally, the intervention also aimed to improve mothers' and parents' general well-being. The programme envisioned that through its intervention, mothers' mental well-being and self-esteem may be strengthened. Table 6 below summarises the responses of the mother respondents to each well-being-related statement from the Rosenberg (1965).[22]

Table 6: Overview of mother wellbeing scores

---

[22] Rosenberg, M. (1965). Rosenberg Self-Esteem Scale (RSE). Acceptance and Commitment Therapy. Measures Package, 61, 5

Wellbeing statement	Value	Percentage of respondents	Standard error	Sample size[23]
"On the whole, I am satisfied with myself"	Strongly agree	33.17	3.34	199
	Agree	54.27	3.53	
	Disagree	11.56	2.27	
	Strongly disagree	1.01	0.71	
"I feel I have a number of good qualities"	Strongly agree	24.62	3.05	199
	Agree	60.80	3.46	
	Disagree	14.57	2.50	
	Strongly Disagree	0	0	
"I am able to do things as well as most other people"	Strongly agree	29.60	3.06	223
	Agree	61.43	3.26	
	Disagree	8.52	1.87	
	Strongly disagree	0.45	0.45	
"I feel that I'm a person of worth, at least on an equal plane with others"	Strongly agree	50.75	3.54	199
	Agree	46.23	3.53	
	Disagree	3.02	1.21	
	Strongly Disagree	0	0	
"I take a positive attitude towards myself"	Strongly agree	39.91	3.28	223
	Agree	54.26	3.34	
	Disagree	5.38	1.51	
	Strongly disagree	0.45	0.45	
"At times I think I am no good at all"	Strongly agree	6.73	1.68	223
	Agree	53.81	3.34	
	Disagree	30.49	3.08	
	Strongly disagree	8.97	1.91	
"I feel I do not have much to be proud of"	Strongly agree	4.02	1.39	199
	Agree	37.69	3.44	
	Disagree	40.70	3.48	
	Strongly disagree	17.59	2.70	
"I certainly feel useless at times"	Strongly agree	2.69	1.08	223
	Agree	27.80	3.00	
	Disagree	26.01	2.94	
	Strongly disagree	43.50	3.32	
"I wish I could have more respect for myself"	Strongly agree	16.14	2.46	223
	Agree	70.40	3.06	
	Disagree	13.00	2.25	
	Strongly disagree	0.45	0.45	
	Strongly agree	2.51	1.11	199

[23] Similarly to the CPR section, in order to keep to the agreed length of 30-40 minutes for the mothers tool, EDI Global used random assignment to assign only half of the statements to each respondent. There is therefore a lower sample size recorded across the study respondents, however, as this questionnaire logic was implemented uniformly across treatment and control villages, the sample remains balanced and the takeaways from these questions should remain unbiased.

"All in all, I am inclined to feel that I am a failure"	Agree	16.58	2.64	
	Disagree	25.63	3.09	
	Strongly disagree	55.28	3.52	

Note: Sample sizes vary across indicators due to missing responses, "Don't know/Refused" answers or skip patterns coded into the questionnaire logics

Overall, mothers scored moderate to high on the MWB Index. When adjusted for our scale, the MWB index is supposed to correspond to the following broad categorisation:[24]

- Below 2.5 – Low-self, esteem
- Between 2.5 and 3.49 – Normal range
- Above 3.49 – better than normal range

With an average MWB score of 2.9 across the sample at baseline, most mothers fall in the normal range for self-esteem.

8. EMPIRICAL STRATEGY

8.1. MODEL

The intervention was rolled out at a sub-county level, with the two treatment and two control areas having been purposively selected from the twelve sub-counties of Mayuge district. In order to get more accurate estimates on the programme's impact from a quasi-experimental set-up, we estimate a difference-in-differences (DiD) model with kernel-based propensity-score matching. This method would be effective in controlling for time-varying confounding variables that may bias the estimated effect of the intervention on mother and child outcomes. By comparing the change in outcomes over time between the treatment and control groups, we can better isolate the effect of the treatment from other factors that might affect the outcome.[25]

The estimation model takes the following form:

$$\gamma_{it} = \beta_0 + \beta_1 D_{it} + \beta_2 T_{it} + \beta_3 T * D_{it} + \varepsilon_{it}$$

---

[24] Martín-Albo, J., Núñez, J. L., Navarro, J. G., & Grijalvo, F. (2007). The Rosenberg Self-Esteem Scale: translation and validation in university students. *The Spanish journal of psychology*, *10*(2), 458-467.

[25] Abadie, A., & Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *econometrica*, *74*(1), 235-267.

Where γ_{it} is the outcome variable[26] (see section 6) for an individual at time t^{27}. D is a binary indicator for treatment variable, where D = 1 if an individual receives the LM intervention, and D = 0 if individual i does not receive the programme. Note that since respondents with poor programme attendance have been dropped out of the programme, treatment refers to mothers in treatment villages whose children have been consistently attending the LM programme for at least seven months.

X_i denotes a set of control variables[28] (co-variates) that that may influence the outcome variable and/or possibly the treatment assignment. The term β_0 denotes the intercept, β_1 is the average treatment effect, β_2 corresponds to the average time effect and β_3 refers to the average treatment effect over time.

To estimate the impact of the Mayuge LM programme, we want to compare two groups: the group of mothers and children who received the treatment, and the group of mothers and children who other did not. But there might be differences between these groups that make it difficult to determine if the treatment caused the observed changes, or some other factors. To overcome this challenge, we use a special method called "kernel-based propensity score matching". This method helps us make the groups more comparable by finding individuals in the control group who seem similar to those in the treated group, on observable characteristics.

For instance, when we're studying the effects of the programme on children's emergent literacy and language, we collect information about factors that can influence this skill, such as each child's age, gender or mother's education level as discussed in section 6. Based on this information, we calculate a score called the "kernel-based propensity score" for each mother-child pair.

---

[26] As outlined in section 6, the "outcome" of the model refers to changes of the following indicators. The impacts were estimated for all of the indicators, and the results are displayed in Appendix D.
- Total IDELA Score
- IDELA Sub-domain Score: Socio-Emotional Learning
- IDELA Sub-domain Score: Emergent Numeracy
- IDELA Sub-domain Score: Gross and Fine motor development
- IDELA Sub-domain Score: Emergent Literacy and Language
- HEQ index
- Handwashing index
- CPR index
- MWB index

[27] Time "t" in this context refers to two time periods: "pre-treatment", which refers to the time of baseline data collection in 2022 May; and "post-treatment", which refers to the time of endline data collection in 2023 January.

[28] The list of control variables refers to the indicators outlined in section 6, namely 1) Child age; 2) Child gender; 3) Mother age, 4) Poverty Probability Index (PPI); and 5) Mother's education.

The propensity score represents the likelihood of a child receiving the Mayuge LM programme programme, given their age, gender, PPI score, mothers' age and mother's education. Next, using the propensity scores, we search for mother-child pairs in the control group (those who did not receive the programme) who have similar scores to those in the treatment group (those who received the programme). By creating these matched pairs, we can compare the outcomes of the treatment group with those of the control group. Since the pairs have similar propensity scores, we can be more confident that any differences we observe in their literacy skills are due to the treatment itself, rather than other factors like age, gender, mother's education or household wealth (as measured by PPI).

In technical terms, the propensity score for individual *i* is defined as the conditional probability of participating in the programme given the covariates X_i:[29]

$$\rho_i = P(D_i = 1 \mid X_i)$$

We can then use kernel-based matching to match each treated individual with a set of untreated individuals who have similar propensity scores. This creates a matched control group that is comparable to the treated group in terms of observable characteristics.

As further checks for robustness, programme impacts were re-estimated by adding mother fixed effect, in order to control for unobserved differences that remains constant over time or across groups. By including mother-fixed effects, the model accounts for factors that are specific to mother-child pair, but do not change over time. As an example, such factor could be mother's inherent aptitude at certain tasks, which may influence how effectively they implement the positive parenting practices at home.

Impact estimates were further refined by clustering standard errors at the sub-county-level. Clustering standard errors is a technique used when the observations within a group or cluster are likely to be correlated – in this case, since the programme was rolled out at the sub-county level, any significant differences between treatment and sub-counties (such as different rates of infrastructure development, average quality of schools ..etc.) would make it difficult to isolate the impact of the programme activities alone.

This model makes some of the following assumptions in creating the impact estimates:

---

[29] Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. The Review of Economics and Statistics, 84(1), 151-161

1. The treatment assignment was not affected by any unobserved confounding variables that may affect the outcome.

2. The propensity score distributions for the treatment and control groups must overlap, which means that there are individuals in both groups with similar probabilities of receiving the treatment.

3. The parallel trends assumption requires that the treatment and control groups would have followed similar trends in the absence of the treatment. This means that respondents in the treatment group who participated in the LM intervention would have had similar outcomes in the absence of the programme to the outcomes of control participants.

4. The treatment effect must be stable over time, which means that the effect of the LM intervention does not meaningfully change over time. This assumption is important to ensure that the treatment effect estimated using DiD remains valid over the entire period of the analysis.

8.2. Sample Balance tests

Sample imbalance refers to dissimilarity of observations between the treatment and control groups, either at baseline or over time, which can lead to a lack of comparability between the groups, making it difficult to establish a clear counterfactual.

Table 10 in Appendix C outlines the sample balance tests between treatment and control groups at baseline. Overall, at baseline the sample of respondent mother-child pairs was not perfectly balanced between the two study groups, as some of the indicators of interest seemed to be at higher levels at baseline in the treatment units. Notably in the sub-domain of emergent literacy and emergent numeracy, respondent children in the treatment areas did seem to perform significantly higher than their counterparts in the control units. Additionally, respondent mothers in treatment areas scored higher on handwashing index[30] (likelihood of washing hands before and after certain types of activities) and respondents in the treatment sample were also found to be considerably more likely to have a dedicated place for handwashing at home. Mothers in treatment villages also appeared more likely to spend quality time together with their child. Moreover, at baseline treatment villages saw significantly higher levels of LMTRP listenership than control villages with about ten percentage points less likely to have heard of the programme there. Mothers in treatment

---

[30] The handwashing index captures mother respondents' reported practices around whether or not they wash their hands after/before certain types of activities. Please see appendix B for more information.

villages were also slightly older at baseline, with an average age of 34 compared to 32 in control communities.

Sample balance at baseline was achieved when it comes to the age of the target child, mothers' education score, total IDELA score, child-parent relationship and mothers' wellbeing index.

Additional demographic indicators around asset ownership and literacy were collected at endline to verify comparability of the respondents' household wealth. While these indicators were only collected at the endline round of data collection, the LM intervention is unlikely to have affected these factors as the programme does not involve any form of monetary or in-kind support or compensation to the participants. As shown in Table 10 in Appendix C, the treatment and control samples slightly differed on composite poverty scores, with respondents in the treatment areas marked as wealthier on average compared to the respondents in control areas. Since the survey did not collect PPI measures at baseline, we cannot verify to what extent this difference may have been driven by differential attrition[31].

8.3. COMPARISON OF DROPOUT AND REPLACEMENT SAMPLES

Attrition and differential replacement protocols may undermine reaching valid inferences by introducing systematic differences between the retained and dropout samples. When taking a closer look at the attrition rates (Table 11 in Appendix C), at baseline those mother-child pairs that dropped out were significantly more likely to have scored higher on the HEQ index, but only in the treatment villages. Beyond this however, mothers and their children that dropped out from the study, were relatively similar to those that remained in terms of their school-readiness scores, secondary outcomes such as mothers' wellbeing, child-parent relationship or handwashing. Moreover, target children that dropped out did not seem to differ in their age from those that remained part of the study. These similarities were consistent across both treatment and control geographies.

While the sample that dropped out had relatively similar characteristics at baseline, the replacement respondents seem to differ more significantly from the rest of the sample. At endline, replacement respondents were scoring lower on all IDELA sub-domain and total scores when compared to children in the original sample. This can partly be explained through other differences as well, for instance replacement children were younger than the rest of the sample, on average by one full year.

---

[31] Differential attrition refers to a situations where respondents dropping out or leaving a study are not the same across the groups that being compared (in the case of the current impact evaluation, between treatment and control groups).

While these differences are consistent across both treatment and control villages, the magnitude of the difference may suggest that the original sample of mothers would not be representative of the Mayuge LM programme participants in the community, which can be a problem for the validity of the differences-in-differences model. This is because if the attrition is not random, we cannot rule out if there were different confounding effects between treatment and control villages that were driving the results that we see. For instance, it may be that in treatment villages those children are likely to drop out, whose parents can afford their children to be sent to available boarding schools in the surrounding area and hence were not around during the endline data collection. On the other hand, if in control villages there were fewer available school opportunities, then children whose parents can afford boarding school may still be more likely to remain for the endline evaluation, as their children would still live in the community. In this case, the attrition would not be random, but rather related to the programme and the outcome we're studying.

As a result, the differences we observe in the outcomes between the control and treatment groups after the attrition may not solely be due to the program's effect. Instead, it could be influenced by the differences in the characteristics of the students who dropped out.

9. RESULTS OF THE QUANTITATIVE ANALYSIS

9.1. IMPACTS ON PRIMARY OUTCOME MEASURES

The results of the DiD estimator are shown in Tables 12-19 in the Appendix D. Overall, **most of the estimated results do not seem robust to different model specifications.** This means that when we tried different ways of analysing the data using different model specifications (section 8), the results tended to change, which suggest there is a great room for bias in how the indicators are selected for the estimation. This level of bias indicates that we cannot make robust conclusions about the impact of the programme when it comes to children's school readiness.

For instance, the model estimates do not pick up significant effects on children's school readiness as measured by the total IDELA scores. When taking a closer look at the sub-domain scores, there are no significant effects detected for either the sub-domain of emergent literacy and language or gross- and fine motor skills. When subcounty-level effects are controlled for[32], children in the treatment villages on average are estimated to perform better on the sub-domain by 0.1 standard deviations, an effect size similar to the one

---

[32] Clustering standard errors at the sub-county level is a technique employed when there is a likelihood of correlation among observations within a group or cluster. Here it is relevant, as the program was implemented at the sub-county level, and it may be that the observed differences are due to systemic differences between the treatment and control sub-counties (i.e. there can be significant differences in economic opportunities, demographic patterns or schooling quality that may influence the outcomes).

detected for the RCT carried out on LM's Ghana programme.[33] Furthermore, when it comes to the domain of social and emotional development measured by the IDELA tool however, children in the treatment villages appear to perform more poorly by about 0.2 standard deviations. This is an unexpected result given the LM programme's focus on promoting social interaction and empathy in children through the PS activities.

When breaking the impact estimates down by sub-groups (Tables 14-19), among male children, the negative estimate on socio-emotional development appears to be even more pronounced. Similarly, less wealthy households also experienced a stronger negative impact in this aspect. On the other hand, for female children, a positive effect was observed in the emergent numeracy sub-domain, indicating an improvement in their numerical skills.

Overall, the average level of bias in the kernel-based PSM scores is ranging between 10-20 percent, indicating an imperfect match on key demographic indicators. This suggests, that that the pairs of respondents we tried to match from treatment to control groups based on factors like child age, mother education or household wealth are not as comparable to each other as our estimation strategy would assume. This could indicate that further important information on the households involved weren't captured, and impact estimates are not trustworthy – further clouding the reliability of the impact estimates.

These impact estimates and what they mean for the overall impact of the LM intervention are further discussed in section 11.

9.2. IMPACTS ON SECONDARY OUTCOME MEASURES

When it comes to the intermediary outcomes of the evaluation, the main model did not pick up any significant effects on home environment quality, as measured by HEQ index. Once we account for the starting sample imbalance (discussed in section 8.2) by including the baseline HEQ index values as covariates, this estimate however turns significant, suggesting that on average, mothers in treatment villages were more likely to engage children at home through pro-learning activities by 7 percentage points. Additionally, mothers in treatment villages were more likely to report more positive handwashing behaviours compared to respondents in the control villages (an increase of 24 percentage points on the handwashing index).

Beyond these effects, there were no additional impacts detected on either mother wellbeing or child-parent relationship from the quantitative study.

---

[33] Amadu, S., Attanasio, O.P., Caeyers, B., Cattan, S., Sosa, L.C., Krutikova, S., Leighton, P., Masselus, L. and Yakubu, M., 2018. *Improving early childhood development in rural Ghana through scalable low-cost community-run play schemes: Baseline report* (No. R144). IFS Report.

When we examined the effects within different subgroups, that among boys, there was a negative estimate on the MWB index, however again, this effect was not robust across model specifications, suggesting unreliable estimates. Furthermore, it seems that mothers with female target children reported a more significant improvement in handwashing practices at home.

Additionally, mothers in wealthier households (as measured by the PPI index) were estimated to see a more significant impact on their child's learning environment at home, compared to less wealthy households. Conversely, mothers from less wealthy households were estimated to have seen greater impacts on their wellbeing compared as a result of the Mayuge LM programme.

Nonetheless, as discussed in the previous section, the level of bias in kernel-based PSM, as well as the lack of consistent effects suggest that the evaluation has failed to obtain trustworthy conclusions about the impact of the programme.

10. QUALITATIVE FINDINGS

10.1. PERCEIVED IMPACTS OF THE PROGRAMME

Overall, both the participants and implementers of the programme have talked of various meaningful changes that the programme has brought about in their communities, as well as families. Feedback was collected from mothers, VHT members and various government functionaries, each group revealing a unique perspective on the programme's benefits. The accounts are summarised in the sections below.

10.1.1. IMPACT ON CHILDREN

Children participating in the LM programme have displayed significant behavioural and interpersonal growth, as observed by their mothers. Participants are seen as more disciplined, cooperative, and able to manage their emotions more healthily. Their conflict resolution abilities and emotional management have also notably improved. An important aspect that mothers noted was the improved communication skills among children – as they saw it, the programme has helped children to communicate more respectfully and effectively, using language appropriately, and displaying superior interpersonal abilities that allowed them to form more positive peer relationships. Furthermore, the programme's interactive activities have been instrumental in teaching children effective communication strategies and better understanding of social norms. These, in turn, have translated to behaviours such as sharing, engaging in cooperative play, and demonstrating better control over emotions.

Children were also noted to be better at following instructions, demonstrating more self-control and reduced aggression. There is a noted increase in children's interest in family bonds, positively affecting their parents' behaviour towards them.

Importantly, mothers felt that the programme has also enhanced children's physical skills, with mothers reporting noticeable improvements in their children's motor abilities and general dexterity. Additionally, children have been seen to adopt healthier hygiene practices, including regular handwashing, drinking only from safe sources of water as well as opportunities for healthier eating as mothers understood how to cook more nutritious food – all of which, mothers and VHTs all felt lead to the reduction in instances of illness.

Many children have gained improved cognitive skills, showing progress in counting, problem-solving tasks, and identifying colours. VHTs also noticed a rise in children's confidence and knowledge sharing among peers, with a noted shift towards cooperative behaviour, fostering an environment of camaraderie and mutual support. Children with disabilities have also significantly benefited, experiencing heightened self-confidence, happiness, and active participation.

10.1.2.   IMPACT ON MOTHERS, FAMILIES AND COMMUNITIES

Another fundamental theme from the qualitative interviews has been the LM programme's transformative effect on the participant mothers, families, and communities. Mothers involved in the programme reported a strengthened bond with their children, marked by increased love, closeness, and improved communication. The quality time they've spent with their children through activities like storytelling, teaching, and playing games, has significantly strengthened their relationships. Moreover, mothers participating in the programme have also reported a boost in mood, confidence, and self-esteem. They have felt more empowered to navigate the challenges of parenting with discipline, confidence, and competence, even experiencing an elevated sense of pride due to participation in the programme. The programme has also fostered a sense of unity and mutual support within communities, further strengthening the familial and community bonds.

Community-wide, the Mayuge LM programme filled a gap in parental and child development resources. Mothers have become more proactive in promoting the health, hygiene, and overall well-being of their children, even including children with disabilities more actively. There's a marked increase in community interest in child health and education. Fathers too are more involved in their families, providing better home support and school fees. There has also been a noted decrease in domestic violence cases, pointing to healthier family dynamics.

10.1.3.   IMPACT ON VHTs AND OTHER GOVERNMENT FUNCTIONARIES

VHTs reported significant changes in their own well-being and skills since the programme began. They've noticed a positive impact on their relationship with participant mothers, fostering a strong bond and friendship. The programme has empowered them with the knowledge to teach children and address the community confidently, improving their communication skills. They've also gave accounts of how their English reading and translation skills have improved as a result of the programme, increasing their confidence and self-esteem.

VHTs have felt they have gained respect from local leaders, being viewed as change-makers and problem solvers. Their ability to assist with government initiatives has increased, strengthening their community relationships. In some cases, VHTs reported helping to set up savings schemes for the mothers in the community in order to better provide community support to each other and their children.

The programme has been instrumental in creating a bridge between government functionaries and local communities, fostering a mutual understanding of needs, concerns, and strategies. The initiative has promoted teamwork among officials while implementing programme activities, leading to improvements in efficiency, proactiveness, and engagement. The functionaries have also reported improvement in their planning, problem-solving, and supervisory skills due to the programme.

10.2. ADDRESSING CRITICAL GAPS

The Mayuge LM programme was established to address a variety of gaps in rural Ugandan communities, particularly in relation to ECD and parenting. Of particular interest to the impact evaluation, was finding out whether some of the practices the mothers were trained on were already being followed in the communities (such as following hygienic behaviours, active play or engaging children through reading/storytelling). This was important to establish given the parallel trends assumption highlighted for the DiD estimator in section 9.

The study examined how the programme, which includes a radio show and parent training sessions, has positively impacted these communities.

10.2.1.   PRE-EXISTING PRACTICES

From the qualitative accounts of mothers, VHTs and government functionaries, it seems that before the Mayuge LM programme was introduced there were considerable knowledge and service gaps in the communities. The most apparent gaps were in child support and education for children aged 3-6. While nursery school

was available in some areas, it was not universally accessible and was considered inadequate in providing quality learning experiences. The Mayuge LM programme has filled these gaps, providing novel, relevant and actionable information about quality care and education to parents who had previously been under-informed.

The study revealed differing parental habits regarding pro-learning engagements at home before the programme's implementation. Some parents, particularly in the treatment villages, reported regularly spending time with their children in engaging ways such as singing, storytelling, and games. Other mothers admitted they rarely engaged their children in these activities before participating in the programme. However, as the Mayuge LM programme encouraged more active engagement, parents who were already actively engaging their children on a regular basis were further motivated, and they even supplemented their activities with what they learned from the programme.

10.2.2. OTHER PROGRAMMES DURING THE EVALUATION PERIOD

In terms of other programmes, the study found that there were no other notable programmes besides the Mayuge LM intervention that affected the treatment and sub-counties systematically differently during the evaluation period (May 2022-February 2023). There were mentions of other programmes mentioned in some villages like those by World Vision, but their focuses were on different issues such as child protection, sponsorship, and provision of bed nets, rather than ECD. VHT members also confirmed that the Mayuge LM programme offered unprecedented opportunities for parents to learn about providing care and education for young children. According to them, there were no similar opportunities available prior to the programme and the level of mobilisation that the programme could raise has really helped ensure that participants were more committed to its mission.

10.2.3. ROLE OF THE LMTRP

A significant part of the Mayuge LM initiative is the LMTRP), which broadcasts instructions on promoting young children's learning, reached over a third of mothers across both treatment and control villages. Its listenership grew from 2021, especially during the COVID-19 pandemic, and it was praised for providing valuable information on various topics, including quality time with children, community development, cognitive development activities, and parenting practices. This in fact was an additional concern to the internal validity of the impact evaluation, as the radio programme was accessible to the mothers in the control villages as well. Specifically, families in the control villages could have been encouraged to start

implementing some of the pro-learning activities with children even before the full programme arrived in their community.

Nonetheless, despite the usefulness of the LMTRP, mothers and VHT members felt the parenting workshops provided through the Mayuge LM programme offered much more effective learning experiences. They noted the workshops were more hands-on, interactive, and allowed for questions and demonstrations, thus making concepts easier to understand and apply. These workshops and the overall Mayuge LM programme have evidently begun to fill the previously existent gaps in the community, enhancing ECD and parenting abilities.

10.3. OVERALL PROGRAMME EXPERIENCE AND COMPLIANCE

Overall, interviewed mothers reported personal stories of meaningful growth, improved welfare and health for their children, new friendships, and a stronger sense of community support. They also praised the accessible and valuable education their children received, which provided them with important skills and knowledge. Mothers acknowledged the VHTs for their valuable support and guidance during the programme, particularly in the play schemes' implementation. Participant mothers appreciated the recognition of their efforts and the knowledge they gained, fostering a sense of respect and community cohesion. With all these positive experiences, when asked about the programme's sustainability, respondents expressed optimism about continued participation, affirming that they see how the programme is providing direct value to their families. They pointed to the unity and collaboration among community members, which further reinforced their confidence in the programme's longevity. On the other hand however, VHTs and participating mothers observed that after a while, mothers with older children felt the program was losing relevance for them, as they thought the parenting workshops and play schemes were primarily targeting parents with young children. Additionally, VHTs reported that some mothers dropped out of the program due to expecting payment for their involvement, potentially affecting the composition of respondents who remained by the time the qualitative study was conducted. These learnings highlight the potential limitations and considerations regarding the program's effectiveness and the profile of participants involved.

Furthermore, VHTs also noted the programme's effectiveness in uniting their communities and highlighted the transformative potential of the programme, as evidenced by the introduction of savings groups in some communities. They also mentioned the visible benefits of the programme to participating mothers, which motivated them to continue their participation.

With the significant improvements in their knowledge, capacity, and communication skills that VHTs reported obtaining through the programme, they also felt committed to sustaining the programme. VHTs felt more proactive, innovative, and confident, which positively influenced their relationship with the community and increased participation. This change was also experienced by district and sub-county officials, who were confident they'd remain committed to the programme goals.

11. DISCUSSION AND CONCLUSIONS

The results of the quantitative survey indicate that the intervention had only limited impact on children's school readiness, as well as on many secondary indicators. In this section we dive deeper into the assumptions of the DiD model, discussing how the findings relate to the results of the qualitative survey and present the takeaways from the evaluation of the LM programme as a whole.

11.1. REVIEW OF MODEL ASSUMPTIONS

As discussed in Section 8.1, the DiD estimator used in the analysis makes a series of assumptions when producing estimates of programme impact – in this section, we review them one-by-one to give more colour on the discussion.

1. **The treatment assignment was not affected by any unobserved confounding variables that may affect the outcome.**
   In the case of the LM intervention, all study areas were set to eventually implement the programme and the phasing of the intervention was arbitrarily chosen by the district staff. Nonetheless, this alone could not necessarily rule out systematic differences between the two geographies[34].

   An example situation in which this assumption would be violated is if in the treatment villages another programme or infrastructural development (i.e., the building of new schools) were to be implemented that will affect the primary outcomes, but that programme is absent in control villages. As mentioned in section 10, the follow-up qualitative survey indicates that no notable NGO activities, government programmes or infrastructural development have affected the treatment and control villages systematically differently during the evaluation period.

---

[34] While we have tried to use the Living Standards Measurement Survey (2021) , as well as the Demographic and Household Survey (2021) from Uganda to verify this assumptions, we found no readily available secondary datasets to match key socioeconomic or other indicators at the level of the study villages.

2. **The propensity score distributions for the treatment and control groups must overlap, which means that there are individuals in both groups with similar probabilities of receiving the treatment.**

An example situation in which this assumption would be violated would be if all children in treatment areas would be aged five and above, while survey participants in control villages would be aged four and below. In that case, there wouldn't be a large enough number of "similar" child respondents in the control group that the treatment children could be compared against. As the balance tables in Appendix C show, this was not the case with most of the included control variables, such as child age, mother's education or mothers' age.

In the case of PPI scores, a notable limitation is the lack of indicators capturing household wealth at baseline. This means that we are unable to control for any variation that could have impacted household wealth between baseline and endline. Nonetheless, given the Theory of Change specified in Figure 1, our understanding is that the intervention was not designed to directly affect household income, but primarily children's and mothers' attitudes, capacity and behaviour.

As mentioned is section 9, the mother-child respondent pairs could not be reliably matched between treatment and control groups based on observable characteristics. This in turn means that given that the model could not establish a credible counterfactual, or "what would have happened to the treatment mother-child pairs, if they hadn't participated in the Mayuge LM programme".

3. **The parallel trends assumption requires that the treatment and control groups would have followed similar trends in the absence of the treatment. This means that respondents in the treatment group who participated in the LM intervention would have had similar outcomes in the absence of the programme to the outcomes of control participants.**

The presence of sample imbalance in the DiD analysis made it challenging to accurately verify the parallel trends assumption. The unequal distribution of observations between the treatment and control groups introduced potential biases and confounded the estimates. Unfortunately, since no pre-baseline data is available from the study communities, this assumption cannot be satisfactorily verified.

One way the model aims to correct for any imbalances, is by including mother-fixed effects in its estimates. One way the model aims to correct for any imbalances, is by including mother-fixed effects in its estimates. This helps us consider in the model any

differences between mothers that don't change over time and might impact the outcome we're looking at (i.e. mother's motivation).

Additionally, the co-variates of child gender, child and mother age, PPI score and mother education aim to improve the accuracy and reliability of the results. These covariates are variables that are related to both the treatment and outcome variable but are not part of the treatment assignment process based on the sampling protocols. In the model, respondents from control and treatment areas that are similar along these characteristics are each assigned a propensity score, which is incorporated in the sample-level impact estimates.

4. **The treatment effect must be stable over time, which means that the effect of the LM intervention does not meaningfully change over time**.

For example, the intervention may not have been intensive or sustained enough to influence these outcomes, or the measures used to assess them may not have been sensitive enough to detect changes. In the case of the LM intervention, respondents in treatment villages have on average participated in the programme for seven to eight months already by the time the endline data was collected. Accounting with an average of two hours of PS activities in the week per child, for any child this would involve an approximate exposure of between 20 to 30 sessions before the endline data collection. For comparison, in the case of the randomised evaluation in Ghana, sizeable effects on key IDELA scores were detected after 30 sessions.[35]

From the qualitative fieldwork, respondents reported a generally strong implementation quality on the ground – suggesting that implementation standards were in theory sufficient to bring about more positive impacts on the treated children.

Many measurements used in the study don't show enough differences between the participants. This can happen when the outcome variable is already very close to its highest or lowest value – this would be the case, if for instance most mothers had scored 95% or above in the wellbeing score (MWB index), as there wouldn't be more "room" for that score to improve. This limited range of scores can make it hard to see any meaningful changes or differences.

---

[35] [35] Amadu, S., Attanasio, O.P., Caeyers, B., Cattan, S., Sosa, L.C., Krutikova, S., Leighton, P., Masselus, L. and Yakubu, M., 2018. *Improving early childhood development in rural Ghana through scalable low-cost community-run play schemes: Baseline report* (No. R144). IFS Report.

To handle this, we take into account the starting score of each participant for a given outcome variable (i.e. Total IDELA Score). This helps us better understand the effect of the programme by considering how much their scores change from the beginning. Most notably, in the case of the HEQ index this adjustment greatly increases the estimated effect size, as well as matching quality (

Table 13).

11.2. TAKEAWAYS

The overall aim of this study was to evaluate the impact of the Mayuge LM programme on participant mothers and their children. The quantitative survey aimed to measure changes in children's school-readiness (estimated by IDELA scores), as well as effects on a series of secondary outcomes such as positive learning environment at home, mother's wellbeing, child-parent relationship or hygienic behaviours of mothers.

The evaluation utilised a quasi-experimental research design, where establishing comparable treatment and control units is crucial for obtaining valid impact estimates due to the non-random selection of study areas. Nonetheless, this objective was not fully achieved due to factors such as statistically significant imbalance at baseline between treatment and control groups; differential attrition (different types of children and mothers dropping out from the endline study); poor-quality statistical matching; and unstable point estimates. As a result, the quantitative survey failed to produce reliable answers on the impact of the Mayuge LM programme.

In the Play Schemes, children displayed increased confidence with solving difficult problems, improved discipline, cooperation, emotional management, and conflict resolution abilities. At home, mothers noted how children improved their communication skills, used more respectful language, and developed more positive friendships. The LM programme also had a transformative effect on mothers, families, and communities. Mothers reported strengthened bonds, increased love, and improved communication with their children. Quality time spent together strengthened relationships and boosted mothers' mood, confidence, and self-esteem. Mothers felt empowered in parenting and experienced a sense of pride. At the community level, programme fostered unity, mutual support, and community interest in child health and education. Fathers became more involved in their families, leading to healthier family dynamics and a decrease in domestic violence cases. The consistency of these accounts of positive changes across different types of participants, functionaries or communities speaks to the transformative changes that the programme brought about.

In conclusion, although the quantitative analysis faced limitations that prevent us from making definitive claims about the program's impact, the qualitative accounts provide

encouraging insights into how innovative, and resource-friendly programmes like Lively Minds can fill critical gaps in communities with poor access to quality ECCE services.

In conclusion, unfortunately the methodological limitations of the study prevent us from drawing conclusive findings regarding the program's impact. While the qualitative accounts point to meaningful impacts in the participating communities, more extensive research will be needed to estimate the impact of the programme, as implemented by government functionaries in low-resource settings. In line with that next step, LM is currently conducting a large-sample Randomised Controlled Trial in Ghana in collaboration with Yale University, which we hope can further corroborate the feedback of mothers, VHTs and other functionaries of the Mayuge LM programme.

APPENDIX A : OVERVIEW OF SURVEY PARAMETERS

Table 7 below shows the total number of mother-child pairs interviewed at baseline and endline.

In some communities, there were no sufficient replacement mother-child pairs available during the survey. In order to ensure that the final sample size is still retained, respondents in other study villages have been oversampled - hence in certain villages the number of respondents surveyed at endline is higher than at baseline.

Table 7: Overview of data collection targets by cluster

Study Geographies				Total number of mother-child pair interviews collected	
Group	Sub-county	Parish	Village	Baseline	Endline*
Treatment	Bukabooli	Bugoto	Bugoto B	13	12 (25%)
		Bugoto	Butumbula	16	16 (13%)
		Bugumya	Kirongo A	17	17 (65%)
		Mairinya	Lwandera	20	19 (0%)
		Mairinya	Nawandegeyi	20	20 (45%)
	Kigandalo	Bugondo	Walukoko	19	19 (11%)
		Isenda	Katalakabi	15	16 (19%)
		Kigandalo	Nakazigo	22	23 (13%)
		Kigulu	Namugolo	20	20 (10%)
		Kyoga	Nakavule	15	15 (13%)
		Maleka	Kigulamo	20	20 (0%)
		Maleka	Maleka B	20	20 (25%)
Control	Busakira	Bukunja	Namisu B	17	17 (24%)
		Butangala	Namiro	12	13 (31%)
		Kaluba	Busakira	22	22 (14%)
		Maumo	Busere B	17	17 (12%)
	Kityerera	Bubinge	Busimo	17	17 (6%)
		Bubinge	Kyikandwa	17	17 (0%)
		Bukalenzi	LutaleA	17	17 (24%)
		Bukalenzi	Ngayama	18	18 (33%)
		Kitovu	Kitovu	19	18 (22%)
		Kitovu	Lutale B	15	15 (0%)
		Ndaiga	Nakyilumila C	16	16 (19%)
		Ndaiga	Ndaiga A	18	18 (17%)
				422	**422**

* Percentages in brackets display the proportion of baseline respondents that that had to be replaced due to attrition.

Table 8: Demographic breakdown of sampled mothers and children 43

Indicator	Value	Percentage of respondents	Standard Error	N
Levels of education attained (mother)	None	14.49	1.72	
	Some primary school	49.88	2.44	
	Primary school	15.91	1.78	
	Some O'level	14.25	1.70	421
	O'level certificate	4.75	1.04	
	Some A'level	0.48	0.34	
	A' level certificate	0.24	0.24	
Age groups (mother)	18 to 24	20.62	1.97	
	25 to 29	20.85	1.98	
	30 to 34	15.17	1.75	
	35 to 39	14.93	1.73	422
	40 to 44	9.72	1.44	
	45 to 50	9.24	1.41	
	. > 50	9.48	1.43	
Percentage of mothers who are heads of their household		20.62	1.97	422
Number of biological children living in the household	One	8.06	1.32	
	Two	16.11	1.79	
	Three	15.17	1.75	
	Four	20.38	1.96	422
	Five	15.64	1.77	
	Six	13.74	1.68	
	Seven or more	10.90	1.52	
Age of the child	3.00	12.32	1.60	
	4.00	18.72	1.90	422
	5.00	31.99	2.27	
	6.00	36.97	2.35	
Gender of the child	Female	55.69	2.42	422
	Male	44.31	2.42	

Note: Sample sizes vary across indicators due to missing responses, "Don't know/Refused" answers or skip patterns coded into the questionnaire logics

APPENDIX B : ADDITIONAL INFORMATION ON KEY VARIABLES

Primary Outcomes

1. IDELA summary variable creation

In order to offer better measurements for ECD, Save The Children created a new evaluation tool in 2011, based on a comprehensive review of existing child development assessment tools and insights from years of work in early childhood programming.

The aim of IDELA is to provide early childhood care and development programmes, donors, and government partners with clear evidence on the status of children from three to six years. The IDELA tool is designed to empower actors to improve access and quality in early childhood development through data driven decision-making. The IDELA tool is used to understand what skills children bring to their first day of school, to measure development and early learning, and to understand issues of equity in young learners.

The questionnaire items of the IDELA tool fall into one of the four major sub-domains of child development, that together make up children's proxy for "school readiness". The table below presents each questionnaire item with its corresponding sub-domain, as per the IDELA. Note that there is a 5th, optional sub-domain of "Executive Function". During the baseline measurements, these optional items were omitted from the questionnaire to boost the survey productivity and reduce the required days of data collection.

The section below outlines the method for the construction of IDELA total and sub-domain scores:[36]

- ✓ Each sub-question within one of the 24 question items is added up, to create a total score for the given question item. The value range of most of these sub-questions is between 0 and 1, or 999 for missing observations. The total score achieved is obtained as a simple total of all the sub-question items. Some questions allow the child to receive extra points for understanding the question or just being motivated throughout the exercise.

- ✓ Create ratio-based score for each item. This includes dividing the sum of all sub-question scores by the highest number attainable.
  - o Example: Total score of shape identification item/5.

- ✓ Create total domain scores by adding the ratio correct for all domain items and divide by total number of items.

---

[36] Guide to creating summary IDELA scores, made available by Save The Children at: https://idela-network.org/

o Example: Add percentage correct columns for all numeracy items and divide by 7 = (measurement+sorting+shape id+number id+one-to-one correspondence+simple operations+puzzle)/7.

✓ Create the total IDELA score by adding the proportion of correct scores across all core domain scores EXCEPT executive function and learning approaches and divide by the total number of domains.
   o Example: (motor+literacy+numeracy+social-emotional)/4.

✓ The values of the IDELA scores range from 0-1, with higher values corresponding to a higher percentage of questions that the child has scored correctly on.

2. Handwashing index

The handwashing index is constructed by summing together the "occasions" of a series of dummy variables (0=Does not wash hands during occasion "X", 1=Washes hands during occasion "X"). The occasions "X" include:
  ➢ *Before preparing food*
  ➢ *Before eating*
  ➢ *After eating*
  ➢ *Before feeding a child*
  ➢ *After cleaning a child's anus/changing nappy*
  ➢ *After disposing of children's faeces*
  ➢ *After toileting/defecation*
  ➢ *After returning from outside the compound*
  ➢ *After touching a sick person*
  ➢ *After handling livestock*
  ➢ *After handling cow dung or any animal faeces*
  ➢ *After greeting a lot of people*
  ➢ *After waking up in the morning*
  ➢ *Respondent never washes hands with soap/ash*

  ➢ The final index is obtained by the formula handwash_index = [total score]/[total number of non-missing statements]. If the person chose "Respondent never washes hands with soap/ash", the index automatically took the value of 0 – however from the 422 respondents, this was not the case for any mother.
  ➢ The values can range from 0-1, with higher values indicating more positive handwashing practices.

Secondary Outcomes

1. Home Environment Quality index

The HEQ index is constructed by summing together the number of times a "pro-learning" activity was being conducted with the child in the mother's household (by any household member) in the three-day period prior to the interview. Outliers were probed for, but for each pro-learning activity, the number of times range between zero and ten typically, with over 95% of cases falling below four occasions for a given activity.

The final index is obtained by the formula HEQ_index = [total score]/[total number of non-missing statements]. The values range from 0-4 in our case, with higher scores corresponding to households where parents spend more quality time with their children.

2. Mother Wellbeing (MWB) index

The Mother wellbeing index is conducted using the **Rosenberg (1965)[37]** self-esteem scale. From the list of 10 items, the questionnaire selected a random subset of five questions in order to reduce survey length.

The MWB index is constructed by summing together the Likert-scale type responses (1= Strongly Agree …4=Strongly disagree) of the following statements):

- **Positively framed** (Value 1= High MWB, Value 4 = Low MWB)
  - *On the whole, I am satisfied with myself.*
  - *I feel I have a number of good qualities.*
  - *I am able to do things as well as most other people.*
  - *I feel that I'm a person of worth, at least on an equal plane with others.*
  - *I take a positive attitude towards myself.*

- **Negatively framed** (Value 1= Low MWB, Value 4 = High MWB)
  - *At times I think I am no good at all.*
  - *I feel I do not have much to be proud of.*
  - *I certainly feel useless at times.*
  - *I wish I could have more respect for myself.*
  - *All in all, I am inclined to feel that I am a failure.*

---

[37] Rosenberg, M. (1965). Rosenberg Self-Esteem Scale (RSE). Acceptance and Commitment Therapy. Measures Package, 61, 5

The above values for <u>positively</u> framed statements are then transformed – 4 is recoded to 1, 3 is recoded to 2, 1 is recoded to 4 and so on. The final index is obtained by the formula MWB = [total score]/[total number of non-missing statements]. The values can range from 1-4.

3. Child-Parent Relationship (CPR) index

The CPR index is constructed by summing together the Likert-scale type responses (1= Definitely does not apply …5=Definitely applies) of the following statements):

- **Positively framed** (Value 1= Low CPR, Value 5 = High CPR)
  - ➢ My child will seek comfort from me.
  - ➢ My child values his/her relationship with me.
  - ➢ When I praise my child, he/she beams with pride.
  - ➢ My child spontaneously shares information about himself/herself.
  - ➢ It is easy to be in tune with what my child is feeling.
  - ➢ My child open shares his/her feelings and experiences with me.

- **Negatively framed** (Value 1= High CPR, Value 5 = Low CPR)
  - My child and I always seem to be struggling with each other.
  - My child is uncomfortable with physical affection or touch from me.
  - My child easily becomes angry at me.
  - My child remains angry or is resistant after being disciplined.
  - Dealing with my child drains my energy.
  - When my child wakes up in a bad mood, I know we're in for a long and difficult day.
  - My child's feelings towards me can be unpredictable or can change suddenly.
  - My child is sneaky or manipulative with me.

The above values for <u>negatively</u> framed statements are then transformed – 5 is recoded to 1, 4 is recoded to 2, 1 is recoded to 5 and so on. (This differs from MWB, given that the direction of coding was different at the start – however for both indices, large values correspond to more positive feelings). The final index is obtained by the formula CPR = [total score]/[total number of non-missing statements]. The values can range from 1-5.

Control Variables

1. PPI score

The PPI score is constructed using the simple Ugandan poverty scorecard following Schreiner (2015).[38]

The questions that make up the score include:

- o 1. How many members does the household have?
- o 2. Are all household members aged 6 to 12 currently in school?
- o 3. Can the female head/spouse read and write with understanding in any language?
- o 4. What type of material is mainly used for construction of the wall of the dwelling?
- o 5. What type of material is mainly used for construction of the roof of the dwelling?
- o 7. What type of toilet facility does the household mainly use?
- o 8. How many mobile phones do members of your household own?
- o 9. Does any member of your household own a radio?
- o 10. Does every member of the household have at least one pair of shoes?

The measures were self-reported by the respondent mothers during the survey at the community centre. A high PPI score is associated with higher levels of estimated household wealth, whereas lower levels of PPI score indicate households with a lower household wealth.

2. Child age
Child age is captured as an integer and refers to the age of the respondent child that the IDELA measurements were conducted with. The measure was self-reported by the mothers with some level of uncertainty, as respondent mothers would often inconsistently estimate their children's ages, requiring additional probing from the field enumerators.

3. Mother's education
Mother's education was captured as the highest level of formal education that the mother has obtained at the time of the survey. This indicator has been captured under the following categories:

---

[38] Schreiner, M. (2012). Simple Poverty Scorecard® Poverty-Assessment Tool. Available at https://www.simplepovertyscorecard.com/UGA_2012_ENG.pdf

Table 9: Overview of mothers' education variable coding

Numeric Value	Highest level of formal education
1	None
2	Some primary school
3	Primary school
4	Some O'level
5	O'level certificate
6	Some A'level
8	Voc./Tech./Commercial
10	Diploma

4. Child gender

Captured as a binary variable, with "1" denoting male children and 0 denoting female children.

APPENDIX C: SAMPLE BALANCE TABLES

Table 10: T-test of sample characteristics between treatment and control respondents, across two rounds

Study Group	Indicator	Baseline		Endline	
		Mean	P-value[39]	Mean	P-value
Control	Total IDELA score	0.446*	0.0916	0.528	0.3181
Treatment		0.475*	0.0916	0.545	0.3181
Control	IDELA Sub-domain score: Socio-Emotional Development	0.42	0.4189	0.516	0.2716
Treatment		0.438	0.4189	0.493	0.2716
Control	IDELA Sub-domain score: Gross and Fine motor skills	0.607*	0.0926	0.73	0.2241
Treatment		0.649*	0.0926	0.756	0.2241
Control	IDELA Sub-domain score: Emergent Literacy	0.289**	0.0326	0.34	0.4523
Treatment		0.326**	0.0326	0.353	0.4523
Control	IDELA Sub-domain score: Emergent Numeracy	0.469	0.2515	0.528***	0.0053
Treatment		0.49	0.2515	0.577***	0.0053
Control	Handwashing Index	0.281***	0.0000	0.304**	0.0121
Treatment		0.346***	0.0000	0.341**	0.0121
Control	Mother wellbeing (MWB) index	2.886***	0.0001	2.936	0.2050
Treatment		3.062***	0.0001	2.983	0.2050
Control	Child-Parent Relationship (CPR) index	3.997	0.2301	4.067	0.2377
Treatment		4.065	0.2301	4.135	0.2377
Control	Home Environment Quality (HEQ) index	1.089***	0.0000	1.118***	0.0000
Treatment		1.539***	0.0000	1.512***	0.0000
Control	Home Environment Quality (HEQ) index | Reading with child	0.82***	0.0000	1.068***	0.0010
Treatment		1.438***	0.0000	1.484***	0.0010
Control	Home Environment Quality (HEQ) index | Telling stories	0.937***	0.0024	0.937***	0.0001
Treatment		1.3***	0.0024	1.419***	0.0001
Control	Home Environment Quality (HEQ) index | Singing	1.463***	0.0001	1.551***	0.0026
Treatment		1.968***	0.0001	1.945***	0.0026
Control	Home Environment Quality (HEQ) index | Playing with child	0.883***	0.0000	0.98***	0.0001
Treatment		1.53***	0.0000	1.479***	0.0001
Control		0.722***	0.0000	0.61***	0.0000

---

[39] The p-values in the table show if the differences in scores between the baseline and endline for the control and treatment groups are statistically significant. A p-value measures the strength of evidence against the idea that the groups are not different. To interpret these values, we compare them to a significance level (usually 0.05). If the p-value is below this level, it suggests strong evidence of a significant difference. However, in this case, the p-value is higher than 0.05, indicating that the observed differences are not statistically significant – nonetheless, we still report them at 10% level. In the example, the p-value for comparing the Total IDELA score for both groups at baseline is 0.0916, suggesting that there is was a moderately significant imbalance between control and treatment children at the start of the evaluation.

Treatment	Home Environment Quality (HEQ) index | Reading with child	**1.24***	0.0000	**1.097***	0.0000
Control	Home Environment Quality (HEQ) index | Going outside with child	1.707	0.6857	1.561	0.5300
Treatment		1.76	0.6857	1.645	0.5300
Control	PPI score	**36.339**	0.0210	**36.498***	0.0521
Treatment		**38.806**	0.0210	**38.378***	0.0521
Control	PPI sub-score: Wall material	**2.105***	0.0055	**2.185**	0.0114
Treatment		**2.69***	0.0055	**2.667**	0.0114
Control	PPI sub-score: No of HH members	3.363	0.4081	3.439	0.5105
Treatment		3.114	0.4081	3.249	0.5105
Control	PPI sub-score: Schooling levels	**1.665***	0.0060	**1.554***	0.0014
Treatment		**2.011***	0.0060	**1.94***	0.0014
Control	PPI sub-score: Literacy of female head	0.211	0.8830	0.22	0.8719
Treatment		0.223	0.8830	0.207	0.8719
Control	PPI sub-score: Roof material	4.375	0.3178	4.455	0.9234
Treatment		4.543	0.3178	4.47	0.9234
Control	PPI sub-score: Main energy source	0.105	0.9773	0.088	0.5279
Treatment		0.103	0.9773	0.138	0.5279
Control	PPI sub-score: Toilet facility	5.187	0.5498	5.19	0.2277
Treatment		5.091	0.5498	5.018	0.2277
Control	PPI sub-score: Number of mobile phones	10.544	0.9352	10.659	0.6493
Treatment		10.589	0.9352	10.429	0.6493
Control	PPI sub-score: Radio ownership	**3.029***	0.0666	3.176	0.1416
Treatment		**3.72***	0.0666	3.677	0.1416
Control	PPI sub-score: Shoes	**5.842**	0.0438	**5.62**	0.0171
Treatment		**6.737**	0.0438	**6.594**	0.0171
Control	Mother's education | None	0.141	0.8613	0.102	0.3227
Treatment		0.147	0.8613	0.134	0.3227
Control	Mother's education | Some Primary School	0.507	0.6997	0.478	0.1242
Treatment		0.488	0.6997	0.553	0.1242
Control	Mother's education | Primary School	0.141	0.3456	0.151	0.8769
Treatment		0.175	0.3456	0.157	0.8769
Control	Mother's education | Some Secondary School	0.161	0.2837	**0.215***	0.0004
Treatment		0.124	0.2837	**0.092***	0.0004
Control	Mother's education | Secondary School	0.044	0.7436	0.054	0.7214
Treatment		0.051	0.7436	0.046	0.7214
Control	Mother's education | A-levels and above	0.005	0.3443	**0***	0.0510
Treatment		0.014	0.3443	**0.018***	0.0510
Control	Age of target child	4.868	0.1869	5.712	0.4160
Treatment		5	0.1869	5.797	0.4160
Control	Age of mother respondent	**32.868**	0.0464	34.507	0.7977
Treatment		**34.876**	0.0464	34.765	0.7977

Control	% of respondents listening to LMTRP	**0.322***	0.0010	**0.346***	0.0001
Treatment		**0.479***	0.0010	**0.539***	0.0001
Control	# of months the child has been enrolled in a educational institution (January 2023)	13.133	0.5559	13.26	0.2882
Treatment		12.443	0.5559	12.08	0.2882
Control	% of male target children	0.463	0.4160	0.454	0.6049
Treatment		0.424	0.4160	0.429	0.6049
* - significant at 90% level		**** - significant at 95% level**		***** - significant at 99% level**	

Table 11: T-test of dropout, replaced and remaining sample of respondents

Study Group	Indicator	Mean (Remained) N=346	Mean (Dropout) N=76	P-value	Mean (Remained) N=346	Mean (Dropout) N=76	P-value
Treatment	Total IDELA score	0.482	0.446	0.242	**0.571***	**0.435***	0
Control		0.451	0.417	0.302	**0.549***	**0.425***	0
Treatment	IDELA Sub-domain score: Socio-Emotional Development	0.448	0.394	0.16	**0.519***	**0.388***	0
Control		0.423	0.405	0.673	**0.540***	**0.394***	0
Treatment	IDELA Sub-domain score: Gross and Fine motor skills	0.657	0.618	0.359	**0.794***	**0.602***	0
Control		**0.622***	**0.529***	0.07	**0.753***	**0.614***	0.001
Treatment	IDELA Sub-domain score: Emergent Literacy	0.326	0.325	0.98	**0.371***	**0.281***	0.006
Control		0.291	0.276	0.624	**0.358***	**0.247***	0.001
Treatment	IDELA Sub-domain score: Emergent Numeracy	0.498	0.458	0.23	**0.603***	**0.471***	0
Control		0.471	0.460	0.734	**0.544***	**0.444***	0.002
Treatment	Handwashing Index	**0.336***	**0.388***	0.091	0.338	0.352	0.639
Control		0.287	0.251	0.156	0.306	0.296	0.706
Treatment	Mother wellbeing (MWB) index	3.062	3.062	0.998	**3.013**	**2.862**	0.022
Control		2.897	2.829	0.429	2.931	2.959	0.705
Treatment	Child-Parent Relationship (CPR) index	4.088	3.969	0.227	4.122	4.190	0.472
Control		4.021	3.878	0.198	4.079	4.004	0.524
Treatment	Home Environment Quality (HEQ) index	1.568	1.421	0.357	1.550	1.349	0.232
Control		1.057	1.250	0.219	1.130	1.059	0.657
Treatment	PPI score				38.806	36.595	0.195
Control					36.339	37.294	0.609
Treatment	PPI sub-score: Wall material				2.690	2.571	0.717
Control					2.105	2.588	0.198
Treatment	PPI sub-score: No of HH members				3.114	3.810	0.159
Control					3.363	3.824	0.424
Treatment	PPI sub-score: Schooling levels				**2.011***	**1.643***	0.073
Control					**1.665***	**0.969***	0.004
Treatment					0.223	0.143	0.543

Control	PPI sub-score: Literacy of female head				0.211	0.265	0.714
Treatment	PPI sub-score: Roof material				4.543	4.167	0.156
Control					4.375	4.853	0.104
Treatment	PPI sub-score: Main energy source				0.103	0.286	0.239
Control					0.105	0.000	0.439
Treatment	PPI sub-score: Toilet facility				5.091	4.714	0.124
Control					5.187	5.206	0.947
Treatment	PPI sub-score: Number of mobile phones				10.589	9.762	0.351
Control					10.544	11.235	0.483
Treatment	PPI sub-score: Radio ownership				3.720	3.500	0.716
Control					3.029	3.912	0.179
Treatment	PPI sub-score: Shoes				6.737	6.000	0.284
Control					5.842	4.500	0.102
Treatment	Mother's education | None	**0.126***	**0.238***	0.066	0.143	0.095	0.418
Control		0.152	0.088	0.332	0.094	0.147	0.35
Treatment	Mother's education | Some Primary School	0.497	0.452	0.604	**0.520****	**0.690****	0.046
Control		**0.468****	**0.706****	0.011	0.503	0.353	0.111
Treatment	Mother's education | Primary School	0.177	0.167	0.873	0.160	0.143	0.785
Control		0.152	0.088	0.332	0.140	0.206	0.332
Treatment	Mother's education | Some Secondary School	0.131	0.095	0.526	0.097	0.071	0.607
Control		0.175	0.088	0.208	0.216	0.206	0.892
Treatment	Mother's education | Secondary School	0.051	0.048	0.92	0.057	0.000	0.114
Control		0.047	0.029	0.653	0.047	0.088	0.33
Treatment	Mother's education | A-levels and above	0.017	0.000	0.395	0.023	0.000	0.325
Control		0.006	0.000	0.657	0.000	0.000	0
Treatment	Age of target child	4.983	5.071	0.602	**5.983*****	**5.024*****	0
Control		4.895	4.735	0.425	**5.895*****	**4.794*****	0
Treatment	Age of mother respondent	34.320	37.190	0.124	35.320	32.452	0.125
Control		33.205	31.176	0.268	34.205	36.029	0.317
Treatment	% of respondents listening to LMTRP	0.491	0.429	0.466	**0.491*****	**0.738*****	0.004
Control		0.327	0.294	0.705	0.327	0.441	0.205
Treatment	# of months the child has been enrolled in a educational institution (January 2023)				12.443	9.870	0.196
Control					13.133	14.400	0.66
Treatment	% of male target children	0.429	0.405	0.78	0.429	0.429	1
Control		0.439	0.588	0.111	0.439	0.529	0.334

* - significant at 90% level   ** - significant at 95% level   *** - significant at 99% level

APPENDIX D: OVERVIEW OF IMPACT ESTIMATES

The following section shows the impact estimates of the Mayuge LM programme on a series of primary and secondary outcomes as defined in the programme's Theory of Change (Figure 1). Programme impacts were estimated using the control variables specified in section 6, as well as with an additional series of model adjustments to test the sensitivity of estimates to the a) exclusion/inclusion of additional control variables; b) clustering of standard errors at different units and c) adding fixed effects[40].

In the current section we report the results of three distinct specifications in line with the original research strategy:

- ✓ **Model (i):**
    - o **Method**: Differences-in-Differences with kernel-based propensity-score matching
    - o **Outcome variables**: All primary and secondary outcomes
    - o **Control variables**: Mother's age, Mother's education, PPI, Child age, Child gender
    - o **Fixed effects**: N/A
    - o **Standard Error Clustering**: N/A

- ✓ **Model (ii):**
    - o **Method**: Differences-in-Differences with kernel-based propensity-score matching
    - o **Outcome variables**: All primary and secondary outcomes
    - o **Control variables**: Mother's age, Mother's education, PPI, Child age, Child gender, [Outcome] value at Baseline
    - o **Fixed effects**: Mother fixed effects
    - o **Standard Error Clustering**: N/A

- ✓ **Model (iii):**
    - o **Method**: Differences-in-Differences with kernel-based propensity-score matching

---

[40] For additional result tables on the robustness checks conducted and the specific methods used, please contact the author or the researchers involved in the study.

- o **Outcome variables**: All primary and secondary outcomes
- o **Control variables**: Mother's age, Mother's education, PPI, Child age, Child gender, [Outcome] value at Baseline
- o **Fixed effects**: N/A
- o **Standard Error Clustering:** Sub-County level

Most notably, models (ii) and (iii) account for potential ceiling/floor effects by including the baseline value of the outcome variable as a covariate.

Table 12: Estimated programme impact on IDELA scores

Model	Total Idela Score	Sub-Domain score: Emergent Numeracy	Sub-Domain score: Social and Emotional Learning	Sub-Domain score: Gross and Fine motor skills	Sub-Domain score: Emergent Literacy and Language	Sample Size (T0 = Baseline, T1=Endline)
Model (i)	-0.047 (0.679)	0.160 (0.187)	**-0.219*** (0.089)	0.003 (0.979)	-0.084 (0.500)	T0: 346, T1: 422
Model (ii)	-0.067 (0.642)	0.127 (0.387)	-0.202 (0.165)	-0.007 (0.960)	-0.115 (0.452)	T0: 388, T1: 376
Model (iii)	-0.052 (0.414)	**0.153**** (0.049)	**-0.224***** (0.002)	0.022 (0.669)	-0.114 (0.153)	T0: 346, T1: 346
* - significant at 90% level ** - significant at 95% level *** - significant at 99% level						

Table 13: : Estimated programme impact on secondary outcomes

Model	HEQ index	CPR index	MWB index	Handwashing index	Sample Size (T0 = Baseline, T1=Endline)
Model (i)	-0.086 (0.540)	-0.063 (0.665)	-0.172 (0.230)	**0.243*** (0.093)	T0: 346, T1: 422
Model (ii)	**0.278*** (0.061)	0.099 (0.479)	0.056 (0.697)	**0.247*** (0.092)	T0: 388, T1: 376
Model (iii)	-0.035 (0.730)	-0.125 (0.227)	-0.111 (0.112)	0.283 (0.223)	T0: 346, T1: 346
* - significant at 90% level ** - significant at 95% level *** - significant at 99% level					

Table 14: Sub-group Analysis | Estimated programme impact on primary outcomes by child gender

Subgroup	Model	Total Idela Score	Sub-Domain score: Emergent Numeracy	Sub-Domain score: Social and Emotional Learning	Sub-Domain score: Gross and Fine motor skills	Sub-Domain score: Emergent Literacy and Language	Sample Size (T0 = Baseline, T1=Endline)
Male children	Model (i)	-0.240 (0.170)	-0.137 (0.464)	**-0.375*** (0.057)	-0.019 (0.918)	-0.244 (0.175)	T0: 150, T1: 186
	Model (ii)	-0.187 (0.399)	-0.050 (0.826)	-0.320 (0.153)	-0.054 (0.797)	-0.201 (0.369)	T0: 167, T1: 164
	Model (iii)	**-0.219*** (0.098)	-0.109 (0.157)	**-0.365*** (0.056)	0.008 (0.948)	**-0.238*** (0.005)	T0: 150, T1: 150
Female children	Model (i)	0.096 (0.511)	**0.382** (0.016)	-0.101 (0.552)	0.018 (0.903)	0.029 (0.865)	T0: 196, T1: 236
	Model (ii)	-0.005 (0.981)	0.194 (0.315)	-0.144 (0.450)	-0.037 (0.841)	-0.037 (0.861)	T0: 221, T1: 212
	Model (iii)	0.074 (0.127)	**0.354** (0.012)	**-0.118** (0.016)	0.032 (0.311)	-0.025 (0.692)	T0: 196, T1: 196

* - significant at 90% level   ** - significant at 95% level   *** - significant at 99% level

Table 15: Sub-group Analysis | Estimated programme impact on secondary outcomes by child gender

Subgroup	Model	HEQ index	CPR index	MWB index	Handwashing index	Sample Size (T0 = Baseline, T1=Endline)
Male children	Model (i)	-0.159 (0.456)	-0.230 (0.322)	**-0.353*** (0.082)	0.067 (0.766)	T0: 150, T1: 186
	Model (ii)	**0.374*** (0.096)	-0.143 (0.524)	0.198 (0.340)	0.291 (0.202)	T0: 167, T1: 164

	Model					
	Model (iii)	-0.121 (0.665)	-0.329 (0.239)	-0.259 (0.189)	0.157 (0.574)	T0: 150, T1: 150
Female children	Model (i)	-0.028 (0.880)	0.044 (0.810)	-0.039 (0.844)	**0.381**** (0.045)	T0: 196, T1: 236
	Model (ii)	0.200 (0.314)	**0.302*** (0.091)	0.147 (0.458)	0.295 (0.130)	T0: 221, T1: 212
	Model (iii)	0.029 (0.404)	0.014 (0.932)	0.001 (0.988)	0.384 (0.300)	T0: 196, T1: 196
* - significant at 90% level ** - significant at 95% level *** - significant at 99% level						

Table 16: Sub-group Analysis | Estimated programme impact on primary outcomes by child age

Subgroup	Model	Total Idela Score	Sub-Domain score: Emergent Numeracy	Sub-Domain score: Social and Emotional Learning	Sub-Domain score: Gross and Fine motor skills	Sub-Domain score: Emergent Literacy and Language	Sample Size (T0 = Baseline, T1=Endline)
Younger children (aged 3-5 at EL)	Model (i)	-0.167 (0.373)	0.055 (0.785)	-0.334 (0.104)	-0.179 (0.457)	-0.072 (0.660)	T0: 108, T1: 162
	Model (ii)	**-0.400*** (0.062)	-0.163 (0.470)	**-0.402*** (0.058)	**-0.461*** (0.068)	-0.267 (0.155)	T0: 134, T1: 128
	Model (iii)	**-0.221**** (0.042)	0.090 (0.288)	**-0.367***** (0.000)	**-0.252*** (0.053)	**-0.185**** (0.047)	T0: 108, T1: 108
Older children (aged 6-7 at EL)	Model (i)	0.020 (0.886)	0.213 (0.162)	-0.152 (0.357)	0.105 (0.395)	-0.092 (0.582)	T0: 238, T1: 260
	Model (ii)	0.058 (0.698)	0.211 (0.188)	-0.130 (0.435)	0.153 (0.247)	-0.040 (0.826)	T0: 254, T1: 248
	Model (iii)	0.016 (0.738)	**0.169**** (0.027)	**-0.156**** (0.027)	**0.144***** (0.000)	-0.105 (0.245)	T0: 238, T1: 238
* - significant at 90% level ** - significant at 95% level *** - significant at 99% level							

Table 17: Sub-group Analysis | Estimated programme impact on secondary outcomes by child age

Subgroup	Model	HEQ index	CPR index	MWB index	Handwashing index	Sample Size (T0 = Baseline, T1=Endline)
Younger children (aged 3-5 at EL)	Model (i)	-0.394 (0.105)	-0.254 (0.325)	-0.240 (0.327)	0.198 (0.444)	T0: 108, T1: 162
	Model (ii)	-0.029 (0.914)	-0.027 (0.911)	-0.052 (0.833)	0.091 (0.711)	T0: 134, T1: 128
	Model (iii)	**-0.320*** (0.093)	**-0.299***** (0.000)	-0.039 (0.807)	0.212 (0.190)	T0: 108, T1: 108
Older children (aged 6-7 at EL)	Model (i)	0.048 (0.783)	0.016 (0.926)	-0.118 (0.509)	0.283 (0.111)	T0: 238, T1: 260
	Model (ii)	**0.325*** (0.073)	0.156 (0.363)	0.083 (0.641)	**0.316*** (0.085)	T0: 254, T1: 248
	Model (iii)	0.085 (0.682)	-0.049 (0.776)	-0.115 (0.374)	0.327 (0.311)	T0: 238, T1: 238
* - significant at 90% level		** - significant at 95% level		*** - significant at 99% level		

Table 18: Sub-group Analysis | Estimated programme impact on primary outcomes by household wealth

Subgroup	Model	Total Idela Score	Sub-Domain score: Emergent Numeracy	Sub-Domain score: Social and Emotional Learning	Sub-Domain score: Gross and Fine motor skills	Sub-Domain score: Emergent Literacy and Language	Sample Size (T0 = Baseline, T1=Endline)
Wealthier Households	Model (i)	-0.057 (0.735)	0.096 (0.572)	-0.068 (0.720)	-0.025 (0.877)	-0.178 (0.356)	T0: 166, T1: 197
	Model (ii)	-0.049 (0.821)	0.116 (0.587)	0.002 (0.994)	-0.037 (0.854)	-0.198 (0.401)	T0: 208, T1: 178

(upper 50% of PPI distribution)	Model (iii)	-0.079 (0.462)	0.051 (0.701)	-0.047 (0.687)	-0.042 (0.686)	**-0.218**** (0.011)	T0: 166, T1: 166
Less Wealthy Households (lower 50% of PPI distribution)	Model (i)	-0.050 (0.740)	0.207 (0.229)	**-0.372**** (0.033)	0.020 (0.900)	-0.013 (0.935)	T0: 180, T1: 225
	Model (ii)	-0.120 (0.543)	0.109 (0.598)	**-0.466**** (0.023)	0.008 (0.967)	-0.020 (0.918)	T0: 180, T1: 198
	Model (iii)	-0.033 (0.627)	**0.249***** (0.000)	**-0.407**** (0.014)	**0.083**** (0.039)	-0.026 (0.672)	T0: 180, T1: 180

* - significant at 90% level   ** - significant at 95% level   *** - significant at 99% level

Table 19: Sub-group Analysis | Estimated programme impact on secondary outcomes by household wealth

Subgroup	Model	HEQ index	CPR index	MWB index	Handwashing index	Sample Size (T0 = Baseline, T1=Endline)
Wealthier Households (upper 50% of PPI distribution)	Model (i)	-0.159 (0.460)	-0.184 (0.359)	-0.048 (0.817)	0.237 (0.261)	T0: 166, T1: 197
	Model (ii)	0.276 (0.236)	0.126 (0.512)	0.026 (0.900)	0.192 (0.371)	T0: 208, T1: 178
	Model (iii)	-0.053 (0.585)	**-0.265*** (0.078)	-0.037 (0.859)	0.232 (0.279)	T0: 166, T1: 166
Less Wealthy Households (lower 50% of PPI distribution)	Model (i)	-0.016 (0.931)	0.030 (0.886)	**-0.331*** (0.096)	0.258 (0.198)	T0: 180, T1: 225
	Model (ii)	0.215 (0.280)	0.131 (0.528)	0.181 (0.363)	**0.404**** (0.043)	T0: 180, T1: 198
	Model (iii)	-0.015 (0.941)	-0.021 (0.918)	**-0.232***** (0.001)	0.331 (0.244)	T0: 180, T1: 180

* - significant at 90% level   ** - significant at 95% level   *** - significant at 99% level

Independent Evaluation of the Improving Early Childhood Care and Education in rural Uganda programme // Management Response

Response to evaluation findings

Quantitative results

Unfortunately, the quantitative aspect of the evaluation failed to produce a reliable assessment of the impact of the programme. We are in agreement with the evaluation team that this was due to methodological challenges, as outlined in the Executive Summary (p.4) of the report. These include: imbalance on key outcomes at baseline between treatment and control, differential attrition, challenges with statistical matching; and the lack of robust point estimates. Whilst it is disappointing that the evaluation experienced these challenges, we don't believe that the evaluation indicates a lack of impact by the programme itself.

Qualitative results

The results from the qualitative aspect of the evaluation were not affected by methodological challenges and therefore we believe these to be an accurate representation of the programme's impact. These results also align with the anecdotal evidence reported by LM technical team.

The results obtained through the qualitative research suggest that the Mayuge Lively Minds programme is achieving objectives regarding child development, shift in parents' mindsets (in both mothers and fathers), improving familial relationships and an increasing the focus on child health and development across the community.

Results from the interviews with implementers also suggest that programme is bringing benefits for implementing VHTs, including personal growth, improved communication skills, and increased respect from local leaders.

Programmatic lessons learned

Whilst the findings from the qualitative research were broadly positive, individual transcripts made reference to issues such as requests for payment by mothers and VHTs,

and to children dropping out of the programme. These issues have been discussed with the Lively Minds Uganda evaluation team to understand how widespread they are and to develop appropriate solutions. For example:

- Children are unsupervised and demonstrate poor behaviour, e.g. fighting each other and damaging the Play Scheme games: subcounty officials and VHTs will ensure that there is proper supervision of children during the Play Scheme, and mothers will be taught about the risks of leaving children unsupervised.
- Mothers believe that VHTs are being paid to deliver the programme and that they are also entitled to payment or compensation for delivering Play Schemes: subcounty officials and VHTs to provide a refresher to mothers on the Mayuge Lively Minds progrmme, and to clarify that VHTs are not paid but instead receive transport reimbursements.
- Mothers are bringing underage or overage children to the Play Scheme: VHTs to verify the ages of children attending Play Schemes and remind mothers of the intended age range of the programme. If older or younger children must accompany their mothers, they should engage in the outdoor games only.
- Mothers are not listening to the radio programme often enough, or are not practicing what they have learned from it: VHTs should promote the radio programme to the mothers, reminding them when it is aired and what they can learn from it. Lively Minds to investigate airing via the community tower megaphone.

Beyond this, no significant changes will be made to the programme as the quantitative results were not reliable, and qualitative results suggest that programme is already having intended impact.

Lessons learned for future evaluations

Key lessons learned for the inception and management of future evaluations include:

- Ensure that evaluation team are fully onboarded and have excellent understanding of the programme (e.g. PS visits, reading list). This includes practices around community entry
- Take a more considered approach to selecting the evaluation methodology – in some cases a fully qualitative approach may be more productive, or a full-scale RCT may be the only way to achieve the evidence required
- Check sample balance tests after baseline analysis to ensure an early warning for any challenges in sampling that may affect overall results
- Balance need to evaluate around implementation and within budget with need to include large enough sample size, enough clusters, correct level of randomisation
- Agree detailed structure of report and tools (especially. qualitative tools) with evalution team prior to development
- Consider whether qualitative research may be more appropriate than quantitative if we anticipate challenges such as poor sample balance or insufficient power.