



THE AGA KHAN UNIVERSITY

eCommons@AKU

Department of Paediatrics and Child Health

Division of Woman and Child Health

8-17-2021

Evaluating implementation of LEAPS, a youth-led early childhood care and education intervention in rural Pakistan: Protocol for a stepped wedge cluster-randomized trial

Aisha Khizar Yousafzai

Christopher R. Sudfeld

Emily E. Franchett

Saima Siyal

Karima Rehmani

See next page for additional authors

Follow this and additional works at: https://ecommons.aku.edu/pakistan_fhs_mc_women_childhealth_paediatr

Authors


Aisha Khizar Yousafzai, Christopher R. Sudfeld, Emily E. Franchett, Saima Siyal, Karima Rehmani, Shelina Bhamani, Quanyi Dai, Chin R. Reyes, Günther Fink, and Liliana A. Ponguta

STUDY PROTOCOL

Open Access



Evaluating implementation of LEAPS, a youth-led early childhood care and education intervention in rural Pakistan: protocol for a stepped wedge cluster-randomized trial

Aisha K. Yousafzai^{1*} , Christopher R. Sudfeld¹, Emily E. Franchett¹, Saima Siyal², Karima Rehmani¹, Shelina Bhamani², Quanyi Dai¹, Chin R. Reyes³, Günther Fink⁴ and Liliana A. Ponguta³

Abstract

Background: The Sustainable Development Goals (SDGs) highlight the importance of investments in early childhood care and education (ECCE) and youth development. Given Pakistan's large young population, and gender and urban-rural inequalities in access to education, training, and employment, such investments offer opportunities. LEAPS is a youth-led ECCE program that trains female youth, 18–24 years, as Community Youth Leaders (CYLs) to deliver high-quality ECCE for children, 3.5–5.5 years, in rural Sindh, Pakistan.

Methods: We use a stepped wedge cluster-randomized trial to evaluate implementation of LEAPS. Ninety-nine clusters will be randomized to receive the intervention in one of three 7-month steps (33 clusters/step). The primary outcome is children's school readiness (indexed by the total score on the International Development and Early Learning Assessment (IDELA)). Secondary child outcomes are children's IDELA domain scores and executive functions. Data are collected in cross-sectional surveys of 1089 children (11 children/cluster from 99 clusters) aged 4.5–5.5 years at four timepoints (baseline and at the end of each step). Additionally, we will enroll three non-randomized youth participant open cohorts, one per step (33 CYLs: 66 comparison youth per cohort; 99:198 in total). Youth cohorts will be assessed at enrollment and every 7 months thereafter to measure secondary outcomes of youth personal and professional development, depressive symptoms, and executive functions. A non-randomized school cohort of 330 LEAPS students (10 students/cluster from 33 clusters) will also be enrolled and assessed during Step 1 after intervention rollout and at endline. The quality of the learning environment will be assessed in each LEAPS ECCE center and in a comparison center at two timepoints midway following rollout and at endline. A concurrent mixed-methods implementation evaluation will assess program fidelity and quality, and the extent to which a technical support strategy is successful in strengthening systems for program expansion. A cost evaluation will assess cost per beneficiary. Data collection for implementation and cost evaluations will occur in Step 3.

* Correspondence: ayousafzai@hsph.harvard.edu

¹Harvard T. H. Chan School of Public Health, Boston, MA, USA

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Discussion: Youth-led models for ECCE offer a promising approach to support young children and youth. This study will contribute to the evidence as a means to promote sustainable human development across multiple SDG targets.

Trial registration: [ClinicalTrials.gov NCT03764436](https://clinicaltrials.gov/ct2/show/study/NCT03764436). Registered on December 5, 2018.

Keywords: Early childhood care and education, Youth development, Vocational training, Workforce development, School readiness, Pakistan

Background

With the ratification of the Sustainable Development Goals (SDGs), the global community embraced a life-course approach to promoting equitable human development for all ages [1]. Central to this agenda is the investment in opportunities for lifelong learning and development [2] including Early Childhood Care and Education (ECCE) to bolster early child development and learning outcomes (Target 4.2); teacher training and workforce development initiatives to enhance quality of educational services (Target 4.c); and youth development programs to promote education, training, and employment for youth (Targets 4.4 and 8.6) [1]. Innovative, integrated policy, and strategies are needed to leverage interconnections across SDG targets in order to promote human development across the lifecourse [3].

An estimated 250 million children under the age of 5 years in low- and middle-income countries (LMICs) are at risk of not meeting their cognitive developmental potential [4]. There is increasing support for ECCE programs as a means to enhance young children's cognitive, social-emotional, and language development and early literacy and numeracy skills in order to better prepare children to enter primary school [5]. A recent meta-analysis found that participation in formal and community-based ECCE programs enhanced cognitive and psychosocial development outcomes among children [6]. Although there has been an increase in ECCE programs, inequitable access and low-quality program implementation remains a concern [7]. Only 17% of children in LMICs have access to ECCE programs, and disparities in access and quality persist by gender, socio-economic status, and urbanicity [8]. Supporting ECCE program quality is critical in order to benefit young children [6]. Investments in ECCE workforce development are essential for maintaining program quality at scale [8], although questions remain regarding how ECCE programming can be successfully integrated within national systems with fidelity and quality [2].

Lack of access to quality ECCE is also linked to increased school dropout and poor achievement at the primary and secondary levels of education [9, 10]. In Pakistan, almost 60.7% of preschool-aged children are

out of school, and 22.8 million primary and secondary school-aged children are out of school [11]. One in six Pakistani children does not transfer from first grade to second grade [12]. With many children being out of school or achieving poor learning outcomes, investments in high-quality equitable ECCE programs are needed for girls and boys.

As youth prepare to transition into the labor force, training and employment opportunities are lacking in many settings. Globally, 500 million youth are unemployed, underemployed, or have insecure jobs [13], and youth are nearly three times more likely to be unemployed than adults [14]. Approximately 30% of youth aged 15 to 29 years in LMICs are not enrolled in education, employment, or training opportunities [15]. These opportunity deficits for youth can in turn lead to longer-term consequences, as youth who are not in employment, education, or training upon finishing school are more likely to experience job instability and unemployment and have lower earnings as adults [16]. In countries such as Pakistan where 64% of the population is under the age of 29 years [17], programs providing opportunities for youth to pursue education and employment are in high demand. Positive Youth Development (PYD) interventions, defined as interventions that seek to “build skills, assets, and competencies; foster youth agency; build healthy relationships; strengthen the environment; and transform systems” [18], have emerged as a strategy to support youth during this critical phase. While PYD interventions have been well-studied in high-income countries and found to significantly reduce risky behaviors and improve youth mental health, physical health, economic outcomes, and general wellbeing, few youth development programs have been implemented and rigorously evaluated in LMICs [18, 19]. A recent systematic review identified just 94 PYD programs targeting youth in LMICs implemented between 1990 and 2016, of which only 37% of programs had been rigorously evaluated [18]. Results indicated that outcome measures were diverse, ranging from measures of impact mediators, to youth behaviors, to youth development [18]. Further research is needed to evaluate the effectiveness of youth development interventions in LMICs and

to better understand the mediating effects and mechanisms of youth development interventions, particularly among more vulnerable populations such as female youth.

From a life-cycle perspective, investing in young children and youth is critical as both the early childhood and adolescent years are sensitive periods of brain development. The early years of a child's life are the first window of opportunity to positively contribute to children's development, establishing a strong foundation for lifelong health and learning [20]. Late adolescence and youth is a second period when the high plasticity of the brain leads to significant intellectual and social-emotional maturity, referred to as the second window of opportunity [21, 22]. During both of these developmental periods, executive functioning skills are undergoing a phase of rapid development. Executive functions are the skills that support emotional regulation, social skills, attention, and reasoning skills and are essential to children's and youth's successful performance in school and work and in fostering healthy relationships [23, 24]. As children's executive functioning skills are thought to develop through interactions with adults possessing strong executive functions [25], supporting executive function development among caregivers may bolster gains in these skills among children [26]. Engaging youth as ECCE providers offers an innovative solution to enhance executive functioning skills and increase opportunities for both young children and youth. A recently published conceptual model for youth-led ECCE outlines a strategy to promote both child and youth development, by providing appropriate training and mentoring for youth to enable them to deliver high-quality ECCE programming [27]. While examples of youth-led ECCE programs are limited [27], a few key examples can be found in LMICs, including the Gandhi Fellowship program in western India [28] and the Early Childhood Development (ECD)-practitioner training track of the Economic Empowerment of Adolescent Girls and Young Women (EPAG) project in Liberia [29, 30]. Youth-led ECCE offers a promising model for implementing a lifecourse approach to human development.

Evidence for Youth-led ECCE from Pakistan

Youth Leaders for Early Childhood Assuring Children are Prepared for School (LEAPS) is a youth-led ECCE program which was conceptualized as a response to the education and training needs of children and youth in rural Sindh, Pakistan. LEAPS was developed following formative research to identify local barriers and enablers of ECCE and PYD interventions. LEAPS trains female youth aged 18–24 years with a minimum 10th grade education as Community Youth Leaders (CYLs), who serve as local advocates for ECD and deliver quality

ECCE programming for children aged 3.5–5.5 years in community-based preschools. A pilot of LEAPS was implemented in partnership with the Government of Pakistan's National Commission for Human Development (NCHD) in 2016 in district Naushahro Feroze [26]. As proof-of-concept, LEAPS was evaluated through a small-scale cluster-randomized controlled efficacy trial. Five villages were randomized per arm, and a total of 240 children aged 3.5–5.5 years were enrolled in the pilot study (170 children per study arm). Intervention clusters received the LEAPS program, and control clusters received education services as usual. The results showed significant improvements in children's school readiness as assessed by the International Development and Early Learning Assessment (IDELA; [31]) (Cohen's $d = 0.3$) [26]. A qualitative analysis of CYL exit interviews also indicated improved professional and personal development benefits for female youth leaders including aspirations for education and career, mental health benefits, and higher self-confidence [32, 33].

Objectives

The present study has five objectives:

- Objective 1: To evaluate the impact of the larger-scale implementation of the LEAPS model on children's school readiness and executive functioning;
- Objective 2: To evaluate the impact of the larger-scale implementation of the LEAPS model on female youth development;
- Objective 3: To evaluate the larger-scale implementation of the LEAPS model as integrated in NCHD platforms and delivered through NCHD systems in order to assess fidelity, demand, enablers, and barriers to implementation and quality;
- Objective 4: To evaluate the readiness of NCHD to uptake LEAPS for sustainable scaling and replication in other districts and provinces in Pakistan, with respect to leadership and governance, training, supervision, advocacy and communication, monitoring and evaluation, and financing and resource mobilization; and
- Objective 5: To examine cost-per-beneficiary ratios through a cost evaluation.

It is hypothesized that after implementation of LEAPS, clusters will experience significantly improved population-level outcomes for school readiness among children aged 4.5–5.5 years as assessed by the IDELA, as well as improvements in children's executive functioning skills. We also hypothesize that female youth who participate in LEAPS will have improved outcomes for executive functioning, and personal and professional

development relative to youth who are not LEAPS participants.

Methods

Methods are described according to the SPIRIT Guidelines [34].

Study setting

The current trial will be conducted in rural Sindh, Pakistan, across four districts: Naushahro Feroze, Khairpur, Sukkur, and Dadu. The ECCE participation in this region is low. A recent report indicated that 54% of children aged 3 to 5 years in rural Sindh were not enrolled in any ECCE program [11]. The lack of school readiness is also reflected in child learning levels in Grade 1 with less than half of children surveyed able to identify Sindhi letters or English letters, or to recognize numbers one to 99. Low literacy rates among parents present additional challenges with three quarters of mothers and over half of fathers without primary school completion (Grades 1–5), which can limit parents' support for their child's at-home learning [11] (see Table 1 in Additional file 1). Communities in rural Sindh also face significant gender disparities in school enrollment and literacy rates among youth: the literacy rate for male youth (61%) is more than twice that of female youth (30%) [35]. Trends in Pakistan more broadly indicate that female youth are also at a significant disadvantage in employment [36] and thus require more targeted support in education, training, and work (see Table 2 in Additional file 1).

Implementation partner and LEAPS intervention support team

Pakistan's National Commission for Human Development (NCHD) was identified as the partner for program design and implementation during the formative research phase. The NCHD is an autonomous federal body established under the office of the President in 2002 mandated to support and augment human development efforts in Pakistan. NCHD works in conjunction with the government's health and education sectors and with local communities to fill service implementation gaps. Currently, NCHD is present in 124 districts of Pakistan [37]. NCHD's services include provision of primary education for children, vocational training programs for youth, and health services for communities in the most remote areas of Pakistan. The NCHD's Universal Primary Education (UPE) program trains teachers who are then hired to teach in "Feeder" primary schools (Grades 1–5), located in remote areas where government schools are not operating. The NCHD has also previously provided vocational training for young women, including tailoring courses and the provision of sewing machines.

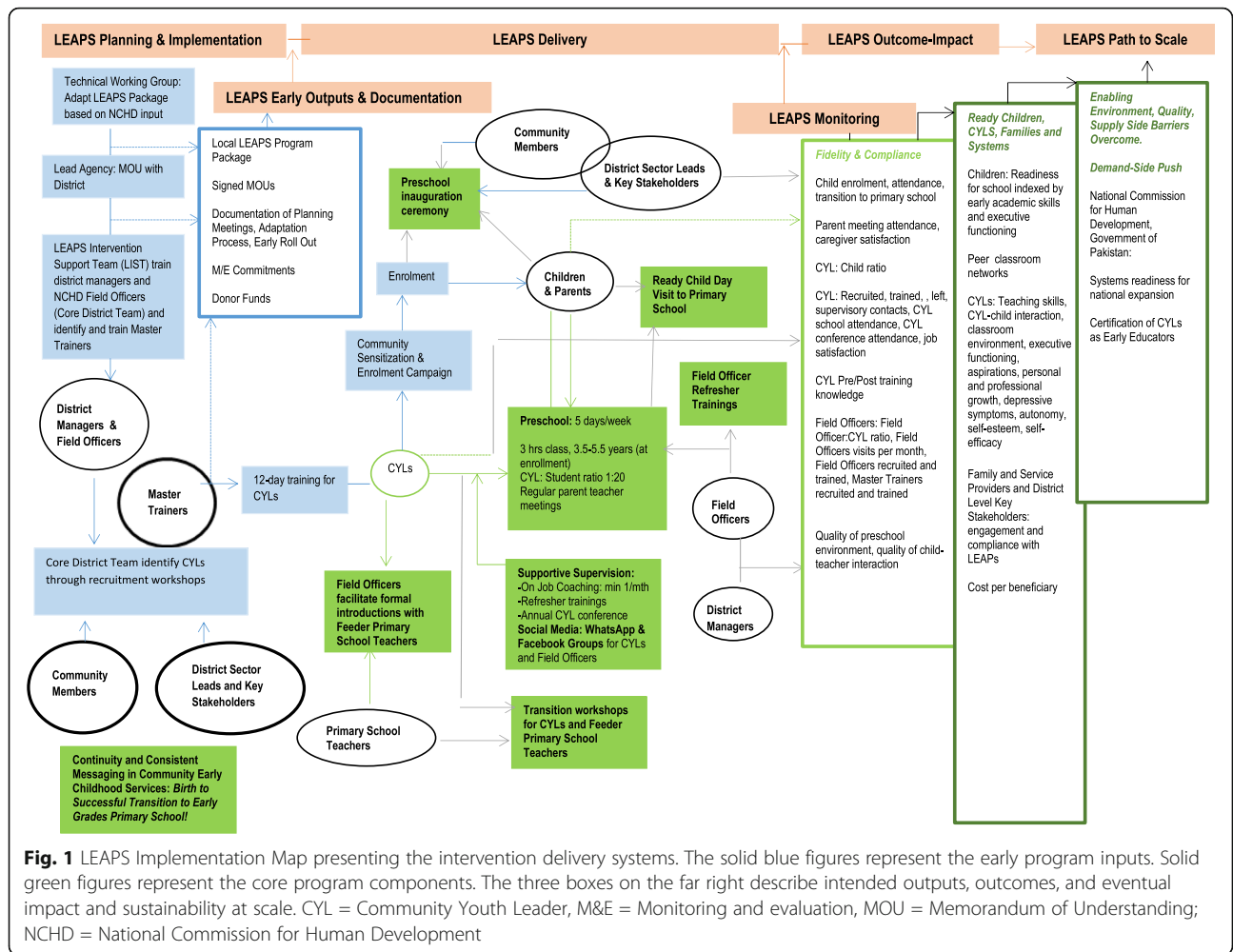
The current study engages NCHD leadership at the Federal Office, Sindh Province Office, and four district offices, as well as the district-level NCHD Field Officers (FOs), who will be responsible for supervising the CYLs in their catchment area. All FOs have a minimum bachelor's level education and several years of experience working with NCHD's Feeder primary school teachers and local community. Implementation of LEAPS activities will be led by NCHD, with technical support from the LEAPS Intervention Support Team (LIST), a five-member female team, all of whom are local to the implementation communities, with a Master's level education and several years of experience delivering community interventions including ECCE. The purpose of the LIST group is to provide training and technical support for NCHD to promote the sustainable rollout and integration of LEAPS within NCHD's UPE program. All LIST members will receive a 3-month training prior to rollout of LEAPS, followed by a 3-day refresher training and ongoing on-the-job mentorship and support in coaching, supervision, and coordination from the LEAPS research team.

Description of the LEAPS intervention

The details of the intervention are reported in accordance with the TiDiERS guidelines [38]. The Implementation Map (Fig. 1) details the program delivery systems and anticipated outputs, outcomes, and impact. A summary of the dosage, location, modes of delivery, and monitoring tools for key LEAPS intervention activities is presented in Table 3 of Additional file 1.

Training of NCHD field officers and master trainers

An initial 5-day training of trainers (TOT) will be held for the 42 FOs and eight district managers who will be engaged in LEAPS. The TOT will be implemented by the LIST members using a training manual. The training introduces basic concepts of ECCE, familiarizes participants with the LEAPS program model, and targets essential supervision and coaching skills, to prepare FOs for their role as CYL supervisors. The LIST members will use an observation checklist to assess FOs' performance during the training. After the training, LIST members and district managers will jointly select the ten highest-performing FOs to be the LEAPS Master Trainers (MTs). In addition to their supervision responsibilities, the MTs will lead all CYL training activities in their district. The MTs will receive training per diems aligned with the government per diem rates for their employee grade level. LIST members will provide a 1-day preparatory training for MTs prior to each CYL training, along with coaching and feedback during training sessions. LIST members will additionally provide ongoing support for all FOs during weekly group



coordination meetings at district offices, periodic group refresher trainings providing targeted support for specific skills, and monthly one-on-one field visits to provide on-the-job coaching during CYL supervision.

CYL recruitment

Following the TOT, district managers and FOs will initiate the CYL recruitment process in their catchment areas. NCHD will engage Feeder school teachers, community health workers, village leaders, and other community members to inquire about potential female youth candidates meeting the CYL recruitment criteria (i.e., aged 18 to 24 years old, have completed Grade 10 schooling, have their family’s support to participate). Candidates will be invited to attend a local CYL recruitment workshop. Recruitment workshops will be led by district managers and FOs, with support from LIST members, and comprise a range of activities designed to assess candidates’ literacy, numeracy, general knowledge, attitudes toward working with young children, problem solving, creativity, communication, and leadership

potential. Youth participants’ performance on workshop activities will be scored using a standardized rubric. Participants who meet the CYL recruitment criteria and have the highest scores will be selected as the CYL.

CYL training

Selected CYLs will be invited to attend their district’s 12-day teacher training. Training will introduce basic concepts of early childhood development and ECCE, orient CYLs to the LEAPS curriculum, and provide hands-on opportunities to practice activities from the curriculum (see Table 4 in Additional file 1 for training overview). We intend to modify consecutive trainings based on lessons learned from early training activities. The CYLs will continue to receive on-the-job vocational training through supportive supervision visits from FOs and LIST members, as well as through participation in CYL refresher trainings and other youth development activities, including an annual CYL conference. Once hired, CYLs will receive a monthly salary of Rs. 8000 (approximately \$50 USD), which aligns with the NCHD

salary guidelines for teachers. CYLs will additionally receive travel support for all training days and professional development activities.

Preschool space identification and setup

NCHD district offices will be tasked with identifying a space for the LEAPS preschool in each cluster through a community engagement strategy. To be considered viable, a preschool space must have a roof, enclosed walls, access to a washroom, access to a water source, and a door that locks. The preschool space will be an in-kind donation from the local community. The CYLs will each receive a LEAPS classroom starter kit, which includes learning materials, furniture, and first aid and cleaning supplies. The CYLs and community members will collaborate to set up the preschool space, keeping in mind the elements for a high-quality preschool program that are relayed during trainings.

Preschool enrollment drives

After the preschool is set up, CYLs will conduct enrollment drives in the village. Children will be eligible for enrolment in the preschool program if they are aged 3.5–5 years and reside locally. CYLs will be coached to ensure equal opportunities for both girls and boys to enroll in LEAPS. Target enrollment will be 20 children per classroom. CYLs will place any additional children on a waiting list.

Inauguration ceremony

Prior to opening the LEAPS preschool, the FO and CYL will plan an inauguration ceremony to celebrate with parents of LEAPS children, local leaders, and the wider community. Inauguration ceremonies offer an essential opportunity to further engage with the community and raise awareness for LEAPs and the importance of ECCE. Following the inauguration ceremonies, preschool classes will officially commence.

Preschool classes and curriculum

LEAPS preschools will follow the NCHD UPE program academic calendar and 6-day school week. LEAPS school days are 3 h long and follow a structured classroom routine (see Table 5 in Additional file 1). The LEAPS curriculum was developed through formative research, drawing on the principles of the HighScope preschool curriculum and aligned with the Government of Pakistan's early childhood education curriculum. It is designed to support the achievement of 50 Key Development Indicators (KDIs) across seven learning areas (presented in Table 4 of Additional file 1). CYLs will record children's attendance daily. CYLs will use a Child Progress Form to track children's progress toward achieving the 50 KDIs. CYLs will be coached to monitor

the overall class performance and to share general strengths and areas for additional support with parents during Parent Teacher Meetings (PTMs). PTMs will provide an opportunity for CYLs to teach parents more about the LEAPS program and their child's learning needs.

Transition from LEAPS to primary school

In alignment with NCHD's Feeder school enrollment criteria, at the end of the academic year, LEAPS children who are aged 5 years and above will transition to the NCHD Feeder school. The NCHD may recommend that children who are between ages 5 and 5.5 years who have not yet achieved their 50 KDIs continue to attend LEAPS to boost their KDI achievement prior to transitioning to primary school at age 5.5 years. A Transition Workshop will be conducted for all CYLs and NCHD Feeder school teachers with the goal of preparing the Feeder school teachers to welcome and support the LEAPS graduates. Additionally, the CYLs will accompany LEAPS children to visit their Feeder school to familiarize students with their new teacher, school environment, and routine.

Supervision visits

In order to ensure CYLs are supported to offer high-quality preschool programming to children, CYLs will receive monthly on-the-job supervision visits from their FO. Visits will last for the entire preschool session. The LIST members will in turn provide ongoing support and coaching to FOs by accompanying them on their preschool supervision visits once a month. The NCHD FOs use a supervision checklist tool during their visits to assess program quality and fidelity, and to provide constructive feedback to the CYL on her strengths and areas for improvement. Similarly, the LIST member uses a separate supervision checklist tool to observe and provide constructive feedback to the FO.

Refresher trainings for CYLs and FOs

Both CYLs and FOs will participate in 1-day professional development workshops or refresher trainings. For CYLs, these will be led by NCHD MTs, and for FOs these will be led by LIST members.

1. Core refresher trainings:
 - a. CYL refresher trainings: A set of core training topics including preschool routine, child progress forms, and preschool behavior management will be offered to CYLs to reinforce key ECCE skills. These are intended to be held quarterly.
 - b. FO refresher trainings: A set of training topics including use of the supervisory checklist tool,

child progress forms, and gender and youth-sensitive supervision are offered to FOs to ensure high-quality preschool supervision visits. These are intended to be held bi-annually.

- Needs-based training: Additional professional development workshops may be held to address observed needs of CYLs and FOs identified by NCHD and LIST members.

Annual CYL conference

An annual conference will be held for all CYLs in one of the four targeted districts. The conference will provide CYLs with a platform to meet their peers, share ideas, discuss career-related challenges and ideas, and brainstorm solutions. The purpose of the conference is three-fold: (i) to promote female youth agency; (ii) to support female youth personal and professional development; and (iii) to showcase the CYLs’ work.

LEAPS response during emergencies and conflict

In the past, there have been cases of political unrest and major flooding in Pakistan, which has led schools to shut down for weeks [39]. Table 6 in Additional file 1 details the actions that LEAPS will undertake to support CYLs and LEAPS children in the event of emergency response.

Study design

This transition-to-scale trial uses a stepped wedge cluster-randomized design with three steps to evaluate the impact of the LEAPS program on children’s school readiness and female youth development in 99 clusters. A cluster-randomized approach was selected as the point of intervention will be at the community level; therefore, this approach minimizes the potential for contamination between intervention and non-intervention recipients. A stepped wedge approach was selected as a pragmatic approach to support program implementation on a larger scale [40]. The LEAPS program will be introduced in clusters in a randomized sequence at three discrete timepoints over the course of the 31-month trial. At each of the three timepoints, 33 clusters will cross over from the control condition (i.e., NCHD services as usual)

to the intervention condition (i.e., initiating the LEAPS program) (see Fig. 2). Once a cluster crosses over it will remain exposed (i.e., the program will remain in operation) for the duration of the trial. Each step will be 7 months in duration. In Sindh, the academic year runs from April to March, with longer holidays falling in May and June, and an extended enrolment period that runs from April to September. Our initial rollout is planned to align with this academic cycle. Prior to opening the LEAPS schools in each step, targeted clusters will undergo a transition period in which NCHD, with technical assistance from the LIST members, will initiate CYL recruitment and trainings, preschool space identification, and community sensitization activities. No concomitant care or interventions are prohibited during the trial due to the pragmatic design of the trial.

Cluster identification and randomization

A cluster is defined as the catchment area of an NCHD Feeder primary school. We restricted the study clusters to NCHD Feeder primary school catchment areas because the possibility to transition from ECCE to primary school was considered essential. The research team coordinated with the NCHD to compile a list of all villages across the four districts with NCHD Feeder schools ($n = 119$), excluding the villages from the LEAPS pilot trial. With the NCHD, the research team screened potential clusters against predefined exclusion criteria, including (i) lack of an eligible CYL candidate, (ii) safety and security concerns, (iii) lack of proper space for the preschool, or (iv) NCHD Feeder school had closed. Twenty clusters were excluded through this process. We applied stratified randomization to the 99 clusters using the 4 districts as strata with 3 blocks in each stratum. The 99 clusters were randomly assigned in blocks to one of the three steps ($n = 33$ clusters per step). The stratified randomization was completed using a computer-generated random allocation sequence in Excel. Block randomization was used to ensure that clusters were evenly dispersed by geography across each step, so that intervention rollout was feasible for implementation partners. The co-investigator and project Senior

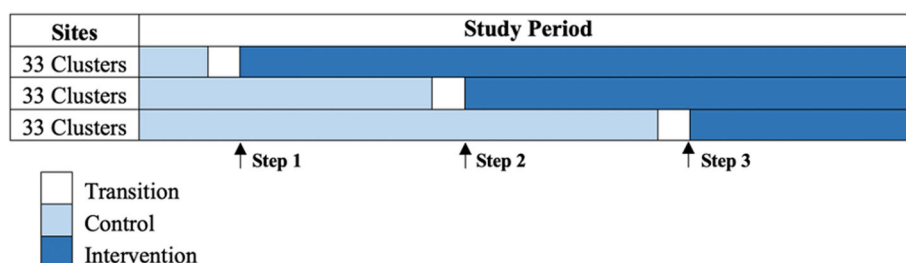


Fig. 2 Stepped wedge cluster-randomized trial timeline

Statistician (CS) who conducted the randomization is not involved in any of the field activities. The final trial includes six clusters in Dadu, nine clusters in Sukkur, 28 clusters in Khairpur, and 56 clusters in Naushahro Feroze (Fig. 3).

Description of data collection activities & study sample

Data collection activities and the participant timeline are summarized in the SPIRIT figure (Table 1) [34]. The trial will include the following data collection activities:

Population survey

A cross-sectional survey of child-caregiver dyads will be conducted at four timepoints: at baseline (Step 0; mos. 1–3) and again at the end of each consecutive step, with 7 months between data collection rounds. Eleven

children aged 4.5–5.5 years and their caregivers will be enrolled in the population survey from each of the 99 clusters (1089 child-caregiver dyads total), regardless of the child's enrollment in the LEAPS preschool program. The primary objective of the population survey will be to assess changes in the average school-readiness over time. With this trial design, rather than assessing program impact on specific individuals, we are aiming to assess the impact of LEAPS on local (cluster-level) school readiness, which is a function of both the proportion attending LEAPS and the impact on the treated. By including both LEAPS participants and children who are not participating in the LEAPS program, this design will allow us to separate treatment effects on the treated from average (intent-to-treat) effects at the population level. The sample will be drawn from village birth

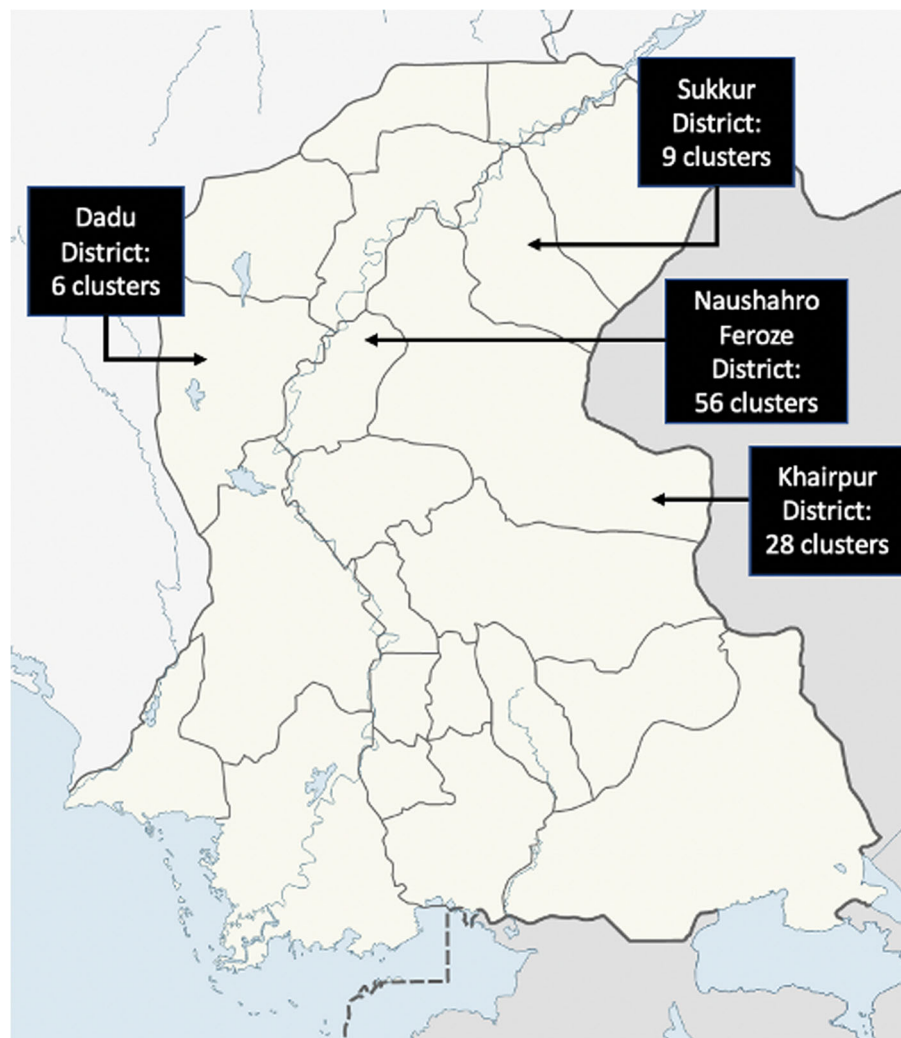


Fig. 3 Map of Districts for the LEAPS Transition-to-Scale Trial in Rural Sindh. Note for Fig. 3: One cluster from district Khairpur was randomized with the district Naushahro Feroze clusters, as it was geographically closer and a multiple of 3 was needed in each group. This cluster fell under the responsibility of the Naushahro Feroze NCHD District Office

Table 1 SPIRIT figure depicting schedule of enrollment, interventions, and assessments. ⁱThe population survey is a cross-sectional survey enrolling 11 children aged 4.5–5.5 years and their primary caregiver per cluster in each of the 99 clusters (total of 1089 child-caregiver dyads per round) conducted at four timepoints. ⁱⁱThe school cohort is a closed cohort enrolling 10 LEAPS preschoolers and their primary caregiver per cluster (stratified by gender) in each of the 33 clusters receiving the intervention during Step 1 (total of 330 child-caregiver dyads). ⁱⁱⁱAn open cohort of youth participants will be enrolled in each of the steps (3 cohorts in total). Each cohort includes all Community Youth Leaders (CYLs) hired by NCHD, as well as two comparison youth participants per cluster (33 CYLs and 66 comparison youth per cohort; 99 CYLs and 198 comparison youth in total). *During the final 2 months of Steps 0, 1, and 2, the 33 clusters that are targeted for intervention rollout in the upcoming step will undergo a transition period to prepare for opening the LEAPS preschools, including recruiting and training CYLs, identifying the preschool space, and setting up the preschool classroom

Timepoints	Study period				
	Step 0 1–3 mo.*	Step 1 4–10 mo.*	Step 2 11–17 mo.*	Step 3 18–24 mo.	Close out 25–31 mo.
Cluster randomization	X				
Allocation	X				
Study phase					
33 clusters	Control*	Intervention	Intervention	Intervention	
33 clusters	Control	Control*	Intervention	Intervention	
33 clusters	Control	Control	Control*	Intervention	
Assessments					
Child, population surveyⁱ					
Demographics	X	X	X	X	
School readiness	X	X	X	X	
Executive functioning skills	X	X	X	X	
Child, school cohortⁱⁱ					
Demographics		X		X	
School readiness		X		X	
Executive functioning skills		X		X	
Youth cohortsⁱⁱⁱ					
Demographics		X	X	X	
Personal and professional development		X	X	X	
Executive functioning skills		X	X	X	
Self-reported depressive symptoms		X	X	X	
Mediators					
Classroom structural and process quality			X	X	
Other					
Implementation/ cost/ system strengthening evaluations				X	

registries provided by NCHD, which will be verified (a) by cross-checking against birth registries from local community health workers; and (b) through household visits conducted by research team members. The sample will be re-drawn at each of the four data collection timepoints. Children may be enrolled for assessment at multiple timepoints. Assessors will use a random number generator to select 11 participants/cluster using the verified lists of eligible children prepared for each cluster. If a participant refuses, assessors will select the next participant in the random sequence.

- Eligibility criteria: Child is age 4.5–5.5 years \pm 2 weeks (i.e., 53.5–66.5 months at the time of the survey) and resides in the targeted village
- Exclusion criteria: Child or caregiver shows signs of severe clinical health condition or disability

School cohort

A closed cohort of 330 LEAPS children and their caregivers (10 children/cluster, stratified by gender) will be enrolled from the 33 clusters receiving the intervention in Step 1. The main purpose of the school cohort is to

assess individual learning and growth trajectories among children participating in the program. Child-caregiver dyads will be enrolled and assessed during Step 1, after the intervention has been initiated in month 8 of the trial and followed up again at the end of Step 3 in month 24. Assessors will use a random number generator to randomly select five girls and five boys from the CYL's enrolment register. If a participant refuses, the next participant in the random sequence will be invited. If fewer than five participants of one gender are registered in the LEAPS class, additional children of the other gender may be enrolled to meet the ten-student sample for that cluster.

- Eligibility criteria: Child must be registered and regularly attending the LEAPS school
- Exclusion criteria: Child or caregiver shows signs of severe clinical health condition or disability

Youth participants

A case:comparison design will be employed in order to examine program impacts on youth participants. In each step, we will enroll an open cohort comprising the 33 CYLs from the 33 clusters introducing the intervention in that step (one CYL/cluster), as well as 66 youth comparisons for a 1:2 CYL: youth comparison ratio (i.e., 3 youth cohorts (one/Step) enrolling 33 CYLs:66 youth comparisons; in total, 99 CYLs and 198 youth comparisons enrolled). In each step, the CYLs will be enrolled in the evaluation and assessed as they are hired by NCHD, during the month before and first month of the step (i.e., in mos. 3–4 for Step 1, in mos. 10–11 for Step 2, and in mos. 17–18 for Step 3). The youth comparisons will be enrolled and assessed beginning the month before through the first 2 months of the step (i.e., in mos. 3–5 for Step 1, in mos. 10–12 for Step 2, and in mos. 17–19 for Step 3). Youth cohort participants will be followed up every 7 months. Given the context, some dropouts are expected, particularly among comparison youth. The sampling ratio of 2 comparison per 1 CYL is intended to compensate for these dropout levels. If a CYL drops out of the LEAPS program after having initiated her in-classroom teaching experience, she will be invited to remain in the evaluation as part of the CYL cohort. If a CYL drops out of the LEAPS program prior to completing her basic training (i.e., prior to having any in-classroom experience), she will be invited to remain in the evaluation as part of the comparison youth cohort. Similarly, if NCHD hires a comparison youth as a CYL, she will be invited to continue in the study as a CYL, and a new comparison youth will be identified and enrolled.

Youth comparisons will be identified using a screening tool during CYL recruitment activities and through

door-to-door household visits. The screening tool for youth comparisons includes a series of questions and short case study exercises to ensure that the participant (i) meets the same gender, age, and education requirements as the CYL, (ii) has support from her family to pursue employment outside of the workplace (as is the case with CYLs), and (iii) demonstrates confidence, and advocacy and problem-solving skills. The tool takes approximately 15 min to implement. These screening criteria represent the qualities that the research team considers most central to the CYL role. Given that CYLs may be positive deviants in their community and have demonstrated high-selection workshop performance relative to their peers, the purpose of the screening tool is to ensure to the best of our ability that youth participants are a fair comparison for CYLs in terms of family environment and skillsets baseline.

- Eligibility criteria:
CYLs must be hired by NCHD.
Youth comparisons must be female, aged 18–24 years \pm 2 weeks, have a minimum tenth grade education, have support from family members to pursue employment, meet capacity-checks for confidence, advocacy, and problem-solving skills, and have no in-classroom teaching experience with LEAPS.
- Exclusion criteria: Youth participant shows signs of severe clinical health condition or disability

Classroom observations

To examine quality of the preschool environment, classroom observations will be conducted in both LEAPS classrooms and comparison preschool classrooms. As LEAPS schools are by design typically the only preschool service operating in an intervention cluster, in order to select the most suitable comparison, the comparison preschool classrooms will be selected based on the following: (i) location: school should be located as close to the LEAPS evaluation cluster as possible; and (ii) prioritizing enrollment of government schools (standard service) where available, followed by NCHD-operated schools and other options (private/Madrasa) to reflect the distribution of these types of schools in rural communities in the targeted districts. Comparison classrooms may vary with respect to the preschool classroom arrangements, including grade-level groupings (stand-alone preschool classrooms or multi-grade classrooms), single- or mixed-gender classroom, and informal or formal settings. One comparison classroom teacher will be enrolled for each LEAPS school. All Step 1 and Step 2 LEAPS schools ($n = 66$) and equivalent number of comparison schools will be observed at two timepoints: once in Step 2 (mos. 14–16) and once during Step 3 (mos.

21–24). All Step 3 LEAPS schools ($n = 33$) and equivalent number of comparison schools will be observed at one timepoint in Step 3 (mos. 21–24).

- Eligibility criteria:
 - LEAPS school: Must be located in an intervention cluster.
 - Comparison preschools: Fits criteria for location, type of school, and gender make-up.
- Exclusion criteria: None

Implementation, costing, and systems strengthening evaluations

A phenomenological design will be employed to collect original qualitative data in addition to administrative and programmatic data, Global Positioning System (GPS) and location data, and costing data to inform the concurrent mixed-methods implementation evaluation, cost evaluation, and systems strengthening evaluation (all three of which are described in detail in the subsections below). Data for these activities will be collected in mos. 22 to 24 at the end of Step 3. Participants in qualitative data collection activities will be purposefully sampled to reflect variation in district, demographics, and stakeholder role. Qualitative data collection will

include focus group discussions (FGDs) and in-depth interviews (IDIs). Administrative, classroom, and programmatic data will be collected from NCHD district offices and around training and supervision activities. Costing data will be collected from internal project budgets (NCHD program and university research partners).

- Eligibility criteria: Sampling and participant details for qualitative data collection are detailed in Table 2
- Exclusion criteria: Participants demonstrate signs of severe clinical health condition or disability

Implementation evaluation Implementation evaluations are critical in transition-to-scale research for a number of reasons: (i) implementation evaluations provide key insights on the challenges of scaling up, in terms of replicability, coverage, appropriate monitoring and evaluation and sustainability [39]; (ii) they help to explain why certain results were achieved, specifically, understanding which components of the intervention contributed to observed impact [41]; (iii) they help to make effectiveness results more interpretable and to avoid type III error, i.e., evaluating an intervention that has not been adequately implemented and therefore drawing incorrect conclusions about an intervention’s

Table 2 Qualitative data collection activities and participants. Table 2 summarizes planned qualitative data collection activities by stakeholder, including sampling details. FGD = Focus Group Discussion, IDI = In-depth Interview

Stakeholder	No. of IDIs/ FGDs	Participant sampling
CYLs	4 FGDs	<ul style="list-style-type: none"> • Representation from all districts and steps • Maximize variation in CYL age, marital status, education, parenting status
CYLs	16–17 IDIs	<ul style="list-style-type: none"> • Representation from all districts and steps • Variation in marital status, parenting status, age, and urbanicity
CYLs’ families	4 Family Interviews	<ul style="list-style-type: none"> • Sub-set families of the same CYLs in IDIs • Variation in marital status and family structure
Children’s mothers	4 FGDs	<ul style="list-style-type: none"> • Representation from all districts • Child must be enrolled in LEAPS preschool • From rural communities
Children’s fathers	4 FGDS	<ul style="list-style-type: none"> • Representation from all districts • Child must be enrolled in LEAPS preschool • From rural communities
NCHD Feeder school teachers	4 FGDs	<ul style="list-style-type: none"> • Representation from all districts and steps
NCHD field officers	4 FGDs	<ul style="list-style-type: none"> • Representation from all districts and Steps • All female field officers will be enrolled • Variation in number of supervisees
NCHD master trainers	1 FGD	<ul style="list-style-type: none"> • Including all field officers who have been selected as master trainers
LIST team	1 FGD	<ul style="list-style-type: none"> • Including the LIST members
LIST members	1 FGD (Pile sorting activity)	<ul style="list-style-type: none"> • Including the LIST members
NCHD district team managers	1–2 FGDs	<ul style="list-style-type: none"> • Including all NCHD district managers
District stakeholders	4–8 IDIs	Sample most relevant/influential stakeholders per district as per Intervention Team’s inputs: <ul style="list-style-type: none"> • Village leaders • District government officials • Social Welfare Department • Education Department

effectiveness [42]; and (iv) they are tools to assess the quality and accuracy of the intervention delivered to participants. This trial's implementation evaluation will capture information relating to facilitators and barriers for participation in and delivery of LEAPS, as well as fidelity to the LEAPS program model and the quality of the ECCE and youth vocational training services delivered. In doing so, we seek to better understand the processes and conditions required to support program expansion, replication, and sustainable scale-up.

Cost evaluation Attention to cost-effectiveness of interventions is a crucial component of intervention planning and evaluation, particularly in resource-limited contexts [41]. The LEAPS cost evaluation will provide information about costs to government providers to inform strategies for program expansion and replication. This evaluation will apply an ingredients approach to assessing program costs, cost per beneficiary, and cost per impact [43].

Systems strengthening evaluation This component of the study aims to evaluate the readiness of NCHD to uptake LEAPS for sustainable scaling and replication for expansion in Sindh and other provinces in Pakistan. Our research team will be facilitating a workshop series for NCHD to provide technical capacity building around organizational capacities identified as central to sustainable, quality implementation of LEAPS. Six "pillars" for systems strengthening were identified based on a review of frameworks for supporting, monitoring, and evaluating organizational capacity development [44–47], transition readiness [48–50], and community-based management training [51, 52], and informed by inputs and feedback from the NCHD provided during a needs assessment conducted by the research team. The LEAPS System Strengthening Pillars include (i) training, (ii) supervision, (iii) monitoring and evaluation, (iv) leadership and governance, (v) advocacy and communication, and (vi) financing and resource mobilization. A set of indicators will be co-created with NCHD and used to assess NCHD's "systems readiness" across each pillar. The systems strengthening evaluation has three aims: (i) to identify barriers and enablers for implementation at scale; (ii) to assess the effectiveness of the capacity development strategy on NCHD's competencies related to the LEAPS System Strengthening Pillars; and (iii) to evaluate efforts of advocacy in NCHD with partners for replication, expansion, and investment (e.g., awareness about ECCE and youth development, enactment of policies, funding commitments made, sustainability plan developed).

Outcomes, mediators, and covariates

The trial's primary and secondary outcomes and covariates are described in Table 3. All measures will be adapted and piloted for the local socio-cultural and language context [53]. The classroom observation tool will be adapted for use in LEAPS from the Measure of Early Learning Environments (MELE) module of the Measuring Early Learning Quality and Outcomes (MELQO) tool [54] and the Instrument for Measuring Quality of Early Childhood Education in Colombia (IMCEIC) [55]. Targeted outcomes and covariates are selected to align with the LEAPS program theory of change (see Fig. 4).

Sample size and power calculation for stepped wedge cluster-randomized trial

A power calculation was conducted with respect to the changes in average child outcomes using representative population surveys, applying the methods as outlined by Hemming and Taljaard [70]. The primary study outcome is children's school readiness as indexed by scores on the International Development and Early Learning Assessment (IDELA). An intraclass correlation coefficient (ICC) of 0.175 was calculated based on the results of the LEAPS efficacy trial [26] using total IDELA score at end-line for the control group. With a fixed number of clusters ($n = 99$) and a step-size of three ($n = 33$ clusters randomized per step), we estimate that at least 8 children per cluster per round, or 792 children total per round, will need to be enrolled to detect a 0.3 SD effect size at 80% power. To achieve adequate power and take refusal rate into account, we will enroll 11 child-caregiver dyads per cluster for a total of 1089 child-caregiver dyads per round, which should detect a 0.3 SD effect size at 90% power with full recruitment. In other words, we estimated sufficient sample sizes to achieve at least 80% power with up to 30% refusal/under-recruitment rate in each round of the data collection under the intent-to-treat principle. This includes refusals from caregivers and children, or under-recruitment due to a decrease in children available at the point of data collection, particularly in smaller villages.

Blinding

Blinding of subjects is not feasible given the nature of the intervention. While the intent is for the evaluation team to be blinded to intervention status, there are limitations to the extent that assessors can be blinded to participants' intervention status, given that assessors live locally and intervention activities are visible in communities. However, team members are taking measures to support blinding where possible, including having separate intervention and evaluation field team members and holding separate team trainings and meetings.

Table 3 Description of trial outcomes, assessment tools, mediators, and covariates. *The starred IDELA, executive functioning skills assessments, anthropometric, and demographic measures were previously adapted for the rural Pakistan context for use in the LEAPS efficacy trial [26]. **The SRQ-20 was previously adapted and validated for use in rural Pakistan [56]. All other measures were adapted or developed for use in this trial. All outcomes will be recorded as absolute values, and cluster-level outcomes will be aggregated by averaging the individual outcomes within each cluster

Variable	Tool	Description
Primary outcome		
School Readiness	International Development & Early Learning Assessment (IDELA) [31]*	IDELA assesses the development and early learning, namely school readiness, of young children aged 3.5–6 years old. The tool includes sub-tests for emergent math, socio-emotional development, emergent literacy, and motor skills. The primary outcome will be the IDELA total score, which will be the parameter estimates computed through a Graded Response Model based on the recommendation of the tool developers. Higher scores reflect better outcomes.
Secondary outcomes		
Emergent Math	International Development & Early Learning Assessment (IDELA) [31]*	IDELA subdomain comprising seven sub-tasks assessing emergent math skills. The domain score is computed through a Graded Response Model recommended by the tool developers. Higher scores reflect better outcomes.
Socio-emotional Development	International Development & Early Learning Assessment (IDELA) [31]*	IDELA subdomain comprising five sub-tasks assessing socio-emotional development. The domain score is computed through a Graded Response Model recommended by the tool developers. Higher scores reflect better outcomes.
Emergent Literacy	International Development & Early Learning Assessment (IDELA) [31]*	IDELA subdomain comprising six sub-tasks assessing emergent literacy skills. The domain score is computed through a Graded Response Model recommended by the tool developers. Higher scores reflect better outcomes.
Motor Skills	International Development & Early Learning Assessment (IDELA) [31]*	IDELA subdomain comprising four sub-tasks assessing fine and gross motor skills. The domain score is computed through a Graded Response Model as recommended by the tool developers. Higher scores reflect better outcomes.
Child Executive Functioning Skills: Working Memory	Corsi Block-Tapping Test [57]*	Direct assessment of child's working memory skills. The child's score reflects the longest sequence for which two block series were repeated correctly, plus 0.5 if one longer series was also correctly processed (range 0–4 points). Higher scores indicate better outcomes.
Child Executive Functioning Skills: Inhibitory Control	Knock and Tap assessment [58]*	Direct assessment of child's inhibitory control skills. Total score is the sum across 16 items (range 0–16 points). Higher scores reflect better outcomes.
	Peg Tap assessment [58]*	Direct assessment of child's inhibitory control skills (range 0–16 points). Total score is the sum across 16 items. Higher scores reflect better outcomes.
Child Executive Functioning Skills: Cognitive Flexibility	Dimensional Change Card Sort [59, 60]*	Direct assessment of child's cognitive flexibility skills. Five sub-scores will be examined: total number correct trials on the separated level, pre- and post-switch (0–6 points each); total number correct trials on the integrated level, pre- and post-switch (0–6 points each); and total correct on the border version (0–12 points). Higher scores reflect better outcomes.
Child General Executive Functioning Skills	Head Toes Knees Shoulders task [61]*	Direct assessment of child's general executive functioning skills. The assessment contains two sections. Child participants may receive 0 to 20 points for each section. Scores will be presented separately for each section. Higher scores reflect better outcomes.
Youth Executive Functioning Skills: Working Memory	Backward Word Span Task [62]*	Direct assessment of youth's working memory skills. The participant's score reflects the longest sequence for which two series were repeated correctly, plus 0.5 if one longer series was also correctly processed (range 0–6 points). Higher scores reflect better outcomes.
	Corsi Block-Tapping Test [63]*	Direct assessment of youth's working memory capacity. Participants will be administered two tests: a forward and a

Table 3 Description of trial outcomes, assessment tools, mediators, and covariates. *The starred IDELA, executive functioning skills assessments, anthropometric, and demographic measures were previously adapted for the rural Pakistan context for use in the LEAPS efficacy trial [26]. **The SRQ-20 was previously adapted and validated for use in rural Pakistan [56]. All other measures were adapted or developed for use in this trial. All outcomes will be recorded as absolute values, and cluster-level outcomes will be aggregated by averaging the individual outcomes within each cluster (*Continued*)

Variable	Tool	Description
		backward test. A score will be awarded for each test, reflecting the longest sequence for which two series were repeated correctly, plus 0.5 if one longer series was also correctly processed (range for each 0–7 points). Scores will be presented separately for each test. Higher scores reflect better outcomes.
Youth Executive Functioning Skills: Inhibitory Control	Number Stroop Effect [64]*	Direct assessment of youth's inhibitory control skills. Points are summed across 20 items to calculate the overall sum score (range 0–20 points). Higher scores reflect better outcomes.
Youth Executive Functioning Skills: Cognitive Flexibility	Trail Making Task (TMT) [65, 66]	Direct assessment of youth's cognitive flexibility skills. Scores reflect the time for completion in seconds, with the maximum time limit being 300 s (range 0–300). Lower scores indicate a better outcome.
Youth Personal and Professional Development	Self-reported Questionnaire	Self-reported questionnaire developed for study context which includes items examining aspects of professional and personal development pertaining to autonomy, self-efficacy, self-esteem, professional growth, and aspirations. Mean item scores will be calculated for each domain. This does not include a standardized scale.
Youth Self-Reported Depressive Symptoms	Self-Reporting Questionnaire 20-Item (SRQ-20) [67]**	Direct assessment of youth self-reported depressive symptoms over the last 30 days. Items are summed for a total score of depressive symptoms over the last 30 days (range 0–20 points). Higher scores reflect poorer outcomes.
Mediator		
Classroom Structural and Process Quality	To be adapted from the Measure of Early Learning Environments (MELE) module of the Measuring Early Learning Quality and Outcomes (MELQO) tool [54] and the Instrument for Measuring Quality of Early Childhood Education in Colombia (IMCEIC) [55]	Direct Assessment of children's pre-primary learning environments in LEAPS preschool and comparison classrooms. The 7 major domains are: environment and physical setting, family and community engagement, personnel, interactions, inclusiveness, pedagogy, and play. Mean scores will be calculated from all items (range 1–4 points). Higher scores reflect better outcomes.
Covariates		
Child's Household Demographics	Child Household Questionnaire, including Socio-Economic Status Index [68]*; Household Food Insecurity Access Scale (HFIAS) [69]*	Household questionnaire including items examining child's age, child's gender, parental education, parental employment, assets-based household socio-economic status, and household food insecurity.
Child Anthropometric Indices	Anthropometric Questionnaire*	Anthropometric Questionnaire collects child age and child growth indicators. Child's anthropometric z scores will be calculated using WHO Child Growth Standards.
Youth's Household Demographics	Youth Enrollment Questionnaire, including Socio-Economic Status Index [68]*; Household Food Insecurity Access Scale (HFIAS) [69]*	Youth Enrollment Questionnaire including items examining the youth's demographics, such as age, education, marriage status, assets-based household socio-economic status, and household food insecurity.

Data collection methods

Data collection team

All data collection will be conducted by trained assessors, most of whom have experience working in prior early childhood development trials. Assessors have at least a bachelor's level education. Data collection will be overseen by the evaluation field lead and field research manager. All trainings will be conducted by the evaluation field lead and field research manager. Assessors

will receive an initial 10-day training for all child and youth assessments, followed by a 1-day refresher training prior to round of data collection. To complete training, assessors must have > 95% agreement for all child and youth direct assessments with the evaluation field lead. Data collectors engaged in conducting qualitative data collection will receive an additional 1-week training in qualitative methods, followed by continued refreshers and debriefing sessions for each qualitative data

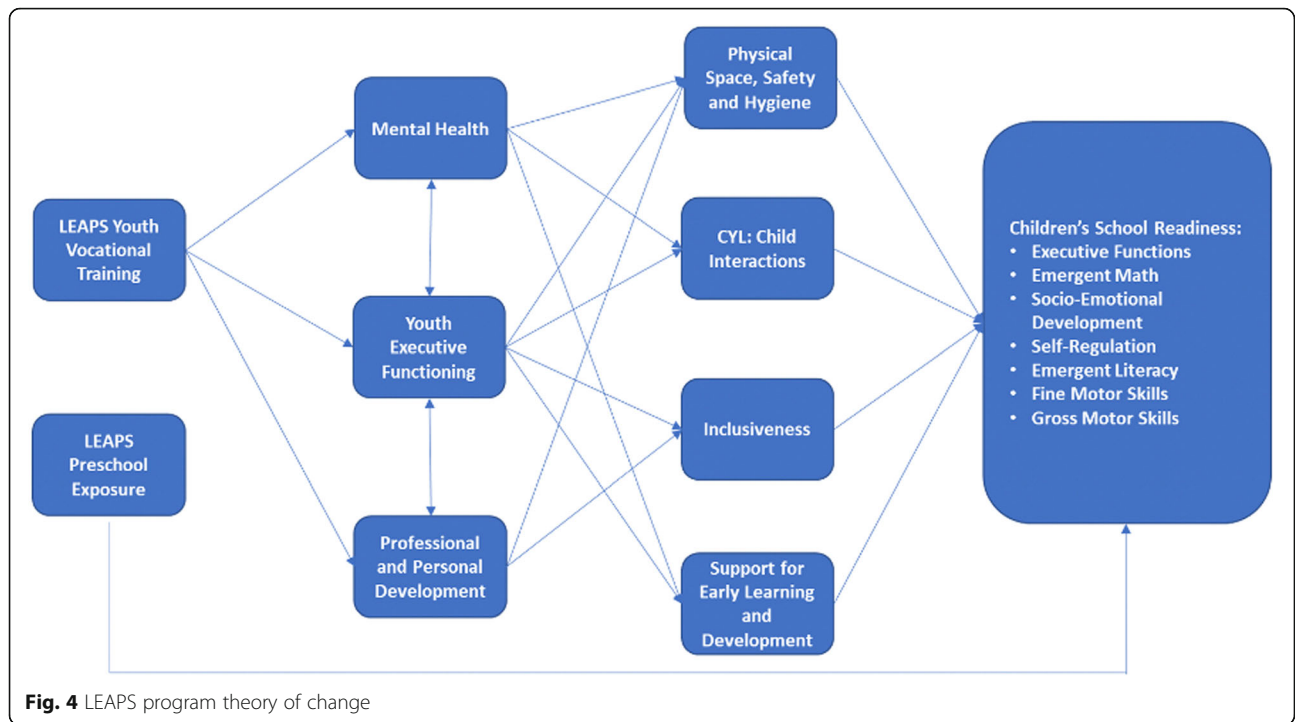


Fig. 4 LEAPS program theory of change

collection activity. For the classroom observation tool, the field research manager and project coordinator will complete a training-of-trainers with an expert at the Universidad de los Andes, Bogota, Colombia. The field research manager will then train the evaluation field lead with support from the project coordinator. The field research manager and evaluation field lead will then deliver a 1-week training to a team of four assessors. Assessors will be required to achieve an interrater reliability of > 70% agreement with a Master Code in coding three 30–45-min videos in order to complete training. Coordination meetings will be held daily, with intensive debriefing meetings including capacity building activities held once per week. The evaluation field lead will conduct regular field visits to observe data collection to maintain quality. In each of the three steps, two assessments per assessor will be observed and independently scored for interrater reliability. Assessments will be video recorded and reviewed with team members during weekly debriefing meetings.

Data collection procedures for child and caregiver assessments

Assessors will conduct household visits to issue invitations to the study. If accepted, the assessor will enroll the caregiver in the study on site and administer the household questionnaire to the caregiver during the home visit. The assessor will evaluate the child at the household, and the child will be provided with a rest

break and refreshments as needed. Questionnaires will be completed with caregivers using tablets. Data collection for the direct assessments will be conducted manually and then later entered in tablets. At the end of the session the child/caregiver will be given a thank you gift (e.g., book or toy).

Data collection procedures for youth participant assessments

CYLs will be issued an invitation to the study during the recruitment and training process. Assessors will issue invitations to comparison youth after completing the screening tool. The youth participant will either be enrolled and assessed on site or at a follow-up appointment, as is practical for the participant. The assessment will be conducted either at a recruitment workshop, training session, or during a home visit with the youth participant. All consecutive follow-up assessments will be conducted at the participant’s household, at the LEAPS school, or at an alternative community space identified for assessments. Location of assessments will be selected based on convenience for the youth participant and the availability of private space. All questionnaires will be collected using tablets. Executive functioning assessments will be conducted manually and later entered in tablets. Youth participants will receive a thank you gift (e.g., a LEAPS mug) and refreshments for participation.

Data collection procedures for classroom observations

Assessors will first obtain permission from the local government and school leadership for the observations, after which they will visit schools to coordinate with the teacher and supervisor to schedule the observation. Assessors will arrive before the start of the school day when possible. Assessors should ensure that they are able to observe a minimum of 2.5 h of the classroom session. If there is less than 2.5 h of the session remaining, the observation should be rescheduled. If a teacher/CYL is absent, the observation will be rescheduled. Assessors should conduct a maximum of two visits: if the teacher/CYL is absent on the first visit, the assessor will reschedule and conduct a second visit. If the teacher/CYL is absent on the second visit, then the school will no longer be included in the data collection round. The observation may be video recorded as an additional resource to reference when determining final scores. Assessors will observe the class without interrupting the school routine or activities. Assessors will take notes throughout the session and manually complete the observation tool after the session. Videos may be reviewed for items where there is a discrepancy in scoring.

Qualitative data collection procedures

All IDIs and FGDs will be conducted by data collectors trained in interview techniques. All activities will be conducted to ensure the participant's privacy and comfort, and to minimize risk of disturbances. The IDIs and FGDs will be semi-structured, using pre-piloted topic guides and will be recorded. The FGDs will additionally be observed by a data collector trained as a note-taker.

Data management

Data is collected on encrypted tablets and will be uploaded from the tablets to an encrypted server maintained by the Aga Khan University Data Management Unit (DMU). Paper-based assessments (i.e., classroom observation tool) will be entered in the DMU database using FoxPro by the DMU research assistant at the field office. Initial checks for quality assurance will be conducted by the DMU research assistant and a DMU data coordinator in collaboration with the evaluation field lead. Data will be shared via regular secure transfers with the team members at Harvard, where it will be uploaded to a password-protected shared Dropbox folder. Access to the data folder will only be given to the research team. At Harvard, the database will be managed by the data analyst and project coordinator, with regular cleaning and checks for quality assurance. All data will be de-identified and the keys linking identifiers to participants will be stored in a password-protected folder. Data will be wiped from the tablets after the round of data collection activities have been completed. All video and audio

recordings will be securely stored on a password-protected computer at the field office managed by the evaluation field lead and will be deleted from the camera/recorder after uploading. Recordings will be deleted after analysis is complete (i.e., within 1 year of endline). Paper-based data assessments will be stored in a locked cabinet at the field site.

Qualitative data will be transcribed in Sindhi by assessors. Five percent of recordings will be randomly chosen by the field research manager and evaluation field lead to check quality of transcription. Transcripts will be translated from Sindhi to English by a translator outside of the project team. Five percent of translations will be chosen randomly for independent back translations. Transcripts, notes, and photos from original qualitative data collection activities will be managed through a password-protected shared Dropbox folder. Only research team members will have access. All data will be de-identified and the keys linking identifiers to participants will be stored in a password-protected folder. Audio recordings will be securely stored on a password-protected computer at the field office managed by the evaluation field lead and will be destroyed after the project has been completed.

Analysis

Impact evaluation

The primary analysis will use the intent-to-treat principle. Baseline balance will be tested between randomized groups to see whether randomization was successful. The primary outcome, IDELA total score, will be computed for each individual using the parameter estimates from the Graded Response Model, as recommended by the tool developers. We will assess the effect of the LEAPS intervention using the analytical framework proposed by Hussey and Hughes [71] for analysis of stepped wedge cluster-randomized trials. Estimation of treatment effects will be obtained from a generalized linear mixed model (GLMM) to produce the mean difference in IDELA scores for LEAPS intervention versus control. The GLMM will include a fixed effect for stratified randomization group, a random effect for cluster, a fixed effect for time, and use an exchangeable correlation structure. In addition, the study will perform a set of sensitivity analyses to explore potential deviations from the basic SW-CRT model proposed by Hussey and Hughes [71]. The closed school cohort data will be to examine the changes in the average school readiness among children participating in the LEAPS preschool between baseline and endline. Longitudinal data collected for school cohort participants will be linked using participant IDs. The preschool classroom structural and process quality data and the child household characteristics will serve as predictors of interest in adjusted

analyses. Youth cohort data will be analyzed to examine differences in outcomes over time among CYLs as compared to the youth comparison group. Longitudinal youth participant data will be linked using participant IDs. Analysis will account for CYLs' length of program exposure and adjust for potential confounding variables, including age, education, marriage status, assets-based household socio-economic status, and household food insecurity. No interim analyses are planned.

Implementation and systems readiness evaluations

Secondary administrative data for program monitoring will be analyzed using Excel and Stata to examine trends in program quality indicators over time and by subgroups (e.g., district, FO, CYL). Translated transcripts of the semi-structured IDIs and FGDs will be analyzed through a thematic analysis. Team members will independently review transcripts, followed by a combination of inductive and deductive coding and discussion to identify themes. A codebook and coding matrices will be constructed to support coding and analysis. Triangulation between stakeholders and data sources will be analyzed to understand both converging patterns and unique perspectives.

Cost evaluation

A cost analysis will be conducted to estimate total project costs for all stakeholders, in addition to cost per beneficiary and cost per impact. Costing data will inform NCHD efforts for expansion and replication of LEAPS.

Ethics and dissemination

Protocol amendments

All modifications to the protocol will be communicated with the PI, co-PI, and co-investigators through regular meetings and email communications. Any amendments to the protocol that have implications for the study design or participant safety will be reported to the affiliated ethics review boards through applications and email communications and to the trial registry through the application process. Protocol amendments impacting trial participants will be communicated in-person by the field research manager, evaluation field lead, or trained assessors.

Reporting and risk of adverse events and harms

The LEAPS research team will be responsible for monitoring regulatory documents and participant files, including reviewing and maintaining all required ERC/IRB documentation, monitoring and updating regulatory supporting documents, recording meeting minutes, documenting the recruitment, enrollment, and informed consent processes, and maintaining logs to track data collection and assessment activities.

There is a small risk of loss of confidentiality for participants. To minimize this risk, interviews will be conducted in private spaces; tablets will be password protected; hard copies of forms will be stored in locked cabinets; data will be de-identified prior to sharing; and assessors will receive training in good clinical practice.

There is additionally a risk of psychological discomfort for youth participants due to survey questions on depressive symptoms. There is minimal risk associated with administration of the child assessments for executive function skills and school readiness and the household questionnaire with caregivers. To mitigate these risks, assessors are trained in how to sensitively administer youth and child assessments, and how to communicate with children's parents about the assessment and their child's performance. Participants will be informed that they may stop at any time without consequence.

The youth participant questionnaire includes a question on suicidal ideation. If the participant responds positively to this item, the assessor will contact the evaluation field lead and field research manager who will assess whether there is an imminent danger of self-harm. If there is immediate danger, then the assessor will escort the participant to a health provider. The team will also discuss with a consultant at the Aga Khan University Department of Psychiatry as needed. In no instance will the person be left alone. If the person has suicidal ideation, but is not at risk of immediate self-harm, the study staff will make an active referral for mental health care.

Finally, corporal punishment is unfortunately widely accepted in the intervention communities, so there is a risk that CYLs may initially use harsh disciplinary practices in the LEAPS classrooms. To mitigate this risk, CYL training includes preschool behavior management, covering strategies for appropriate classroom management and employing gentle discipline. The NCHD's CYL contracts also include a clause prohibiting use of corporal punishment. If a CYL is observed using harsh discipline in the classroom, she will receive targeted training in classroom management. Repeated episodes of harsh discipline by a CYL will result in termination of the CYL's employment.

Consent to participate

For all data collection activities, informed consent will be obtained from participants by the research team's assessors, who have received Good Clinical Practice training. For participants under age 18 years, both child assent and consent from a legal guardian will be obtained by a trained assessor. Consent forms were written to explain the information with accuracy and clarity for individuals with low levels of literacy (on

average mothers in this location have 2 years of formal education) [26, 72]. All forms are written in Sindhi, the local language. All subjects will be free to withdraw from the study at any time and will be assured that this will not affect the standard of care received in the community.

Consent for the preschool classroom observations and video recording will be obtained by a trained assessor from the CYL/comparison preschool teacher, as well as a representative from the school's parent-community board. For schools that do not have a parent-community board, the assessor will instead seek consent from the school's leadership, i.e., the school principal or coordinator. Consent will be required from both the CYL/comparison preschool teacher and the representative from the school board or school leadership in order to proceed with the video recordings. The assessor will then enroll the CYL/teacher at the school and observe the classroom session to complete the assessment.

Confidentiality

All assessments will be conducted in a private location, either in an assessment center, classroom, office, or private household. All data (including recorded interviews and notes, direct assessments, questionnaires, video recordings for quality assurance) will be de-identified by the research team. The keys linking identifiers to participants will be stored in a password-protected file (separated from the other data files), to which only the research team members will have access. All hard copy materials containing identifiable information (i.e., consent forms, interview cover sheets, photos) will be stored in locked cabinets. All tablets used for data entry will be password protected. Tablets and video recorders will be stored in locked cabinets when not in use. Data will be regularly transferred to the central server at Aga Khan University and then wiped. Video and audio recordings will be deleted after analysis is completed (within 1 year of endline). All research team staff will receive training in protecting participant privacy and adherence to Good Clinical Practices.

Ancillary and post-trial care

Given that LEAPS serves vulnerable populations with limited access to education, training, and employment, the research team has a responsibility to ensure that LEAPS participants are supported beyond the trial period. The research team will coordinate with NCHD to ensure that (a) LEAPS children have a clear pathway to transition to primary school at the close of the trial and (b) that CYLs are supported through career counseling exit meetings to identify further options for education and employment.

Dissemination: trial results

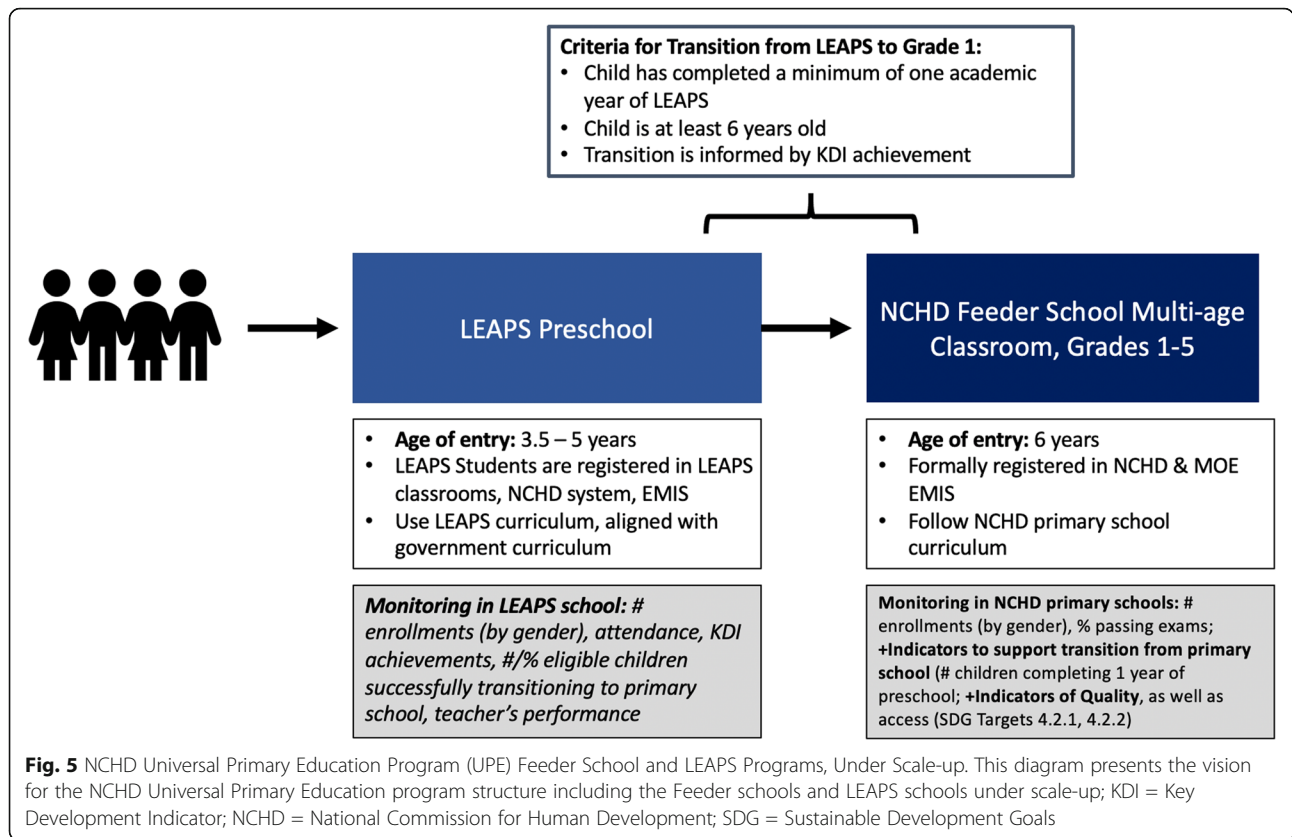
Results will be disseminated through a variety of channels, including workshops with government stakeholders, funder and lay reports, presentations at national and international conferences, publications in peer-reviewed journals, and through networks for early childhood and youth development such as Asia-Pacific Regional Network for Early Childhood, the Pakistan Alliance for Early Childhood, the Early Childhood Development Action Network, Early Childhood Peace Consortium, and the Alliance for International Youth Development.

Discussion

This will be the first larger-scale impact evaluation of a youth-led ECCE program in an LMIC, providing much-needed empirical support for the youth-led ECCE model. The results of this evaluation will enable Pakistan's NCHD to make evidence-based decisions about strengthening and expanding their community-based preschool and youth vocational training systems. The evidence from this evaluation will also be of relevance to regional and global settings where there is a vast unmet need for preschool services and limited employment and learning opportunities for youth. This study will additionally contribute to the literature around capacity development to support sustainable transitions of programming from donors to government ownership with changing levels of support technical partners, providing evidence from a systematic approach to capacity building and program transition with government partners in a rural, resource-limited context. It is our vision that this trial will lay the groundwork for future efforts to expand and replicate LEAPS in other NCHD communities, in order to strengthen NCHD's ECCE programming on a wider, systems level. Figure 5 depicts the end goal for a fully integrated LEAPS program model, linked to NCHD Feeder Primary Schools within the Universal Primary Education Program. An expanded program model would additionally aim to obtain accreditation for CYL vocational training. Finally, findings from this work will contribute to the research around gender equity and youth development, providing a rigorous evaluation of a female-specific youth development program in an LMIC, a context for which there are very few program evaluations [18].

Limitations

A limitation of the stepped wedge design is that time is associated with both the exposure to the intervention and potentially the outcome, and in this way is a possible confounder [40]. Calendar time will therefore be adjusted for in the analysis. Potential interruptions



due to unexpected emergency contexts, such as emergency shutdowns and flooding, also pose a risk to the study design and implementation. Finally, because the project is implemented in collaboration with a government organization, turnover in government partners may create challenges for program delivery.

Trial status

This trial is currently active (protocol version 5, April 9, 2021; see Additional file 2). Recruitment of participants began in May 2021. Final participants are expected to be recruited and complete their assessments at the end of 2020.

Abbreviations

AKU: Aga Khan University; CYL: Community Youth Leader; DMU: Data Management Unit; ECD: Early Childhood Development; ECCE: Early Childhood Care and Education; FGD: Focus group discussion; FO: NCHD Field Officer; HFIAS: Household Food Insecurity Access Scale; IDELA: International Development and Early Learning Assessment; ICC: Intraclass correlation coefficient; IDI: In-depth interview; IMCE IC: Instrument for Measuring Quality of Early Childhood Education in Colombia; KDI: Key Development Indicator; LEAPS: Youth Leaders for Early Childhood Assuring Children are Prepared for School; LHW: Lady Health Worker; LIST: LEAPS Intervention Support Team; LMIC: Low- and middle-income country; MELE: Measure of Early Learning Environments; MELQO: Measuring Early Learning Quality and Outcomes; NCHD: National Commission for Human Development; PI: Principal Investigator; PTM: Parent Teacher Meeting; PYD: Positive Youth Development; SDG: Sustainable Development Goal; SPIRIT: Standard Protocol Items: Recommendations for

Interventional Trials; SRQ-20: Self-Reporting Questionnaire 20-Item; TiDIERS: Template for Intervention Description and Replication; TMT: Trail Making Task; UPE: NCHD Universal Primary Education program

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13063-021-05518-9>.

Additional file 1. Supplemental tables and figures.

Additional file 2. Trial protocol registration and sponsor details.

Acknowledgements

We sincerely thank the National Commission for Human Development Head Office in Islamabad, Province Office in Karachi, Sindh, and four District Offices in Sindh province, including district Dadu, Khairpur, Naushahro Feroze, and Sukkur, for their invaluable partnership and support in carrying out this study, as well as the LEAPS families, children and CYLs for their support and commitment. We are deeply grateful to Ms. Shahnaz Parveen Hakro and Tasleem Soomro for their dedication and leadership. Finally, we thank Ms. Arjumand Rizvi at the Aga Khan University for her support with the sample size and power calculations, as well as Mr. Amjad Hussain, Mr. Zubair Tagar, and Mr. Hassan Naqvi with the Aga Khan University Data Management Unit for their support in evaluation and data collection efforts.

Authors' contributions

AKY conceptualized the study, led the development of the LEAPS intervention, identified outcomes measures, and contributed to trial design and plan of analysis. CS designed the study trial and contributed to the plans for analyses. EF and SS adapted measures for outcome data collection for the study, and EF contributed to the plan of analyses. SS, EF, and KR developed the intervention components of the LEAPS Programme and qualitative topic guides. CR and LP contributed to the development of the

intervention and quantitative and qualitative measures. SB contributed to the qualitative topic guides and to the development of the intervention. GF contributed to the study design, measures, and plans for analyses. RD contributed to the plan of analysis. AKY, EF, KR, and RD prepared the first draft. All authors reviewed, provided critical feedback, and approved the final protocol. Authorship will be decided in accordance with the International Committee of National Journal Editors guidelines.

Funding

Funding for this work is provided by Dubai Cares and Saving Brains Program Grand Challenges Canada (award number TTS-1808-17846). The funder has no role in the development or implementation of the study, nor will they be involved in the analysis, interpretation, report writing, or publication and dissemination.

Availability of data and materials

Following publication of all planned manuscripts, the final de-identified dataset will be stored indefinitely on the Harvard Open Data Dataverse Network, except for in the case of request of withdrawal by the participant for themselves or the child in which case the data will be deleted. Permission to use the data should be sought from the PI in advance, and the source should be cited.

Declarations

Ethics approval and consent to participate

All protocol, consent forms, and data collection tools were reviewed and approved by the Institutional Review Board of the Harvard T. H. Chan School of Public Health, Boston, MA, USA (Protocol #: IRB18-1149); the Ethics Review Committee at Aga Khan University Karachi, Pakistan (5473-Ped-ERC-18); and the National Bioethics Committee Pakistan, Islamabad, Pakistan (Ref: No.4-87/NBC-334-Y2-Extension/19). All participants/legal guardians will have the provided informed consent to participate.

Consent for publication

Not applicable.

Competing interests

The authors have no competing interests.

Author details

¹Harvard T. H. Chan School of Public Health, Boston, MA, USA. ²Aga Khan University Karachi, Sindh, Pakistan. ³Yale Child Study Center, New Haven, CT, USA. ⁴Swiss Tropical and Public Health Institute, Basel, Switzerland.

Received: 1 December 2020 Accepted: 6 August 2021

Published online: 17 August 2021

References

- UN General Assembly. Transforming our world: the 2030 Agenda for Sustainable Development. New York: Division for Sustainable Development Goals; 2015. Available from: <https://www.refworld.org/docid/57b6e3e44.html>
- Richter LM, Daelmans B, Lombardi J, Heymann J, Boo FL, Behrman JR, et al. Investing in the foundation of sustainable development: pathways to scale up for early childhood development. *Lancet*. 2017;389(10064):103–18. [https://doi.org/10.1016/S0140-6736\(16\)31698-1](https://doi.org/10.1016/S0140-6736(16)31698-1).
- Nilsson M, Griggs D, Visbeck M. Policy: map the interactions between Sustainable Development Goals. *Nature*. 2016;534(7607):320–2. <https://doi.org/10.1038/534320a>.
- Lu C, Black MM, Richter LM. Risk of poor development in young children in low-income and middle-income countries: an estimation and analysis at the global, regional, and country level. *Lancet Glob Health*. 2016;4(12):e916–22. [https://doi.org/10.1016/S2214-109X\(16\)30266-2](https://doi.org/10.1016/S2214-109X(16)30266-2).
- McCoy DC, Zuilkowski SS, Yoshikawa H, Fink G. Early childhood care and education and school readiness in Zambia. *J Res Educ Effectiveness*. 2017; 10(3):482–506.
- Britto PR, Lye SJ, Proulx K, Yousafzai AK, Matthews SG, Vaivada T, et al. Nurturing care: promoting early childhood development. *Lancet*. 2017; 389(10064):91–102. [https://doi.org/10.1016/S0140-6736\(16\)31390-3](https://doi.org/10.1016/S0140-6736(16)31390-3).
- Yoshikawa H, Kabay S. The evidence base on early childhood care and education in global contexts; background paper prepared for the education for all global monitoring report 2015, education for all 2000-2015: achievements and challenges. Paris: UNESCO; 2015.
- Neuman MJ, Josephson K, Chua PG. A review of the literature: early childhood care and education (ECCE) personnel in low-and middle-income countries, Early Childhood Care and Education Working Papers Series, vol. 4; 2015.
- UNICEF. School readiness: a conceptual framework. New York: United Nations Children's Fund; 2012.
- UNESCO Institute for Statistics. More than one-half of children and adolescents are not learning worldwide. UIS Fact Sheet. 2017;46 Available from: <http://uis.unesco.org/sites/default/files/documents/fs46-more-than-half-children-not-learning-en-2017.pdf>.
- Pakistan ASER. Annual Status of Education Report 2019 (Provincial); 2020.
- Bartlett S. Learning about learning: reflections on studies from 10 countries. Geneva: Aga Khan Foundation; 2013. Available from: <https://www.akdn.org/publication/learning-about-learning>
- Solutions for Youth Employment (S4YE). Towards solutions for youth employment: a 2015 baseline report. 2015. Available from: <https://www.s4ye.org/resources>
- International Labour Organization. Global employment trends for youth 2013: a generation at risk. 2013.
- Organization for Economic Co-operation and Development (OECD) and International Labour Organization. Promoting better labour outcomes for youth. 2014. Available from: https://www.g20.utoronto.ca/2014/OECD-ILO-better_outcomes_for_youth.pdf
- Patton GC, Sawyer SM, Santelli JS, Ross DA, Afifi R, Allen NB, et al. Our future: a Lancet commission on adolescent health and wellbeing. *Lancet*. 2016;387(10036):2423–78. [https://doi.org/10.1016/S0140-6736\(16\)00579-1](https://doi.org/10.1016/S0140-6736(16)00579-1).
- United Nations Development Programme (UNDP) Pakistan. Pakistan national human development report: Unleashing the potential of a young Pakistan. Islamabad: UNDP; 2017. Available from: <https://www.un.org.pk/undp-pakistan-launches-the-pakistan-national-human-development-report-2017-unleashing-the-potential-of-a-young-pakistan/>
- Catalano RF, Skinner ML, Alvarado G, Kapungu C, Reavley N, Patton GC, et al. Positive youth development programs in low-and middle-income countries: a conceptual framework and systematic review of efficacy. *J Adolesc Health*. 2019;65(1):15–31. <https://doi.org/10.1016/j.jadohealth.2019.01.024>.
- Alvarado G, Skinner M, Plaut D, Moss C, Kapungu C, Reavley N. A systematic review of positive youth development programs in low-and middle-income countries. Washington, DC: YouthPower Learning, Making Cents International; 2017.
- Nelson CA. The neurobiological basis of early intervention. In: Shonkoff JP, Meisels SJ, editors. Handbook of early childhood intervention. 2nd ed. Cambridge: Cambridge University Press; 2000.
- Konrad K, Firk C, Uhlhaas PJ. Brain development during adolescence: Neuroscientific insights into this developmental period. *Dtsch Arztebl Int*. 2013;110(25):425–31. <https://doi.org/10.3238/arztebl.2013.0425>.
- UNICEF Office of Research - Innocenti. The adolescent brain: a second window of opportunity - a compendium. Florence: UNICEF Office of Research - Innocenti; 2017.
- Blair C, Razza RP. Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev*. 2007;78(2):647–63. <https://doi.org/10.1111/j.1467-8624.2007.01019.x>.
- Duncan GJ, Dowsett CJ, Claessens A, Magnuson K, Huston AC, Klebanov P, et al. School readiness and later achievement. *Dev Psychol*. 2007;43(6):1428–46. <https://doi.org/10.1037/0012-1649.43.6.1428>.
- Shonkoff JP, Fisher PA. Rethinking evidence-based practice and two-generation programs to create the future of early childhood policy. *Dev Psychopathol*. 2013;25(4):1635.
- Yousafzai AK, Rasheed MA, Rizvi A, Shaheen F, Ponguta LA, Reyes CR. Effectiveness of a youth-led early childhood care and education programme in rural Pakistan: a cluster-randomised controlled trial. *PLoS One*. 2018; 13(12):e0208335. <https://doi.org/10.1371/journal.pone.0208335>.
- Ponguta LA, Rasheed MA, Reyes CR, Yousafzai AK. A conceptual model for youth-led programs as a promising approach to early childhood care and education. *N Dir Child Adolesc Dev*. 2018;2018(159):71–89. <https://doi.org/10.1002/cad.20233>.
- Fellowship G. Gandhi fellowship handbook. New Delhi; 2016. Available from: <https://gandhifellowship.org/>
- Adoho F, Chakravarty S, Korkoyah DT, Lundberg M, Tasneem A. The impact of an adolescent girls employment program: The EPAG project in Liberia. World Bank. 2014. <https://doi.org/10.1596/1813-9450-6832>.

30. Bowen F, Devercelli A. Economic empowerment of adolescent girls and young women (EPAG) [paper symposium]. Baltimore: SRCD 2019 Biennial Meeting; 2019. Available from: https://www.srcd.org/sites/default/files/file-attachments/srcd19programguide_web_1.pdf
31. Pisani L, Borisova I, Dowd AJ. International development and early learning assessment technical working paper. Save the Children; 2015.
32. Yousafzai AK, Rasheed M, Rizvi A, Shaheen F, Ponguta LA, Reyes CR. Cluster-Randomized Clinical Trial and Process Evaluation of the LEAPS Program in Rural Pakistan [Paper Symposium]. Baltimore: SRCD 2019 Biennial Meeting; 2019. Available from: https://www.srcd.org/sites/default/files/file-attachments/srcd19programguide_web_1.pdf
33. Yousafzai AK, Rasheed MA, Siyal S, Franchett E, Sudfeld C, Ponguta LA, et al. LEAPS: a strategy to benefit young children and youth. *Early Childhood Matters: Advances in early childhood development*; 2019. p. 102–5.
34. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med*. 2013;158(3):200–7. <https://doi.org/10.7326/0003-4819-158-3-201302050-00583> PMID: 23295957; PMCID: PMC5114123.
35. Pakistan Bureau of Statistics. Pakistan Social & Living Standards Measurement Survey 2018-19 National/Provincial (Social Report). 2020.
36. Pakistan Bureau of Statistics. Labor Force Survey 2017-18. 2018.
37. National Commission for Human Development. National Commission for Human Development: About Us. Available from: http://www.nchd.org.pk/ws/index.php?option=com_content&view=article&id=6&Itemid=2 [Accessed 23 Oct 2020].
38. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *Bmj*. 2014;7:348.
39. Yousafzai AK, Rasheed MA, Siyal S. Integration of parenting and nutrition interventions in a community health program in Pakistan: an implementation evaluation. *Ann N Y Acad Sci*. 2018;1419(1):160–78. <https://doi.org/10.1111/nyas.13649>.
40. Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *Bmj*. 2015; 350(feb06 1):h391. <https://doi.org/10.1136/bmj.h391>.
41. Linnan L, Steckler A. Process evaluation for public health interventions and research; 2002.
42. El Fakiri F, Hoes AW, Uitewaai PJ, Frenken RA, Bruijnzeels MA. Process evaluation of an intensified preventive intervention to reduce cardiovascular risk in general practices in deprived neighbourhoods. *Eur J Cardiovasc Nurs*. 2008;7(4):296–302. <https://doi.org/10.1016/j.ejcnurse.2008.01.002>.
43. Levin HM, McEwan PJ. Cost-effectiveness analysis: methods and applications. Thousand Oaks: Sage; 2000.
44. USAID, AIDSTAR-Two. Organizational capacity building framework: a foundation for strong, more sustainable HIV/AIDS programs, organizations & networks, vol. 2: Technical Brief; 2011. Available from: https://www.msh.org/sites/default/files/as2_technical_brief_2_final.pdf
45. USAID E. Human and institutional capacity development handbook: A USAID model for sustainable performance improvement. Washington, DC: US Agency for International Development; 2010. Available from: <https://www.hsdl.org/?abstract&id=14447>
46. LaFond A, Brown L. A guide to monitoring and evaluation of capacity-building interventions in the health sector in developing countries. MEASURE Evaluation Manual Series. Chapel Hill: Carolina Population Center, University of North Carolina at Chapel Hill; 2003.
47. Otoo S, Agapitova N, Behrens J. The capacity development results framework; 2009.
48. Bennett S, Ozawa S, Rodriguez D, Paul A, Singh K, Singh S. Monitoring and evaluating transition and sustainability of donor-funded programs: Reflections on the Avahan experience. *Eval Program Plann*. 2015;52:148–58. <https://doi.org/10.1016/j.evalprogplan.2015.05.003>.
49. Bao J, Rodriguez DC, Paina L, Ozawa S, Bennett S. Monitoring and evaluating the transition of large-scale programs in global health. *Glob Health Sci Pract*. 2015;3(4):591–605. <https://doi.org/10.9745/GHSP-D-15-00221> pmid: 26681706; 2015.
50. Sgaier SK, Ramakrishnan A, Dhingra N, Wadhvani A, Alexander A, Bennett S, et al. How the Avahan HIV prevention program transitioned from the Gates Foundation to the government of India. *Health Aff*. 2013;32(7):1265–73. <https://doi.org/10.1377/hlthaff.2012.0646>.
51. Crocker J, Shields KF, Venkataraman V, Saywell D, Bartram J. Building capacity for water, sanitation, and hygiene programming: Training evaluation theory applied to CLTS management training in Kenya. *Soc Sci Med*. 2016;166:66–76. <https://doi.org/10.1016/j.socscimed.2016.08.008>.
52. Plan International Kenya. Plan International Kenya: CLTS management capacity building for district-level officials, Implementation Narrative. 2015.
53. Fernald LC, Prado E, Kariger P, Raikes A. A toolkit for measuring early childhood development in low and middle-income countries; 2017. <https://doi.org/10.1596/29000>.
54. UNESCO UNICEF Brookings Institution and the World Bank. Overview: Measuring Early Learning Quality and Outcomes (MELQO). 2017
55. Ponguta LA, Maldonado-Carreño C, Kagan SL, Yoshikawa H, Nieto AM, Aragón CA, Mateus A, Rodríguez AM, Motta A, Varela LR, Guerrero PA. Adaptation and application of the measuring early learning quality and outcomes (MELQO) framework to early childhood education settings in Colombia. *Zeitschrift für Psychologie*. 2019;227(2):105–12. <https://doi.org/10.1027/2151-2604/a000361>.
56. Rahman A, Iqbal Z, Waheed W, Hussain N. Translation and cultural adaptation of health questionnaires. *J Pak Med Assoc*. 2003;53(4):142–7.
57. Noël MP. Counting on working memory when learning to count and to add: a preschool study. *Dev Psychol*. 2009;45(6):1630–43. <https://doi.org/10.1037/a0016224>.
58. Molfese VJ, Molfese PJ, Molfese DL, Rudasill KM, Armstrong N, Starkey G. Executive function skills of 6–8 year-olds: brain and behavioral evidence and implications for school achievement. *Contemp Educ Psychol*. 2010;35(2): 116–25. <https://doi.org/10.1016/j.cedpsych.2010.03.004>.
59. Frye D, Zelazo PD, Palfai T. Theory of mind and rule-based reasoning. *Cogn Dev*. 1995;10(4):483–527. [https://doi.org/10.1016/0885-2014\(95\)90024-1](https://doi.org/10.1016/0885-2014(95)90024-1).
60. Carlson SM. Developmentally sensitive measures of executive function in preschool children. *Dev Neuropsychol*. 2005;28(2):595–616 https://doi.org/10.1207/s15326942dn2802_3 PMID: 16144429.
61. Ponitz CE, McClelland MM, Jewkes AM, Connor CM, Farris CL, Morrison FJ. Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Childhood Res Q*. 2008;23(2):141–58. <https://doi.org/10.1016/j.ecresq.2007.01.004>.
62. Davis HL, Pratt C. The development of theory of mind: the working memory explanation. *Austr J Psychol*. 1995;11(1):73–92. <https://doi.org/10.1080/00049539508258765>.
63. Kessels RP, Van Zandvoort MJ, Postma A, Kappelle LJ, De Haan EH. The Corsi block-tapping task: standardization and normative data. *Appl Neuropsychol*. 2000;7(4):252–8. https://doi.org/10.1207/S15324826AN0704_8.
64. Stroop JR. Studies of interference in serial verbal reactions. *J Exp Psychol*. 1935;18(6):643–62. <https://doi.org/10.1037/h0054651>.
65. Lezak MD. Neuropsychological Assessment. New York: Oxford University Press; 1995.
66. Reitan RM, Wolfson D. The Halstead-Reitan neuropsychological test battery: theory and clinical interpretation. 1993. Tucson: Reitan Neuropsychology Laboratory; 1993.
67. Beusenberg M, Orley JH, World Health Organization. A user's guide to the self reporting questionnaire (SRQ). Geneva: World Health Organization; 1994.
68. Vyas S, Kumaranayake L. Constructing socio-economic status indices: how to use principal components analysis. *Health Policy Plann*. 2006;21(6):459–68. <https://doi.org/10.1093/heapol/czl029>.
69. Coates J, Swindale A, Bilinsky P. Household Food Insecurity Access Scale (HFAS) for measurement of food access: indicator guide: version 3; 2007.
70. Hemming K, Taljaard M. Sample size calculations for stepped wedge and cluster randomised trials: a unified approach. *J Clin Epidemiol*. 2016;69:137–46. <https://doi.org/10.1016/j.jclinepi.2015.08.015>.
71. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemporary Clin Trials*. 2007;28(2):182–91. <https://doi.org/10.1016/j.cct.2006.05.007>.
72. Yousafzai AK, Rasheed MA, Rizvi A, Armstrong R, Bhutta ZA. Effect of integrated responsive stimulation and nutrition interventions in the Lady Health Worker programme in Pakistan on child development, growth, and health outcomes: a cluster-randomised factorial effectiveness trial. *Lancet*. 2014;384(9950):1282–93. [https://doi.org/10.1016/S0140-6736\(14\)60455-4](https://doi.org/10.1016/S0140-6736(14)60455-4).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.