Shaping Future Success: Evidence from an Early Childhood Human Capital Formation Intervention

Deepak Saraswat [*] Shwetlena Sabarwal [†] Lindsey Lacey [‡] Natasha Jha [§]
Nishith Prakash [¶] Rachel Chazan-Cohen [‖]

May 16, 2026

Abstract

Nearly 250 million children under five in low- and middle-income countries face developmental deficits despite expanding access to early childhood services. We present evidence from a large-scale randomized controlled trial (N=3,131 children in 201 schools) in Nepal's government school system testing three implementation models that combine classroom quality improvements with parental engagement. Following a 15-day teacher training, schools were randomly assigned to deliver caregiver sessions through teachers alone, teachers with in-class helpers, or external facilitators. All three models raised children's developmental outcomes by 0.10–0.20 standard deviations and improved caregiver engagement by similar magnitudes. The helper model proved most effective, sustaining classroom quality while maintaining teacher-family connections. Benefits are concentrated among households with lower baseline engagement, with the largest developmental gains among children starting furthest behind. Mechanism analysis reveals that the intervention transformed home and school from competing substitutes into complementary drivers of development. These findings demonstrate that feasible interventions embedded within existing systems can deliver scalable gains in early childhood human capital while narrowing inequality.

JEL Codes: J13, J24, I21, I24, O15

Keywords: Early childhood education, Cognitive skills, Non-cognitive skills, Ages and Stages Questionnaire, Child development, Nepal

The early stages of this project benefited greatly from the guidance of Nathan Fiala. We thank Mo Alloush, Ashutosh Bhuradia, Alex Eble, Peter Eggen, Willa Friedman, Alejandro J. Ganimian, Pamela Jakiela, Gaurav Khanna, Matthew A. Kraft, Owen Ozier, Aaryav Sharma, and Sandip Sukhtankar and participants at NEUDC 2025 for valuable comments that strengthened the paper. We are grateful to the Early Learning Partnership and the World Bank for generous financial support, without which this work would not have been possible. Excellent research assistance was provided by Vharun Boopathy, Shivanshi Garodia, Neil Oommen, Jae Wook Jung, Mridul Mishra, Nikhilesh Prakash, and Deepika Shrestha. We are grateful to the Center for Education and Human Resource Development (formerly Department of Education, Nepal), Officials in the local governments (municipalities) of Jhapa District, Provincial Education Training Center, Koshi Province, Seto Gurans National Child Development Services, Solutions Consultant, and the Nepal Development Research Institute for their invaluable collaboration in implementing the program and facilitating data collection. Author order was randomized using the American Economic Association's author-randomization tool. We used ChatGPT to assist with grammar checks and minor sentence-level improvements; all analysis, arguments, and interpretations are our own. Ethical approval for the study was obtained from the Nepal Health Research Council (NHRC Approval No. 862/2020) and from the Institutional Review Board of the University of Connecticut. The study was pre-registered with the AEA Trial Registry (RCT ID: AEARCTR-00113181).

[*]Amazon.com, Inc. Email: saraswat.d@gmail.com
[†]The World Bank. Email: ssabarwal@worldbank.org
[‡]Allegheny County Analytics. Email: Lindsey.Lacey@alleghenycounty.us
[§]The Centre for Advanced Financial Research and Learning. Email: natasha.jha@cafral.org.in
[¶]Northeastern University, BREAD, CESifo, IZA, GLO, HiCN, and CReAM. Email: n.prakash@northeastern.edu
[‖]University of Connecticut. Email: rachel.c.cohen@uconn.edu

1 Introduction

Children's earliest years are foundational for later life outcomes, with early interactions shaping cognitive, socio-emotional, and physical development (Phillips and Shonkoff, 2000). Yet millions of children in low- and middle-income countries (LMICs) begin school without these foundations, limiting their ability to benefit from education and economic opportunities. Nearly 250 million children under age five in LMICs are at risk of not reaching their developmental potential due to poverty, malnutrition, or inadequate learning environments, and this number remains high despite increased service provision (Black et al., 2017). These early deficits compound across the life course, reducing educational attainment, earnings, and health while increasing criminality (Heckman and Karapakula, 2019; Campbell et al., 2014; Brutti and Montolio, 2021). Addressing these deficits requires strengthening the two primary contexts for early learning: homes and classrooms. Research highlights that responsive parenting and enriched early learning environments together provide the basis for healthy development (Jeong et al., 2021; Shonkoff and Fisher, 2013).

As preschool enrollment expands in LMICs, concerns about quality are becoming increasingly urgent (Masino and Niño-Zarazúa, 2016). Governments recognize the long-term value of early childhood education for human capital development and are increasing investments in pre-primary programs (Jakiela et al., 2024; McCoy et al., 2017). Yet fewer than one in three children aged 3–4 attend formal programs, and pre-primary education receives less than 2 percent of education budgets (World Bank and UNICEF, 2020). Moreover, program designs often import standards from high-income settings that fit poorly in resource-constrained contexts where classes are large and cultural practices differ (Alan, 2025; Chen and Wolf, 2021). Without addressing these design and resource constraints, expanding access risks repeating the trajectory of universal primary education, where enrollment gains failed to translate into learning improvements (Pritchett, 2013).

This paper evaluates an early childhood intervention in Nepal that adapts a train-the-trainer model. All teachers receive training in developmentally appropriate practices; schools are then randomly assigned to one of three models for delivering parental education sessions: teachers alone, teachers with in-class helpers, or external facilitators. We test how supply-side constraints—specifically teacher workload and staffing support—influence program effectiveness across these implementation models. This design addresses a major constraint in LMIC education systems: teachers balance instruction with administrative duties, home visits, and community engagement, often with minimal external support (Glewwe and Kremer, 2006). When overburdened teachers adopt new responsibilities without relief

from existing demands, both classroom quality and family engagement efforts may suffer. Our experimental design allows us to identify whether and how implementation support shapes program outcomes.

We randomly assigned 201 schools in Jhapa district, Nepal, to one of three treatment arms or a control group. All teachers in treatment schools first completed a 15-day professional development program: ten days focused on pedagogy, classroom practices, and national early childhood education standards, followed by five days preparing them to lead caregiver sessions on health, nutrition, hygiene, disability awareness, and cognitive and socio-emotional stimulation. Teachers were then expected to deliver 20 two-hour sessions over six months. Treatment Arm 1 (T1, 50 schools) implemented this model with teachers working alone. Treatment Arm 2 (T2, 50 schools) added an in-class helper on a seven-month contract to assist with classroom management. Treatment Arm 3 (T3, 50 schools) assigned caregiver sessions to an external facilitator trained on the same curriculum, allowing teachers to focus on classroom instruction. The remaining 51 schools served as controls.

The three treatment arms test alternative approaches to organizing responsibilities for classroom instruction and parental engagement. T1 preserves the teacher's central role in both domains but risks overburdening teachers in resource-constrained settings. T2 tests whether relieving classroom workload allows teachers to maintain instructional quality while strengthening home–school connections. T3 introduces specialization, with facilitators managing caregiver engagement, testing whether division of labor improves efficiency or weakens complementarities between classroom and home inputs. Comparing outcomes across these models provides direct evidence on whether workload relief and specialization enhance program effectiveness, or whether teacher continuity across both domains yields stronger gains despite the additional burden.

We find improvements across multiple outcomes, including classroom quality, teaching practices, child development, and caregiver engagement. Teacher training successfully enhanced pedagogy, as measured by the Learning Environment and Practices (LEP) Index capturing teacher planning, organization, and use of materials. Child development gains, measured by the Ages and Stages Questionnaire (ASQ), were strongest when teachers worked alone (0.21 SD), despite the additional workload, while the helper model produced the most balanced improvements across classroom quality, child development, and parental engagement, as measured by the Practices of Early Stimulation and Learning (PESL) Index. The facilitator model improved classroom environments but showed no detectable effects on child development, suggesting that separating classroom instruction from parental engage-

ment weakens complementarities between home and school inputs. Overall, the evidence indicates that keeping teachers central to both classroom instruction and caregiver engagement generated the strongest effects on child development, while models adding support produced more balanced but smaller impacts.

We further examine heterogeneity in program impacts by baseline parental engagement, caregiver education, and children's initial developmental scores. Child development gains were largest among families with low baseline parental engagement (0.22 SD), while households with higher baseline engagement saw smaller gains (0.05 SD, n.s.), consistent with ceiling effects. Children starting with the lowest ASQ scores had the most room to improve, and those from less-engaged households showed the largest developmental gains. By contrast, parental engagement improvements showed the opposite pattern: gains were strongest among families already practicing higher baseline stimulation (0.14 SD), though these behavioral changes did not translate into equally large child development improvements because their children were already performing better at baseline. This heterogeneity indicates that the intervention addressed different margins across family types: low-engagement families experienced substantial child development gains despite modest measured changes in home practices, while high-engagement families increased their practices further without generating proportional child development returns.

Two mechanisms help explain these patterns. First, parental participation in caregiver sessions was higher in schools with classroom helpers, indicating that reducing teachers' classroom burden freed capacity to mobilize families. External facilitators in T3 did not generate the same recruitment advantage. Once parents began attending, however, participation rates remained similar across arms, suggesting that initial recruitment rather than ongoing retention was the binding constraint. While T2 generated the strongest parental engagement gains, T1 produced the largest child development effects despite more modest parental engagement, highlighting that teacher continuity across classroom and home domains may matter as much as absolute levels of engagement.

Second, the intervention transformed how home and school inputs interact to produce child development. In control schools, classroom quality and parental engagement functioned as substitutes: children could achieve similar developmental outcomes through either strong classroom environments or engaged parents. Treatment schools created complementarities, where improvements in both classroom quality and parental practices generated larger child development gains than either input alone. This shift from substitutability to complementarity helps explain why integrated mod-

els keeping teachers central to both domains produced stronger developmental outcomes even when parental engagement levels were moderate.

We benchmark our effect sizes against other early childhood programs in LMICs. The intervention improved developmental outcomes by 0.13–0.20 SD (ASQ) and parental engagement by 0.11–0.20 SD (PESL). These magnitudes align closely with estimates from a cross-country synthesis of over 50 studies, which finds average gains of 0.15 SD for cognitive skills and 0.12 SD for socio-emotional skills (Holla et al., 2021), and are comparable to programs combining teacher training with parental engagement, such as Colombia's professional development initiative (0.16–0.30 SD) (Andrew et al., 2024), Rwanda's First Steps parenting program (0.30–0.38 SD) (Justino et al., 2023), and India's urban home-visiting model (0.30 SD) (Andrew et al., 2020). Our estimates fall below more intensive interventions like India's kindergarten subsidy (0.80 SD) (Dean and Jayachandran, 2020) and Tamil Nadu's augmented ICDS program (0.29–0.46 SD) (Ganimian et al., 2024), but exceed programs focused primarily on access expansion, such as The Gambia's preschool initiative (Blimpo et al., 2022), or transfer-based models like Honduras' conditional cash transfer program (0.13 SD) (López Bóo and Creamer, 2019). Taken together, our estimates fall in the middle of the distribution: stronger than access-only or transfer-based programs, weaker than the most intensive or resource-heavy models.

We make three contributions to the literature on early childhood development and human capital formation. First, we extend research on parenting interventions in LMICs, which has shown that providing caregivers with knowledge and skills can improve early stimulation practices (Gertler et al., 2014; Walker et al., 2011; Attanasio et al., 2020). However, these programs face scalability challenges, as most rely on community health workers or paraprofessionals who themselves have limited capacity (Lopez Garcia et al., 2021; Macmillan and Tominey, 2022). By embedding parental education within schools and making teachers central to delivery, we move beyond the traditional distinction between school-based and home-based interventions. Our approach simultaneously strengthens home environments and reinforces school–parent linkages, an integration often missing in stand-alone parenting programs.

Second, we build on recent research examining how staffing constraints shape education program effectiveness. Evidence from India's Integrated Child Development Services shows that adding a half-time worker increased instructional time and improved math and language scores (Ganimian et al., 2024), while studies from Kenya find that teaching assistants can free teacher time for more productive tasks (Duflo et al., 2015; Bold et al., 2018). We extend this literature by testing two alternative staffing

models (helpers inside classrooms versus facilitators outside classrooms) while holding teacher training constant across both. This design allows us to identify which type of staffing support enhances program effectiveness and why. Our results show that only in-class support generated consistent improvements across classroom quality, child development, and caregiver engagement, demonstrating that workload relief must preserve rather than sever the connection between teachers and families.

Third, we contribute to the literature on teacher professional development in early childhood settings by providing large-scale experimental evidence from a government school system. Prior work emphasizes the centrality of teacher–student interactions for learning and socio-emotional growth (Araujo et al., 2016; Wolf et al., 2019), and meta-analyses highlight the effectiveness of teacher coaching in shifting pedagogy (Kraft et al., 2018). However, most evidence from LMICs comes from small pilots or NGO-implemented programs, leaving open questions about whether professional development can work at scale within public systems. Our study demonstrates that a 15-day training program integrated into the government system can substantially improve teacher–student interactions. We document reductions in negative practices such as punitive discipline, alongside increases in warmth, attentiveness, and active engagement. These results show that system-embedded professional development can reshape classroom dynamics even in resource-constrained public schools.

Our findings speak directly to policy debates about scaling early childhood services without sacrificing quality. Evidence from sub-Saharan Africa illustrates the risks of expanding access to low-quality provision (Blimpo et al., 2022), while successful programs in Colombia and India highlight the importance of teacher support and accountability for translating access into learning gains (Andrew et al., 2024). We contribute to this debate by showing that institutional design choices—specifically, whether teachers remain central to both classroom instruction and parental engagement, or whether these functions are divided—shape program effectiveness as much as training content or curriculum design. These results suggest that expanding early childhood programs requires careful attention to how responsibilities are organized to enable coordination between teachers and families.

2 Background

Our intervention takes place in Jhapa District in southeastern Nepal, the country's easternmost district with a population of approximately one million.[1] The district has a young population: over 25 percent

[1] Central Bureau of Statistics Nepal (2021).

of residents are children under age 15, making early childhood investments particularly relevant as this cohort enters school and eventually the labor force. We work across fourteen municipalities within the district, partnering with the local government education system to implement the intervention in 201 government preschools.

Nepal has substantially expanded early childhood education access over the past two decades. ECD enrollment rose from 12 percent in 2000 to 86 percent in 2017, with over 36,000 centers serving nearly one million children (Poudyal et al., 2019; World Bank, 2020). The share of first-graders with prior ECD experience doubled from 33 to 66 percent between 2007 and 2017. Despite constitutional guarantees of early childhood development rights and government commitments to universal access by 2030, this expansion has occurred with minimal resources: ECD spending remains below 0.1 percent of GDP and accounts for only 1–3 percent of the education budget. Quality outcomes reflect this under-investment. In 2014, fewer than one-third of Nepali children aged 36–59 months were developmentally on track in literacy and numeracy, and only 64 percent met benchmarks in socio-emotional development (Poudyal et al., 2019).

Nepal's rapid ECD expansion has occurred within significant resource and capacity constraints. Teachers often lack training in developmentally appropriate practices, and opportunities for professional development are limited. Many parents also have limited knowledge of early stimulation practices. Our intervention addresses these dual constraints by providing structured teacher training while embedding parental education within the school system, leveraging teachers' position as trusted community figures to bridge information gaps in both classroom and home environments.

We implemented the intervention in partnership with Seto Gurans, a local nonprofit specializing in early childhood development, and district education authorities in Jhapa. Seto Gurans recruited and trained facilitators, coordinated with schools, and managed program logistics, while we maintained control over random assignment and data collection. This partnership enabled us to test the intervention within Nepal's government school system while ensuring implementation fidelity across treatment arms.

3 Experimental Design and Method

3.1 Intervention Design

The intervention provides teacher training and parent education to strengthen both classroom instruction and home learning environments, following a two-generation approach (Shonkoff and Fisher, 2013). We partnered with Seto Gurans, a local nonprofit specializing in early childhood development, to design and deliver the curriculum through Nepal's Ministry of Education.[2] Unlike intensive home-visiting models such as Reach Up and Learn (Gertler et al., 2014; Attanasio et al., 2020; Yousafzai et al., 2014), our approach delivers parent education through school-based group sessions led by teachers or school-based facilitators, reducing per-child implementation costs while building on teachers' existing relationships with families.

All teachers received 15 days of training: ten days on developmentally appropriate pedagogy, classroom management, and national ECD standards, followed by five days preparing them to lead parent education sessions. The parent education curriculum comprised 30 topics organized under five components: the role of ECD caregivers, stages of child development, health and nutrition, child protection and care, and early learning. From this curriculum, we selected 18 topics jointly with Seto Gurans as part of the Support for Nepal Early Childhood Development Study. We designed and added two sessions focused specifically on disability awareness and inclusive care. In total, 20 two-hour sessions were delivered over six months in community settings, covering health, nutrition, hygiene, disability awareness, and cognitive and socio-emotional stimulation through play-based activities. To test how implementation models affect program effectiveness, we designed three treatment arms that vary the organization of responsibilities between teachers and support staff while holding training content constant. Full curriculum details appear in Appendix C.

3.2 Treatment Arms

We randomized a sample of 201 schools into the following four arms:

[2]Seto Gurans (www.setogurans.org) is a Nepali organization with extensive experience implementing early childhood development programs.

3.2.1 Treatment Arm 1: Teachers Training

In T1 (50 schools), teachers completed the 15-day training program and then delivered all 20 parent education sessions themselves. Sessions were two hours each, held at the school over six months, and attended by primary caregivers of enrolled children. This model places full responsibility for both classroom instruction and parent engagement on teachers, testing whether they can effectively manage both domains without additional support.

3.2.2 Treatment Arm 2: Teachers Training + Inside Classroom Helper

In T2 (50 schools), teachers received the same training and delivered all 20 parent education sessions, but were supported by an in-class helper who assisted with daily classroom management. The helper's role was to ease the teacher's classroom workload, freeing capacity to maintain instructional quality while engaging parents. This model tests whether reducing classroom burden enables teachers to effectively manage both domains.

3.2.3 Treatment Arm 3: Teachers Training + Outside Classroom Helper

In T3 (50 schools), teachers received the same training but an external facilitator delivered all 20 parent education sessions. Facilitators were recruited and trained specifically for this parent education role, allowing teachers to focus exclusively on classroom instruction. This model introduces specialization: teachers concentrate on pedagogy while facilitators manage parent outreach. The T1–T3 comparison thus tests whether maintaining teacher continuity across classroom and home domains (T1) generates stronger effects than dividing these responsibilities between specialized staff (T3).

3.2.4 Control Group

The 51 schools in this arm served as the control group and did not receive any intervention in the form of teacher training or parental-education sessions - this group was business as usual.

3.3 Conceptual Framework

3.3.1 Production Function and Implementation Costs

We model child development outcomes as produced through classroom quality (Q) and parental engagement (P):

$$D_i = f(Q_i, P_i, X_i; \theta(s)) \tag{1}$$

where D_i represents developmental outcomes for child i, Q_i captures classroom quality (teacher skills, materials, instructional practices), P_i measures parental engagement (stimulation activities and knowledge of child development), X_i denotes child and household characteristics, and $\theta(s)$ governs how inputs combine under staffing model $s \in \{T1, T2, T3\}$.

Both Q and P require teacher time and effort to produce. We represent the "cost" of producing these inputs (in terms of teacher time, attention, and capacity) as $C_Q(s)$ and $C_P(s)$ under each staffing model. This cost framework builds on research emphasizing that teacher workload and capacity constraints are binding in resource-limited settings, where teachers balance instruction with administrative duties, parent outreach, and community engagement (Glewwe and Kremer, 2006; Bold et al., 2018).

Treatment Arm 1 (Teacher Training): The teacher produces both Q and P with full responsibility for classroom instruction and parent education sessions:

$$C_Q(T1) = c_q \tag{2}$$

$$C_P(T1) = c_p \tag{3}$$

subject to a total capacity constraint: $c_q + c_p \leq \bar{C}$, which may bind.

Treatment Arm 2 (Teacher Training + Inside Classroom Helper): An in-class helper reduces the teacher's classroom management burden, lowering the cost of producing Q:

$$C_Q(T2) = c_q - h, \quad h > 0 \tag{4}$$

$$C_P(T2) = c_p \tag{5}$$

This loosens the capacity constraint: $(c_q - h) + c_p < \bar{C}$, potentially allowing the teacher to maintain classroom quality while investing more effectively in parent engagement. Evidence from India

and Kenya suggests that teaching assistants can free teacher time for more productive tasks, though effectiveness depends on integration and accountability (Ganimian et al., 2024; Duflo et al., 2015).

Treatment Arm 3 (Teacher Training + Outside Classroom Helper): Division of labor assigns parent education to an external facilitator, eliminating this responsibility from the teacher's workload:

$$C_Q(T3) = c_q \tag{6}$$

$$C_P(T3) = 0 \text{ for the teacher} \tag{7}$$

However, this model sacrifices a continuity premium arising from the teacher–parent relationship. While economic theory suggests that specialization according to comparative advantage can improve efficiency (Fryer Jr., 2018), in educational settings rigid separation of tasks may reduce effectiveness if classroom instruction and family engagement are inherently complementary activities.

3.3.2 Complementarity versus Substitution and the Continuity Premium

We ask whether classroom quality (Q) and parental engagement (P) function as complements or substitutes in producing child development. The parameter $\theta(s)$ in equation (1) governs this interaction:

$$\frac{\partial^2 f}{\partial Q \partial P} = g(\theta(s)) \tag{8}$$

where the sign and magnitude of the cross-partial derivative capture whether improvements in classroom quality raise (complements) or lower (substitutes) the marginal productivity of parental engagement, and vice versa.

We hypothesize that teacher continuity affects θ:

$$\theta(s) = \begin{cases} \theta_H & \text{if the teacher bridges both classroom and home domains (T1, T2)} \\ \theta_L & \text{if responsibilities are separated across actors (T3)} \end{cases}$$

with $\theta_H > \theta_L$.

This continuity premium arises because the same person delivering both classroom instruction and parent education facilitates information flow (teachers observe children and tailor advice), message alignment (home practices reinforce classroom pedagogy), and trust (established teacher–parent

relationships increase receptivity) (Bergman, 2021; Shonkoff and Fisher, 2013).

In control schools, we hypothesize that Q and P operate primarily as substitutes: families compensate for lower-quality classrooms through stronger home learning investments, and classrooms partially offset weaker home environments. This substitutability reflects households optimizing child development given resource constraints and limited coordination (Cunha and Heckman, 2007). The intervention aims to shift this relationship toward complementarity by coordinating classroom instruction and parent education. Achieving this transformation requires continuity—specifically, the same actor (the teacher) bridging both domains.

When teachers deliver both classroom instruction and parent sessions (T1, T2), complementarity can emerge:

$$\frac{\partial^2 f}{\partial Q \partial P}\Big|_{\theta_H} > 0.$$

When responsibilities are separated (T3), this complementarity weakens or disappears:

$$\frac{\partial^2 f}{\partial Q \partial P}\Big|_{\theta_L} \approx 0.$$

External facilitators lack direct classroom observation and cannot tailor advice to individual children. Their recommendations may also diverge from teachers' pedagogical approaches, creating weaker alignment and reducing the effectiveness of parent outreach.

3.3.3 Testable Predictions

Prediction 1 – Classroom Quality and Child Development Outcomes: We expect differential effects across treatment arms driven by the interaction of capacity constraints, classroom quality production, and complementarity. T2 generates the highest classroom quality (Q) because helper assistance ($c_q - h$) allows teachers to focus on instruction while maintaining capacity for parent engagement. This yields a balanced model with strong Q and P, though child development effects are moderate. T1 operates under the binding constraint $c_q + c_p \leq \bar{C}$, leading teachers to prioritize classroom quality while maintaining continuity with parents (producing moderate P but preserving $\theta = \theta_H$). Because

$$\frac{\partial^2 f}{\partial Q \partial P}\Big|_{\theta_H} > \frac{\partial^2 f}{\partial Q \partial P}\Big|_{\theta_L},$$

strong Q combined with moderate P under high complementarity generates the largest child development gains. T3 allows teachers to focus exclusively on classroom instruction, producing moderate Q, but separation eliminates the continuity premium ($\theta = \theta_L$), preventing child development gains from materializing.

Prediction 2 – Parental Outcomes:

Parental Engagement: T2 generates the strongest improvements in parental engagement because $(c_q - h) + c_p < \bar{C}$ frees teacher capacity to invest in outreach while maintaining continuity and credibility. Both T1 and T3 produce more modest gains. In T1, the binding constraint limits the intensity of engagement despite teacher continuity. In T3, facilitators specialize in parent sessions but lack classroom context and established trust, yielding gains similar to the capacity-constrained T1 model.

Caregiver Stress: Teacher-led models (T1, T2) that maintain the teacher–parent connection should not increase caregiver stress, as teachers provide supportive guidance grounded in classroom observation. The facilitator model (T3), which severs this connection, may increase stress by introducing new expectations without the buffer of established teacher relationships.

Prediction 3 – Shift from Substitution to Complementarity: In control schools, we expect $\frac{\partial^2 f}{\partial Q \partial P} < 0$, indicating substitution. In treatment schools where teachers maintain continuity (T1, T2), we expect $\frac{\partial^2 f}{\partial Q \partial P} > 0$, indicating complementarity. In T3, complementarity should not emerge despite treatment assignment:

$$\frac{\partial^2 f}{\partial Q \partial P}\Big|_{T3} \approx 0.$$

Prediction 4 – Heterogeneity by Baseline Parental Engagement:

Child Development: Low-baseline engagement families should exhibit large gains in child development (compensatory effects, as $\frac{\partial D}{\partial P}\big|_{P_{low}} > \frac{\partial D}{\partial P}\big|_{P_{high}}$), but limited measured changes in parental engagement. High-baseline engagement families should show substantial gains in parental engagement, especially in T2, though child development gains may be smaller due to ceiling effects. This pattern suggests that the intervention operates through distinct channels depending on family context (Attanasio et al., 2020).

Caregiver Stress: For high-baseline engagement families, teacher-led sessions in T1 should reduce stress by providing trusted, personalized support. For low-baseline engagement families, the facilitator model (T3) should increase stress, as new expectations are introduced without the buffer of estab-

lished teacher–parent relationships.

3.4 Sample Selection, Randomization, Training, and Timeline

Sample Selection and Randomization: This study was conducted in Jhapa district, covering schools from all fifteen municipalities. Randomization was done at the school level: 201 schools with Early Childhood Development (ECD) sections were selected and randomly assigned to one of three treatment arms (50 schools each) or to the control group (51 schools). Figure A1 displays the geographic distribution of study schools across the district. All children enrolled in the selected schools were part of the intervention. For data collection, however, budget constraints required sampling a subset of students. Using systematic random sampling, we selected 3,131 children across the study schools. Specifically, every fifth child on the admission list was chosen until 20 children were enrolled per school. In schools with fewer than 20 enrolled children, all ECD students were included in the sample. This approach ensured consistent representation across schools of varying sizes while maintaining feasibility for the endline.

Training: To ensure the effective implementation of this intervention, a structured training program was developed to build the capacity of key stakeholders. The process began with a training pilot conducted from February 3–17, 2022, aimed at refining and contextualizing training materials to improve their relevance and effectiveness. This was followed by a series of targeted capacity-building sessions designed to equip educators and facilitators with the necessary pedagogical expertise and practical skills for successful program delivery. The Early Childhood Development (ECD) teacher training, held from April 3–12, 2022, focused on evidence-based instructional strategies to support early learning, followed by Participatory Education (PE) teacher training from May 29 to June 2, 2022, which provided educators with methodologies for integrating physical activity into the curriculum. To further strengthen program delivery, PE facilitator training was conducted from June 4–10, 2022, followed by classroom helper training from June 12–14, 2022, ensuring adequate support structures were in place across participating schools. This multi-stage training approach aimed to improve instructional quality, ensure consistency in program implementation, and foster a supportive learning environment across the study schools.

Timeline: Figure 1 summarizes the study timeline. Baseline data collection occurred from June 1 to July 22, 2022. The intervention was implemented from July 2022 to January 2023, with schools in treatment arms receiving 20 caregiver education sessions over six months while control schools continued stan-

dard operations. Endline data collection took place from February 9 to March 4, 2023, approximately eight months after baseline.

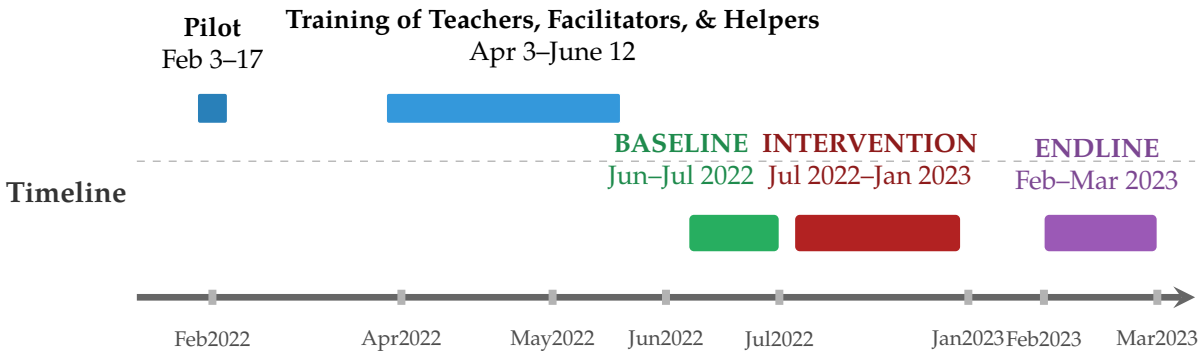


Figure 1: Study Timeline

Notes: This figure shows the study timeline from February 2022 to March 2023. The pilot training occurred in February 2022, followed by training of teachers, facilitators, and helpers from April to June 2022. Baseline data collection took place from June to July 2022, the intervention was implemented from July 2022 to January 2023, and endline data collection occurred from February to March 2023.

4 Data Collection, Outcome Measures, and Validity of the Experimental Design

4.1 Baseline Survey

Baseline data were collected from June to July 2022 through in-person interviews with 3,131 primary caregivers across all 201 schools. Interviews lasted approximately 33 minutes and were conducted by trained enumerators in the local language. Survey modules covered demographic information, household amenities, health status and beliefs (including disability), parental stress, and caregiving practices. From these data we constructed baseline indices for parental engagement (PESL), child development (ASQ), and caregiver stress, which we use to examine heterogeneity in treatment effects by baseline characteristics.

4.2 Learning Environment and Practices Check

At endline, we assessed classroom learning environments across all 201 schools. Teachers completed in-person interviews covering water and sanitation facilities, lesson preparation, pedagogical practices, physical infrastructure, and classroom organization. From these data we constructed the Learn-

ing Environment and Practices (LEP) Index, aggregating indicators of teacher planning, instructional quality, infrastructure, and learning area organization.[3] The LEP Index serves as a first-stage outcome measuring how the intervention affected classroom quality.

4.3 Teacher-Student Interaction

At endline, we complemented survey data with a structured classroom observation to assess teacher–student interactions. Trained enumerators conducted twenty one-minute observations within a two-hour classroom window, each followed by a four-minute period to record assessments. The protocol focused on established best practices for early childhood education. To minimize bias, teachers were not informed about the specific aspects being observed or the evaluation criteria. These observations provide a direct measure of the quality of classroom interactions and the implementation of ECD practices in study schools.

4.4 Parental Education Session Attendance

We recorded attendance at all 20 parental education sessions conducted during the intervention period. Attendance data captures which caregivers participated in each session, allowing us to examine participation patterns across treatment arms and explore attendance as a mechanism linking treatment assignment to outcomes.

4.5 Endline Survey

The endline survey was conducted from February to March 2023 with 3,004 primary caregivers (96% of the baseline sample). Interviews lasted approximately 29 minutes. The questionnaire mirrored the baseline design, covering demographic characteristics, household amenities, health status and beliefs (including disability), parental stress, and caregiving practices. From the endline survey we constructed our primary outcome measures: the Ages and Stages Questionnaire (ASQ) Index for child development, the Practice of Early Stimulation and Learning (PESL) Index for parental engagement practices, and the Caregiver Stress Index.

[3] Appendix E.2 provides details on the items used in constructing the Learning Environment and Practices Index.

4.6 Outcome Measures

First-Stage Outcomes: We assess the intervention's impact on classroom quality through two measures: the Learning Environment and Practices (LEP) Index and the Teacher-Student Interaction Index.

The LEP Index captures classroom features and teacher behaviors reflecting the overall quality of early learning environments. It combines information across four domains: (i) teacher planning and routines, including preparation of lesson plans, following a daily timetable, taking attendance, and maintaining student records; (ii) pedagogical practices, including regular use of stories, poems, songs, learning materials, and inclusive activities for children with disabilities; (iii) physical environment, including cleanliness, safe drinking water, child-friendly toilets, and adequate seating; and (iv) organization of structured learning areas, including designated spaces for language, math, science, creativity, role play, and construction, use of the ECD kit box, and regular supervision by the head teacher. Each component is measured as a binary indicator, with the total score standardized relative to the control group mean to form the index.[4]

To assess teacher-student interactions, we conducted 20 structured one-minute classroom observations within a two-hour window, followed by four-minute recording periods. From these observations we constructed two indices: the Positive Interaction Index and the Negative Interaction Index.

The Positive Interaction Index captures behaviors that foster an engaging and supportive classroom environment, including: greeting children warmly, reading to groups, facilitating discussions, encouraging student participation, providing individual attention, using positive reinforcement, showing enthusiasm and positive affect, and creating opportunities for peer interaction. The Negative Interaction Index measures behaviors that undermine the learning environment, including: harsh or punitive discipline, yelling or threatening, ignoring student questions or needs, showing frustration or impatience, and failing to intervene in conflicts. Both indices are constructed by aggregating observations across the 20 snapshots per classroom and standardizing relative to the control group mean.[5]

Child-Level Outcomes: We measure child development using the Ages and Stages Questionnaire (ASQ), a widely used developmental screening tool adapted for Nepal. The ASQ evaluates children across five domains: communication, gross motor skills, fine motor skills, problem-solving, and personal-social development. Caregivers respond to age-appropriate questions about their child's abilities using a three-point scale: "Yes" (10 points), "Sometimes" (5 points), or "Not Yet" (0 points). We construct six

[4] Appendix E.2 provides details on the items used in constructing the Learning Environment and Practices Index.
[5] Appendix E.2 provides details on the variables used in the construction of the Positive and Negative Interaction Indices.

outcome measures: an aggregate ASQ score summing across all five domains, plus individual scores for each domain. All scores are standardized relative to the control group mean.[6]

We selected the ASQ for four reasons. First, it is cost-effective and does not require specialized training, making it feasible for large-scale implementation in resource-constrained settings. Second, caregiver reporting captures daily observations across multiple contexts such as home, community, and school, thereby providing a more comprehensive picture of child development than single-session direct assessments. Third, reliance on caregiver knowledge rather than direct observation by enumerators reduces observer bias and avoids the artificial testing environment that may affect child performance. Fourth, the ASQ has been validated across diverse populations and used in prior economic evaluations of early childhood programs (Weaver et al., 2025; López Bóo and Creamer, 2019; Araujo et al., 2019; Rubio-Codina et al., 2015). While direct assessments such as the Bayley Scales (Bayley, 2006) or WPPSI offer objective performance measures, their high cost, training requirements, and single-observation limitations constrain their use in large-scale studies in settings like ours.

Caregiver Outcomes: We measure three primary caregiver outcomes: parental engagement practices, caregiver stress, and attitudes toward disability.

Parental Engagement: The Practice of Early Stimulation and Learning (PESL) Index captures whether caregivers engage in developmentally supportive activities with their children, including: reading books, telling stories, singing songs, taking the child outside, playing together, and practicing early learning activities such as naming, counting, or drawing. Each activity is coded as a binary indicator (1 if practiced, 0 otherwise), summed, and standardized relative to the control group. We also construct a PESL Intensity Index that measures the frequency of these activities per week (1–3 days, 4–5 days, or 6+ days), which is similarly standardized.[7]

Caregiver Stress: The Caregiver Stress Index is constructed from a seven-item scale assessing: feeling reinforced or supported by the child, emotional connection with the child, alignment between caregiving expectations and reality, interpretation of the child's behavior, and overall confidence in caregiving abilities. Caregivers respond using a five-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree), with responses summed and standardized relative to the control group.

Attitudes and Beliefs: We measure caregiver attitudes toward disability through four questions covering acceptance, inclusion, and understanding of children with disabilities, summed and standardized to

[6] Appendix E.3 provides details on the variables used to construct ASQ scores across domains.

[7] Appendix E.4 details the variables used in the construction of the PESL and Stress Indices. Appendix E.5 provides details on the PESL Intensity (Intensity of Early Stimulation and Learning) Index.

form the Attitude to Disability Index. We also assess caregiver confidence in supporting children with disabilities through Likert-scale questions on knowledge and ability, forming the Confidence Disability Index. Finally, we measure beliefs about physical discipline through questions on whether caregivers view spanking as necessary or normal and whether they reported using spanking in the previous 14 days.[8]

4.7 Validity of the Experimental Design

4.7.1 Sample Characteristics and Baseline Balance

Table A1 presents descriptive statistics and baseline balance across treatment arms. Column 1 reports overall sample characteristics. Children are on average 48 months old (SD 12), with roughly equal shares of boys and girls. Households average three adults and 1.3 siblings. The primary caregiver is female in 84 percent of cases, with average educational attainment equivalent to six to eight years of schooling. Baseline developmental and caregiver practice measures are standardized around zero with considerable variation: the aggregate ASQ score averages 0.04 (SD 0.95), the PESL Index averages 0.04 (SD 0.99), and the Caregiver Stress Index averages 0.07 (SD 1.07). Attitudes toward disability average slightly negative (–0.10, SD 0.93).

Columns 2–5 report means by treatment arm, while columns 6–8 present tests of mean differences relative to the control group. Demographic characteristics are well balanced across treatment arms. Among baseline outcome measures, we observe small differences in attitudes toward disability and the PESL Index. However, there is no consistent pattern across related measures, and the magnitude of these differences is small relative to overall variation. Given the number of comparisons, such differences likely reflect chance rather than systematic imbalance. Overall, the results are consistent with successful randomization.

4.7.2 Attrition and Compliance

Approximately 4 percent of the baseline sample (127 out of 3,131 children) did not participate in the endline survey. Although attrition is modest, we test whether it is systematically associated with treatment assignment or baseline characteristics in Section 6.7.1. Additionally, 20 of the 150 treatment schools (13 percent) did not implement parental education sessions, resulting in noncompliance. We

[8] Appendix E.5 provides details on the variables used in the construction of the Attitude to Disability and Confidence Disability Indices.

report intent-to-treat estimates in the main analysis, which are unaffected by compliance rates. In Section 6.7.1, we examine baseline characteristics of noncompliant schools to assess whether noncompliance influences the interpretation of our findings.

4.8 Deviations from the Pre-Analysis Plan

Our analysis follows the Pre-Analysis Plan (PAP) registered at the AEA RCT Registry (AEARCTR-0011318), which specified primary outcomes, index construction, multiple-hypothesis corrections, empirical specifications, and attrition protocols, with minor modifications necessitated by field conditions. Most notably, we exclude baseline controls from main specifications because the COVID-19 pandemic disrupted baseline data collection, leaving substantial missing data that would reduce the effective sample by about 30%. Following McKenzie (2012) and recent guidance from Kerwin et al. (2024), we rely on randomization alone to preserve statistical power, as including baseline controls is not required for identification and conventional balance tests are prone to misleading over-rejection. Robustness checks with baseline controls for primary outcomes are reported in Appendix Table A2. We also extend the PAP by disaggregating treatment effects across implementation models ("training only" versus "training plus additional support"), which reveals meaningful heterogeneity (Banerjee et al., 2020), by standardizing variable coding so that positive coefficients uniformly indicate improvements, and by adapting ASQ scoring from narrow age-specific cutoffs to broader developmental categories, enhancing comparability across our heterogeneous sample. Finally, we examine outcomes of direct policy relevance identified during fieldwork, including heterogeneity by age and baseline status. These deviations reflect challenges during the pandemic and field insights while maintaining the integrity of the experimental design.

5 Empirical Specification

We estimate the intent-to-treat (ITT) impact of assigning a trained teacher to ECD classrooms, first by pooling all treatment arms together (Teacher Training, Teacher Training with an Inside Classroom Helper, and Teacher Training with a Facilitator, relative to the control) and then separately by treatment arm. The pooled specification is:

$$Y_{i,s,t=1} = \beta_0 + \beta_1 T_s + \beta_2 Y_{i,s,t=0} + \epsilon_{i,s} \tag{9}$$

where $Y_{i,s,t=1}$ denotes the outcome for individual (student or parent) i in school s measured post-treatment, and $Y_{i,s,t=0}$ is the corresponding baseline value. T_s indicates the treatment assignment of school s: schools were randomly allocated to one of three treatment arms or to a control group. In this specification, $T_s = 1$ if school s was assigned to any of the three treatment arms and $T_s = 0$ if assigned to the control group. $\epsilon_{i,s}$ is the error term. Standard errors are clustered at the school level to account for within-school correlation and because treatment was assigned at that level.[9] The coefficient of interest, β_1, estimates the ITT effect of access to a trained teacher in the ECD classroom.

We next estimate ITT impacts separately by treatment arm, relative to the control. In this specification, indicators are included for assignment to each of the three treatment groups: Teacher Training (T1), Teacher Training with an Inside Classroom Helper (T2), and Teacher Training with a Facilitator (T3). All other covariates are defined as in Equation 9, including baseline outcomes where available. The coefficients β_1, β_2, and β_3 estimate the ITT effects for each treatment arm relative to the control group.

6 Results

We estimate the intervention's effects across three categories of outcomes. First, we examine impacts on classroom environments and teacher-student interactions, which capture how the intervention affected the learning environment. Second, we report effects on our three primary outcomes: child development (ASQ), parental engagement practices (PESL), and caregiver stress. Third, we assess changes in caregiver beliefs and attitudes toward discipline and disability. We present results for pooled treatment effects and separately by treatment arm to compare implementation models.

Because we test multiple hypotheses across outcome domains, we report Benjamini-Hochberg False Discovery Rate (FDR) adjusted p-values alongside unadjusted p-values (Benjamini and Hochberg, 1995) in the main outcomes table (Table 3).

6.1 Impact on Classroom Environments

Table 1 reports treatment effects on the Learning Environment and Practices (LEP) Index. Panel A presents pooled treatment effects, Panel B compares teacher training alone to training combined with any support staff, and Panel C disaggregates by type of support (in-classroom helper versus outside

[9] We report the estimates from Equation 9 with baseline controls $Y_{i,s,t=0}$ in Appendix Table A2.

classroom facilitator).

Pooled treatment schools improved classroom environments by 0.91 SD relative to control schools ($p < 0.01$). This substantial effect indicates that teacher training combined with structured parent education significantly strengthened classroom organization, resources, and instructional practices.

Panel C reveals important heterogeneity across implementation models. The helper model produced the largest gains (1.14 SD, $p < 0.01$), significantly exceeding both the training-only model (0.97 SD, $p < 0.01$) and the facilitator model (0.60 SD, $p < 0.01$). The difference between helper and facilitator models is statistically significant ($p = 0.03$), suggesting that in-class support is particularly effective at improving classroom quality. Helpers assist with classroom management, freeing teachers to focus on instructional organization and child-centered activities. By contrast, facilitators operating outside the classroom had more limited impact on classroom environments, consistent with evidence that additional staff improve outcomes when their roles are directly integrated into classroom routines (Duflo et al., 2015; Ganimian et al., 2024).

Panel B shows that training combined with any form of support (0.87 SD) produces similar effects to training alone (0.97 SD), though both models generate substantial improvements. The helper model's larger effect in Panel C reflects that in-class support specifically enhances classroom quality, while external facilitators do not.

6.2 Improvements in Teacher-Student Interactions

Table 2 reports treatment effects on teacher-student interactions. Panel A shows that treatment classrooms experienced a 0.49 SD increase in positive interactions ($p < 0.01$) and a 0.57 SD decline in negative interactions ($p < 0.01$) relative to controls. Teachers in treatment schools were more likely to read to groups, facilitate conversations, and display warmth, while relying less on harsh discipline and showing reduced disengagement.

Panels B and C reveal that improvements in teacher-student interactions were consistent across all three treatment arms. The uniformity of effects is striking: coefficients range narrowly from 0.44 to 0.53 SD for positive interactions and -0.53 to -0.59 SD for negative interactions across arms, with no statistically significant differences ($p > 0.7$). This pattern contrasts sharply with the LEP Index results, where the helper model produced significantly larger gains than other arms. The consistency of effects across arms, despite their substantial differences in classroom organization (Table 1), demonstrates

that the core teacher training curriculum successfully shifted pedagogical interactions independent of implementation model. The training changed teacher behaviors toward developmentally appropriate interactions regardless of whether teachers received additional classroom support or delegated parent engagement to facilitators.

These findings indicate that professional development focused on child-centered pedagogy can meaningfully improve teacher-student interactions even in resource-constrained settings, with effects concentrated in the domains directly targeted by training (Araujo et al., 2016; Ganimian et al., 2024).

6.3 Impacts on Children's Development and Caregiver Engagement

Table 3 reports treatment effects on our three primary outcomes: child development (ASQ), parental engagement (PESL), and caregiver stress. Panel A presents pooled treatment effects, Panel B compares training alone to training with support, and Panel C disaggregates by type of support.

Pooled treatment improved child development by 0.15 SD ($p = 0.01$, BH-adjusted $p = 0.04$) and parental engagement by 0.13 SD ($p = 0.01$, BH-adjusted $p = 0.04$). The intervention did not increase caregiver stress (0.03 SD, $p = 0.55$, BH-adjusted $p = 0.64$), indicating that the program fostered new developmental activities without overburdening families.

Panel C reveals substantial heterogeneity in child development effects across treatment arms. The training-only arm generated the largest gains (0.21 SD, $p < 0.01$, BH-adjusted $p = 0.00$), despite operating under binding capacity constraints where teachers managed both classroom instruction and parent sessions without additional support. The helper arm produced moderate gains (0.16 SD, $p = 0.02$, BH-adjusted $p = 0.04$), suggesting that in-class support sustained instructional quality while enabling teacher-parent engagement. The facilitator arm showed no detectable effect on child development (0.08 SD, $p = 0.20$, BH-adjusted $p = 0.27$), indicating that separating classroom instruction from parental engagement eliminated developmental gains despite teachers focusing exclusively on instruction.

Parental engagement patterns differ markedly across arms. The helper arm produced the strongest improvements in PESL (0.16 SD, $p < 0.01$, BH-adjusted $p = 0.04$), as reduced classroom burden allowed teachers to invest in parent outreach. Both the training-only arm (0.11 SD, $p = 0.07$, BH-adjusted $p = 0.13$) and facilitator arm (0.11 SD, $p = 0.08$, BH-adjusted $p = 0.13$) produced modest gains of identical magnitude. Despite facilitators specializing in parent sessions, they were no more effective at im-

proving parental engagement than capacity-constrained teachers, suggesting that classroom context and teacher credibility matter more than dedicated time for family outreach.

Caregiver stress showed no significant changes in the training-only (-0.01 SD, $p = 0.82$, BH-adjusted $p = 0.86$) or helper arms (-0.01 SD, $p = 0.86$, BH-adjusted $p = 0.86$), but increased marginally in the facilitator arm (0.11 SD, $p = 0.06$, BH-adjusted $p = 0.12$). The difference between helper and facilitator arms on stress is marginally significant ($p = 0.05$), suggesting that external facilitators introducing new expectations without established teacher relationships may create stress, while teacher-led engagement does not.

These results demonstrate the importance of teacher continuity in generating child development outcomes. The training-only arm's strong performance (0.21 SD), despite having lower classroom quality than the helper model (0.97 vs 1.14 SD LEP) and lower parental engagement than both other models, indicates that maintaining the same person for both classroom instruction and parent sessions creates synergies that outweigh absolute input levels. When teachers bridge both domains, information flows between classroom observations and parent advice, messages reinforce rather than contradict, and established trust enhances receptivity. This continuity creates complementarities between classroom and home environments that amplify developmental gains.

The helper model sustained high classroom quality (1.14 SD) while enabling the strongest parental engagement (0.16 SD), yet produced moderate child development effects (0.16 SD). This suggests that while relieving capacity constraints helps teachers maintain both classroom quality and parental engagement at higher levels, these gains do not fully offset potential dilution effects when helpers manage classroom tasks or the slightly different nature of teacher-student engagement in this model.

The facilitator model's null effects on child development provide the strongest evidence for the importance of continuity. Despite teachers focusing exclusively on classroom instruction, which should maximize classroom quality, separating this from parental engagement eliminated developmental gains. Even moderate classroom quality combined with moderate parental engagement failed to produce child development improvements when different actors delivered each input. External facilitators lacked classroom context to tailor advice, their recommendations may have conflicted with classroom practices, and they lacked the established trust that makes teacher-led engagement effective. When classroom and home inputs operate independently rather than in coordination, their combined effect on child development diminishes substantially. The pattern across arms confirms that maintaining teacher continuity between classroom and home is critical for generating child development

gains.

Our pooled child development effect (0.15 SD) and training-only effect (0.21 SD) are consistent with effect sizes documented in other early childhood interventions combining teacher training with parental engagement. The absence of stress increases distinguishes our intervention from some remote parenting programs that inadvertently raised caregiver anxiety (Arteaga et al., 2025; Noble et al., 2021), indicating that school-based, teacher-led models can enhance parental engagement without overburdening families. Detailed cost-effectiveness analysis and comprehensive benchmarking against comparable programs appear in Section 7.

6.4 Impacts on Caregiver Attitudes Toward Disability

Table 4 reports treatment effects on caregiver beliefs about discipline and attitudes toward disability. Panel A presents pooled effects, Panel B compares training to training with support, and Panel C disaggregates by support type. These outcomes were targeted through two parent education sessions covering disability awareness, identification, and inclusive practices.

The intervention did not change caregiver views on physical discipline (0.00 SD, $p > 0.10$). Caregivers in both treatment and control groups maintained similar beliefs about whether spanking is necessary or acceptable. By contrast, attitudes toward disability improved modestly. The Attitude to Disability Index, which captures views on integrating children with disabilities in school and community settings, increased by 0.09 SD ($p < 0.10$). Caregivers also reported marginally higher confidence in supporting children with disabilities (0.08 SD, $p > 0.10$), though this effect is not statistically significant.

Panel C reveals that the helper model drove these attitudinal changes. The training-only arm shows minimal effects in attitudes toward disability (0.07 SD, $p > 0.10$), while the helper arm generates the strongest improvements in both attitudes toward disability (0.14 SD, $p < 0.05$) and confidence in supporting children with disabilities (0.13 SD, $p < 0.10$). The facilitator arm produces small, statistically insignificant effects (0.05 SD and 0.10 SD respectively). This pattern mirrors the parental engagement results: in-class support enabled teachers to deliver more effective parent education, while capacity-constrained teachers (training-only) and external facilitators both produced weaker attitudinal shifts.

The null effect on discipline beliefs alongside modest improvements in attitudes toward disability

suggests that the intervention was more successful at introducing new frameworks (disability inclusion) than changing deeply entrenched norms (physical discipline). The helper model's advantage in shifting attitudes reinforces that workload relief enables teachers to engage parents more effectively on complex topics requiring sustained discussion and trust-building.

These attitudinal effects are consistent with prior evidence on parenting interventions in LMICs. While intensive two-generation programs like Early Head Start have documented reductions in harsh disciplinary beliefs (Love et al., 2005), shorter parenting components often struggle to shift deeply entrenched norms around physical discipline (Gershoff et al., 2016; Knerr et al., 2013). Our modest improvements in attitudes toward disability (0.09–0.14 SD) align with evidence that parenting interventions can shift beliefs in domains aligned with new information or community values, particularly when delivered through trusted figures like teachers (Gómez-Cotilla et al., 2024; Yousafzai et al., 2014). The pattern of effects (null on discipline, positive on disability) mirrors findings that interventions are more effective at introducing new frameworks than overturning established practices.

6.5 Heterogeneity in Program Impacts

We examine whether program impacts varied by baseline characteristics. Table 5 presents results stratified by baseline parental engagement (PESL). Among caregivers with below-median baseline engagement, the intervention generated substantial improvements in child development (0.22 SD, $p < 0.01$) but modest gains in parental engagement (0.09 SD, $p > 0.10$). In contrast, caregivers with above-median baseline engagement showed strong improvements in parental engagement (0.14 SD, $p < 0.05$) but negligible child development gains (0.05 SD, $p > 0.10$). Children from low-engagement families had the most room to improve developmentally, while high-engagement families were more responsive to teacher outreach and increased their already-strong parenting practices.

Caregiver stress patterns also varied by baseline engagement. Among high-baseline-engagement families, the training-only arm reduced stress substantially (-0.14 SD, $p < 0.05$), indicating that teacher-led parent sessions provided supportive guidance. Among low-baseline-engagement families, the facilitator arm increased stress (0.20 SD, $p < 0.01$), suggesting that external facilitators without established relationships created pressure rather than support for already-disengaged caregivers.

Table 6 examines heterogeneity across baseline outcome quartiles. The main effects in the first three rows show that child development gains are concentrated in the lowest baseline ASQ quartile: the training-only arm produces a 0.31 SD improvement ($p < 0.01$), substantially larger than effects in

higher quartiles. For parental engagement, the opposite pattern emerges: gains are strongest in the highest baseline PESL quartile (0.39 SD for training-only, $p < 0.01$), indicating that already-engaged families were most responsive to teacher outreach. The interaction terms test whether treatment arm effectiveness varies across quartiles. While point estimates suggest some variation, the F-test of interaction terms indicates no statistically significant differences across treatment arms within quartiles ($p > 0.65$ for all outcomes). This confirms that the pattern of effects, large child gains for low-baseline children, large engagement gains for high-baseline families, holds consistently across all three implementation models.

Figure 3 disaggregates impacts by child and household characteristics. Effects are similar for boys and girls on child development, though boys show somewhat larger parental engagement gains. Children of more educated caregivers show larger child development improvements (0.20 vs 0.10 SD for below-median education), while less-educated households show larger parental engagement gains, suggesting that education shapes both the capacity to implement practices and the responsiveness of children to those practices. Household composition shows mixed patterns: households with more adults show slightly larger child development effects but smaller parental engagement gains, possibly reflecting that additional adults support child outcomes directly but reduce the measured role of primary caregiver practices. Effects are broadly similar across households with different numbers of siblings and asset levels, indicating the intervention was effective across diverse family structures and socioeconomic contexts.

These heterogeneity patterns reveal that knowledge transfer was the primary channel through which the intervention operated. Child development gains concentrated among children starting furthest behind (0.31 SD for lowest quartile), while parental engagement gains were strongest among those already practicing stimulation activities (0.39 SD for highest quartile), indicating that responsiveness to new information varied by baseline knowledge. The consistency of these patterns across all three treatment arms indicates that the fundamental mechanism—transferring knowledge about developmentally appropriate practices—operated similarly regardless of implementation model, though the helper model's workload relief enabled teachers to deliver broader and more consistent improvements across diverse family contexts.

6.6 Exploring Potential Mechanisms

We examine three potential mechanisms through which the intervention generated impacts: compensatory and reinforcement pathways, differential parental attendance across treatment arms, and shifts in how home and school environments interact.

The heterogeneity patterns suggest two distinct channels. For households with lower initial engagement, the program operated through a compensatory mechanism: improving classroom quality and providing new information offset deficits in the home environment, leading to substantial child development gains (0.22 SD) despite modest changes in measured parental engagement. For households with higher initial engagement, the program operated through reinforcement: strengthening practices already in place improved measured parental engagement (0.14 SD) but yielded smaller incremental child development gains due to ceiling effects. These contrasting pathways indicate that knowledge transfer was most effective where information gaps were largest, while behavioral responses to new information varied by baseline family context.

Table 7 examines parental attendance at education sessions across treatment arms. Schools with classroom helpers recorded higher parental attendance, averaging 1.24 additional parents per session compared to training-only schools ($p < 0.10$). This attendance advantage persisted throughout the program despite overall declines across all arms (coefficient on session number: -0.03, $p < 0.05$). By contrast, schools with facilitators showed no meaningful difference from training-only schools in either average attendance (0.07, $p > 0.10$) or consistency (-0.02, $p > 0.10$).[10] These patterns indicate that helpers enabled teachers to mobilize parents more effectively by easing classroom workload, while facilitators operating outside the classroom provided no recruitment advantage. Once parents began attending, participation regularity was similar across arms, suggesting that initial mobilization rather than retention drove attendance differences.

Table 8 explores whether home and school environments function as substitutes or complements in producing child development outcomes.[11] In control schools, the interaction between parental engagement (High PESL) and classroom quality (High LEP) is negative and significant (-0.69, $p < 0.01$), indicating substitution: children benefit from strength in one domain compensating for weakness in the other. In treatment schools, this interaction becomes positive (triple interaction coefficient: 0.53, $p < 0.05$), indicating complementarity: children experience the strongest developmental gains when

[10] Consistency is measured as the coefficient of variation in attendance within each school.

[11] We use a two-stage instrumental variables approach to address potential endogeneity. Full methodological details and results appear in Appendix E.6.

both home and school inputs are high. This transformation from substitution to complementarity helps explain why the training-only arm, which maintained teacher continuity between classroom and home, generated the largest child development effects despite operating under capacity constraints. When teachers bridge both domains, improvements in one reinforce rather than replace improvements in the other. For parental engagement and caregiver stress outcomes, the triple interaction terms are small and statistically insignificant, indicating that the shift to complementarity is specific to child development rather than reflecting a general pattern across all outcomes.

These mechanisms explain the pattern of results across treatment arms. The training-only arm's strong child development effects reflect both the compensatory pathway (for low-engagement families) and the shift from substitution to complementarity enabled by teacher continuity. The helper arm's balanced outcomes across domains reflect successful knowledge transfer facilitated by workload relief, enabling teachers to recruit parents effectively and maintain both classroom quality and family engagement. The facilitator arm's null child development effects reflect broken complementarity: even when classroom quality and parental engagement improved independently, the lack of coordination between separate actors delivering each input prevented these improvements from reinforcing each other.

6.7 Robustness Checks

6.7.1 Attrition and Compliance

As noted in Section 4.7.2, 127 of 3,131 children (4%) did not participate in the endline survey. We examine whether attrition threatens internal validity by differing systematically between treatment and control groups or by correlating with baseline characteristics.

Table A4 Panel A reports attrition rates by treatment arm. Attrition in training-only schools (2.9%) and helper schools (2.2%) is slightly higher than in control schools, but differences are not statistically significant ($p = 0.20$ and $p = 0.25$, respectively). Facilitator schools show minimal differential attrition (0.6%, $p = 0.52$). Panel B examines baseline predictors of attrition. Children from lower-asset households were more likely to attrit (-0.003, $p = 0.01$), as were children with less-educated caregivers (-0.004, $p = 0.07$). Female children were less likely to attrit (-0.010, $p = 0.09$). Baseline outcome measures (ASQ, PESL, Stress) do not significantly predict attrition.

Table A5 presents Lee bounds addressing potential selection bias. For child development (ASQ),

the 95% confidence interval ranges from 0.066 to 0.229 SD, with the lower bound remaining positive. This indicates that even under worst-case assumptions about selective attrition, treatment effects on child development remain positive. For parental engagement (PESL), bounds range from -0.116 to 0.121 SD, crossing zero. However, the low overall attrition rate (4%) and absence of baseline outcome differences between attritors and non-attritors suggest limited scope for substantial bias. For caregiver stress, bounds range from -0.119 to 0.051 SD.

We also examine non-compliance in program implementation. Table A6 shows that 20 of 150 treatment schools (13%) did not conduct parental education sessions: 6 training-only, 9 helper, and 5 facilitator schools. Children in non-compliant schools lived in larger households (0.313 additional adults, $p = 0.03$), were less likely to have female caregivers (-0.067, $p = 0.05$), and had less-educated caregivers (-0.336 years, $p = 0.01$). Importantly, baseline outcome measures (ASQ, PESL, Stress) did not differ significantly between compliant and non-compliant schools ($p > 0.67$ for all outcomes). Our intent-to-treat estimates include all treatment schools regardless of implementation, providing conservative estimates of program effectiveness.

6.7.2 Measurement Validity and Social Desirability Bias

Several key outcomes rely on self-reports: caregivers report child development (ASQ) and their own engagement practices (PESL), while teachers report classroom practices (LEP). We address concerns about social desirability bias through multiple validation checks.

First, we collected direct observations independent of self-reports. Trained enumerators conducted twenty one-minute structured observations per classroom to construct the Positive and Negative Teacher-Student Interaction Indices. Teachers were not informed of specific scoring criteria. These directly-observed interaction measures show large treatment effects (0.49 SD for positive interactions, -0.57 SD for negative interactions), confirming that behavioral changes occurred beyond self-reports. The importance of these observed interactions is well-established: seminal frameworks such as CLASS emphasize that emotionally supportive and instructionally rich interactions are strong predictors of child development (Pianta et al., 2008), while empirical evidence demonstrates that variation in teacher quality, as captured through observed interaction quality, has large effects on cognitive and socio-emotional development (Araujo et al., 2016; Andrew et al., 2024). Our direct observations of improved teacher-student interactions provide independent validation that the intervention changed actual classroom practices, not just self-reported perceptions.

Second, we examine whether self-reported outcomes align with observed changes. Child development gains (ASQ, caregiver-reported) are strongest in treatment arms that produced the largest improvements in directly-observed classroom quality (LEP) and teacher-student interactions. If reporting bias drove results, we would not expect this alignment between caregiver reports and independently-observed classroom changes.

Third, heterogeneity patterns are inconsistent with uniform social desirability effects. Child development gains concentrate among low-baseline-engagement families (0.22 SD), while parental engagement gains are strongest among high-baseline-engagement families (0.14 SD). These opposing patterns across subgroups suggest distinct behavioral mechanisms rather than common reporting bias affecting all measures similarly.

Fourth, we benchmark against experimenter demand effects documented in prior research. Studies estimating demand effects in field experiments find magnitudes ranging from 0.05 to 0.15 SD (de Quidt et al., 2018, 2019; Winichakul et al., 2024). Our child development effects (0.15–0.21 SD) exceed this range for training-only and helper models, suggesting that genuine behavioral changes explain the majority of observed effects even if some reporting bias is present.

These validation checks indicate that while self-reports may contain some measurement error, the pattern of results, particularly the alignment between caregiver-reported child outcomes and independently-observed classroom improvements, is most consistent with genuine behavioral changes rather than reporting artifacts.

7 Benchmarking Program Impacts and Costs

We assess the cost-effectiveness of our intervention by comparing implementation costs to developmental impacts. The intervention cost approximately USD 47,500 (NPR 6 million) to implement across 150 schools, equivalent to USD 317 (NPR 40,000) per school or USD 21 (NPR 2,600) per child. These costs covered the complete program: twenty parenting education sessions, teacher training (15 days), remuneration for 50 facilitators and 50 classroom helpers, coaching and mentoring, and monitoring and supervision.

Following standard cost-effectiveness calculations for education interventions (Ganimian et al., 2024), we estimate benefit-cost ratios (BCR) by comparing implementation costs to projected lifetime earnings gains from improved child development. Our observed effect sizes (0.15 SD pooled child

development (ASQ), 0.21 SD for the training-only model, and 0.16 SD for the helper model) translate to BCRs in the range of 11–14 under conservative assumptions about the relationship between early childhood test scores and adult earnings. Under these estimates, even modest tax recovery (approximately 8 percent of projected lifetime earnings gains) would fully offset program delivery costs.

Our effect sizes align with the broader distribution of early childhood interventions in LMICs. Meta-analytic evidence on ECD programs documents average effects of approximately 0.15 SD on cognitive outcomes and 0.12 SD on socio-emotional development (Jeong et al., 2021; McCoy et al., 2017). Global syntheses of education interventions report median effects of 0.10–0.15 SD across more than a thousand studies, emphasizing that even effects in this range can be meaningful when achieved at scale (Evans and Yuan, 2022). Our pooled effect (0.15 SD) falls squarely within this range, while our training-only arm's effect (0.21 SD) approaches magnitudes found in well-resourced interventions despite operating in resource-constrained rural settings. This positions our intervention between intensive urban programs like Boston Pre-K (0.45–0.62 SD) (Weiland and Yoshikawa, 2013) and lower-intensity parenting interventions (0.11 SD) (Gómez-Cotilla et al., 2024).

Table 10 benchmarks our intervention against comparable early childhood programs. Our cost per child (USD 21) and BCR (11–14) compare favorably to other teacher training and parenting interventions. Tamil Nadu's ICDS facilitator program achieved a BCR of 13:1 with effects of 0.29–0.46 SD, though at substantially higher cost per child (Ganimian et al., 2024). Rwanda's "First Steps" parenting program achieved effects of 0.30–0.38 SD at USD 2 per caregiver session, demonstrating exceptional cost-effectiveness for a pure parenting model (Justino et al., 2023). Colombia's "Hogares Infantiles Mejorado" package cost USD 35 upfront plus USD 13 annually per child and produced 0.16 SD average gains (0.30 SD for poorest households) (Andrew et al., 2024). Ghana's preschool quality package achieved 0.11–0.18 SD effects with BCRs of 11–103:1 (Wolf et al., 2019), while Peru's pedagogy reform generated 0.19 SD gains with BCRs of 29–88 (Gallego et al., 2021). Bangladesh's preschool initiative produced 0.25–0.37 SD gains with BCRs of 3–28 (Spier et al., 2020), and Indonesia's ECE expansion achieved 0.16 SD improvements in socio-emotional development (Brinkman et al., 2017).

By contrast, more intensive interventions achieve larger effects at substantially higher costs. India's kindergarten subsidy program produced 0.80 SD gains but required extensive per-child subsidies (Dean and Jayachandran, 2020). Honduras' conditional cash transfer cost nearly USD 500 per household annually yet yielded only 0.13 SD improvements (López Bóo and Creamer, 2019), resulting in substantially lower cost-effectiveness. The teaching-assistant variant of Colombia's Hogares Infantiles

cost USD 300 per child with no measurable effects, highlighting that additional personnel do not guarantee improved outcomes without appropriate implementation design. The Gambia's expansion and quality enhancement trial produced null impacts despite substantial investments (Blimpo et al., 2022).

Three features distinguish our cost structure from higher-cost interventions. First, we leveraged existing government school infrastructure and teachers rather than creating parallel delivery systems, substantially reducing fixed costs. Second, the helper and facilitator models tested whether modest additional personnel (USD 5–8 per child) could enhance effectiveness, finding that in-class helpers improved outcomes while external facilitators did not—a critical design insight for cost-conscious scaling. Third, the training-only model's strong performance (0.21 SD at lowest cost) demonstrates that meaningful impacts are achievable through teacher capacity-building alone when teachers maintain continuity between classroom instruction and parental engagement.

These results indicate that our intervention represents a cost-effective approach to improving early childhood outcomes in resource-constrained settings. The combination of moderate costs (USD 21 per child), meaningful effect sizes (0.15–0.21 SD), and favorable benefit-cost ratios (11–14:1) positions this model as a viable candidate for scaling within government education systems. Our effects place us in the mid-range of documented interventions: more effective than access-only or transfer programs, comparable to combined teacher training and parenting interventions, and approaching the effectiveness of some intensive programs but at a fraction of the cost. The training-only variant offers particular promise for contexts with severe budget constraints, while the helper model provides a pathway for incremental quality improvements where modest additional resources are available.

8 Conclusion

This paper provides experimental evidence on how organizational design shapes the effectiveness of early childhood interventions delivered through government systems. We evaluate three implementation models in Nepal's government schools, finding that program effectiveness depends critically on how responsibilities are structured among frontline workers. Maintaining teacher continuity between classroom instruction and parental engagement generated the largest child development gains, while models that separated these responsibilities, delegating family outreach to external facilitators, produced null effects despite improving classroom environments. These results demonstrate that the organization of service delivery matters as much as program content for achieving developmental

impacts at scale. This underscores the importance of design in contexts where frontline workers face competing demands and institutional constraints (Duflo et al., 2015; Bold et al., 2018).

The central mechanism involves transforming how home and school environments interact. In control schools, classroom quality and parental engagement function as substitutes—families compensate for weak schools, or schools support children from less-engaged families. When teachers coordinate both domains, this relationship shifts toward complementarity: improvements in one reinforce rather than replace improvements in the other. This transformation occurs through teacher continuity, which enables information flow between classroom observations and parent guidance, ensures message consistency, and leverages established trust. The intervention disproportionately benefited households with lower baseline engagement, with child development gains concentrated among children starting furthest behind, indicating potential to narrow rather than widen developmental inequalities.

These findings contribute to understanding human capital formation in resource-constrained settings. Rapid expansions in access to education and social services often fail to generate commensurate gains when implementation design is neglected (Pritchett, 2013; Glewwe and Kremer, 2006). We show that relatively short teacher training, when embedded in structures that support coordination between home and school, can shift both classroom practices and family behaviors, underscoring that frontline capacity and the organization of support within public systems are central determinants of effectiveness (Mbiti et al., 2019). The heterogeneity patterns, largest effects among disadvantaged families, align with frameworks emphasizing that early investments become more productive when they coordinate across environments rather than compete (Johnson and Jackson, 2019). By demonstrating that home-school relationships can shift from substitutes to complements through organizational design, this study identifies a mechanism relevant across sectors where household investments interact with public provision (Heckman and Karapakula, 2019; Pianta et al., 2008). This parallels recent evidence from India showing that pedagogical supports are most effective when they complement rather than substitute for teachers' knowledge (Ganimian et al., 2025).

For policymakers, these results offer both opportunities and cautions. On one hand, relatively low-cost training models embedded within government infrastructure can shift classroom practices, strengthen teacher-parent connections, and generate meaningful developmental gains without increasing caregiver stress. Such approaches are particularly relevant for settings where access has expanded faster than quality (Andrew et al., 2024), and where fiscal constraints limit the feasibility of intensive external interventions. On the other hand, our evidence suggests caution when parental en-

gagement is organized through facilitators disconnected from classrooms. While specialization may appear efficient, its effectiveness depends on close coordination with classroom instruction and established accountability structures (Fryer Jr., 2018). Without such integration, the complementarities between home and school environments weaken, limiting benefits to children. The contrast across implementation models therefore highlights that organizational design, how roles are structured and coordinated within government systems, is as important as curriculum or pedagogy. Strengthening state capacity in early childhood requires not only increasing personnel but also embedding them effectively within structures of accountability and support (Ganimian et al., 2024).

Several limitations warrant attention. The follow-up period is eight months post-intervention, and future work should assess whether improvements persist into primary school and beyond. Our outcome measures rely substantially on caregiver and teacher reports, which are widely validated but could be complemented with direct assessments (Araujo et al., 2019). The program was implemented in one district; replication across institutional contexts would strengthen external validity. Longer-term follow-up and replication will be critical to assess whether the short-run gains observed here translate into lasting improvements in schooling trajectories, labor market outcomes, and intergenerational mobility (Elango et al., 2015; Garcıa et al., 2023).

As investments in early childhood expand globally, achieving lasting improvements in human capital requires grappling with the realities of implementation within public systems (Heckman, 2006; List et al., 2021). Our results underscore that effectiveness depends not only on what programs deliver but on how delivery is organized, particularly whether structures maintain or sever the connections between classroom instruction and family engagement. By embedding reforms within government infrastructure, targeting both teachers and parents, and addressing workload constraints, interventions can generate meaningful developmental gains at modest cost while remaining scalable. The most durable gains are likely to come from two-generation models that explicitly strengthen home-school linkages, ensuring that teachers and caregivers work in tandem rather than isolation. While the effect sizes observed here align with the global distribution of education program impacts (Evans and Yuan, 2022), they were achieved at low cost through existing public systems serving disadvantaged populations. Together with evidence of dynamic complementarity (Johnson and Jackson, 2019), this suggests that even moderate early gains, when embedded within public systems, can yield substantial long-term benefits. We hope these findings inform both the design of future interventions and the broader research agenda on strengthening state capacity to deliver effective early childhood services at scale.

References

Alan, S. (2025): "Shaping Society's Character: The Role of Schools in Developing Social and Emotional Skills," in *Handbook of Culture*, ed. by B. Enke, P. Giuliano, N. Nunn, and L. Wantchekon, Elsevier, forthcoming.

Andrew, A., O. Attanasio, B. Augsburg, M. Day, S. Grantham-McGregor, C. Meghir, F. Mehrin, S. Pahwa, and M. Rubio-Codina (2020): "Effects of a scalable home-visiting intervention on child development in slums of urban India: evidence from a randomised controlled trial," *Journal of child psychology and psychiatry*, 61, 644–652.

Andrew, A., O. P. Attanasio, R. Bernal, L. C. Sosa, S. Krutikova, and M. Rubio-Codina (2024): "Preschool quality and child development," *Journal of Political Economy*, 132, 2304–2345.

Araujo, M. C., P. Carneiro, Y. Cruz-Aguayo, and N. Schady (2016): "Teacher Quality and Learning Outcomes in Kindergarten," *The Quarterly Journal of Economics*, 131, 1415–1453.

Araujo, M. C., M. Dormal, and N. Schady (2019): "Childcare quality and child development," *Journal of Human Resources*, 54, 656–682.

Arteaga, I., A. de Barros, and A. J. Ganimian (2025): "The challenges of scaling up effective child-rearing practices using technology in developing settings: Experimental evidence from India," *Journal of Research on Educational Effectiveness*, 1–19.

Attanasio, O., S. Cattan, E. Fitzsimons, C. Meghir, and M. Rubio-Codina (2020): "Estimating the Production Function for Human Capital: Results from a Randomized Controlled Trial in Colombia," *American Economic Review*, 110, 48–85.

Banerjee, A., E. Duflo, A. Finkelstein, L. F. Katz, B. A. Olken, and A. Sautmann (2020): "In praise of moderation: Suggestions for the scope and use of pre-analysis plans for rcts in economics," Tech. rep., National Bureau of Economic Research.

Bayley, N. (2006): "Bayley scales of infant and toddler development," .

Benjamini, Y. and Y. Hochberg (1995): "Controlling for False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of Royal Statistical Society, Series B (Methodological)*, 57(1), 289–300.

Bergman, P. (2021): "Parent-child information frictions and human capital investment: Evidence from a field experiment," *Journal of political economy*, 129, 286–322.

Black, M. M., S. P. Walker, L. C. H. Fernald, C. T. Andersen, A. M. DiGirolamo, C. Lu, D. C. McCoy, G. Fink, Y. R. Shawar, J. Shiffman, A. E. Devercelli, Q. T. Wodon, E. Vargas-Baron, and S. Grantham-McGregor (2017): "Early Childhood Development Coming of Age: Science Through the Life Course," *The Lancet*, 389, 77–90.

Blimpo, M. P., P. Carneiro, P. Jervis, and T. Pugatch (2022): "Improving access and quality in early childhood development programs: Experimental evidence from the Gambia," *Economic Development and Cultural Change*, 70, 1479–1529.

Bold, T., M. S. Kimenyi, G. Mwabu, A. Ng'ang'a, and J. Sandefur (2018): "Experimental evidence on scaling up education reforms in Kenya," *Journal of Public Economics*, 168, 1–20.

Brinkman, S. A., A. Hasan, H. Jung, A. Kinnell, and M. Pradhan (2017): "The impact of expanding access to early childhood education services in rural Indonesia," *Journal of Labor Economics*, 35, S305–S335.

Brutti, Z. and D. Montolio (2021): "Preventing criminal minds: Early education access and adult offending behavior," *Journal of Economic Behavior Organization*, 191, 97–126.

Campbell, F., G. Conti, J. Heckman, S. Moon, R. Pinto, E. Pungello, and Y. Pan (2014): "Early Childhood Investments Substantially Boost Adult Health," *Science (New York, N.Y.)*, 343, 1478–85.

Chen, S. and S. Wolf (2021): "Measuring the Quality of Early Childhood Education in Low- and Middle-Income Countries," *Frontiers in Psychology*, 12, 774740.

Coffey, D., M. Geruso, and D. Spears (2017): "Cognitive science in the field: A preschool intervention durably enhances intuitive but not formal mathematics," *Psychology*.

Cunha, F. and J. Heckman (2007): "The Technology of Skill Formation," *American Economic Review*, 97, 31–47.

de Quidt, J., J. Haushofer, and C. Roth (2018): "Measuring and bounding experimenter demand," *American Economic Review*, 108, 3266–3302.

de Quidt, J., L. Vesterlund, and A. J. Wilson (2019): "Experimenter Demand Effects," in *Handbook of Research Methods and Applications in Experimental Economics*, ed. by A. Schram and A. Ule, Edward Elgar Publishing, chap. 20, 384–400.

Dean, J. T. and S. Jayachandran (2020): "Attending kindergarten improves cognitive development in India, but all kindergartens are not equal," Tech. rep., mimeo.

Duflo, E., P. Dupas, and M. Kremer (2015): "School governance, teacher incentives, and pupil–teacher ratios: Experimental evidence from Kenyan primary schools," *Journal of Public Economics*, 123, 92–110.

Elango, S., J. L. García, J. J. Heckman, and A. Hojman (2015): "Early childhood education," in *Economics of Means-Tested Transfer Programs in the United States, Volume 2*, University of Chicago Press, 235–297.

Evans, D. K. and F. Yuan (2022): "How Big Are Effect Sizes in International Education Studies?" *Educational Evaluation and Policy Analysis*, 44, 532–540.

Fryer Jr., R. G. (2018): "The "Pupil" Factory: Specialization and the Production of Human Capital in Schools," *American Economic Review*, 108, 616–56.

Gallego, F. A., E. Näslund-Hadley, and M. Alfonso (2021): "Changing pedagogy to improve skills in preschools: experimental evidence from Peru," *The World Bank Economic Review*, 35, 261–286.

Ganimian, A. J., I. M. Mbiti, and A. Mishra (2025): "Leveraging Complementarities between Teachers' Content Knowledge and Pedagogical Supports: Experimental Evidence from India," Unpublished manuscript.

Ganimian, A. J., K. Muralidharan, and C. R. Walters (2024): "Augmenting State Capacity for Child Development: Experimental Evidence from India," *Journal of Political Economy*, 132, 1565–1602.

Garcia, J. L., J. J. Heckman, and V. Ronda (2023): "The lasting effects of early childhood education on promoting the skills and social mobility of disadvantaged African Americans and their children," *Journal of Political Economy*, 131, 1477–1506.

Gershoff, E. T., A. Grogan-Kaylor, J. E. Lansford, L. Chang, A. Zelli, K. Deater-Deckard, and K. A. Dodge (2016): "Parent discipline practices in an international sample: Associations with child behaviors and moderation by perceived normativeness," *Child Development*, 87, 487–502.

Gertler, P., J. Heckman, R. Pinto, A. Zanolini, C. Vermeersch, S. Walker, S. M. Chang, and S. Grantham-McGregor (2014): "Labor market returns to an early childhood stimulation intervention in Jamaica," *Science*, 344, 998–1001.

Glewwe, P. and M. Kremer (2006): "Schools, Teachers, and Education Outcomes in Developing Countries," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek and F. Welch, Elsevier, vol. 2, 945–1017.

Gómez-Cotilla, R., M. d. l. Á. López-de Uralde-Selva, and L. Valero-Aguayo (2024): "Efficacy of Early Intervention Programmes: Systematic Review and Meta-analysis," *Psicología Educativa*, 30, 1–10.

Heckman, J. J. (2006): "Skill formation and the economics of investing in disadvantaged children," *Science*, 312, 1900–1902.

Heckman, J. J. and G. Karapakula (2019): "Intergenerational and Intragenerational Externalities of the Perry Preschool Project," Working Paper 25889, National Bureau of Economic Research.

Holla, A., M. Bendini, L. Dinarte, and I. Trako (2021): "Is investment in preprimary education too low," *World Bank Policy Research Working Paper*.

Jakiela, P., O. Ozier, L. C. Fernald, and H. A. Knauer (2024): "Preprimary education and early childhood development: Evidence from government schools in rural Kenya," *Journal of Development Economics*, 171, 103337.

Jeong, J., E. E. Franchett, C. V. Ramos de Oliveira, K. Rehmani, and A. K. Yousafzai (2021): "Parenting interventions to promote early child development in the first three years of life: A global systematic review and meta-analysis," *PLOS Medicine*, 18, 1–51.

Johnson, R. C. and C. K. Jackson (2019): "Reducing Inequality through Dynamic Complementarity: Evidence from Head Start and Public School Spending," *American Economic Journal: Economic Policy*, 11, 310–349.

Justino, P., M. Leone, P. Rolla, M. Abimpaye, C. Dusabe, M. D. Uwamahoro, and R. Germond (2023): "Improving parenting practices for early child development: Experimental evidence from Rwanda," *Journal of the European Economic Association*, 21, 1510–1550.

Kerwin, J., N. Rostom, and O. Sterck (2024): "Striking the right balance: Why standard balance tests over-reject the null, and how to fix it," .

Knerr, W., F. Gardner, and L. Cluver (2013): "Improving positive parenting skills and reducing harsh and abusive parenting in low-and middle-income countries: A systematic review," *Prevention Science*, 14, 352–363.

Kraft, M. A., D. Blazar, and D. Hogan (2018): "The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence," *Review of Educational Research*, 88, 547–588.

Lee, D. S. (2009): "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies*, 76, 1071–1102.

List, J. A., D. Suskind, and L. H. Supplee (2021): "The scale-up effect in early childhood and public policy," *The Scale-Up Effect in Early Childhood and Public Policy: Why Interventions Lose Impact at Scale and What We Can Do About It Routledge*.

López Bóo, F. and J. Creamer (2019): "Cash, conditions and child development: Experimental evidence from a cash transfer in honduras," Tech. rep., IZA Discussion Papers.

Lopez Garcia, I., U. Saya, and J. Luoto (2021): "Cost-effectiveness and economic returns of group-based parenting interventions to promote early childhood development: Results from a randomized controlled trial in rural Kenya," *PLOS Medicine*, 18, e1003746.

Love, J. M., E. E. Kisker, C. M. Ross, P. Z. Schochet, J. Brooks-Gunn, D. Paulsell, K. Boller, J. Constantine, C. Vogel, A. S. Fuligni, and C. Brady-Smith (2005): "Making a Difference in the Lives of Infants and Toddlers and Their Families: The Impacts of Early Head Start," Tech. rep., U.S. Department of Health and Human Services, Administration for Children and Families, Washington, DC.

Macmillan, L. and E. Tominey (2022): "Parental inputs and socio-economic gaps in early child development," *Journal of Population Economics*, 36.

Masino, S. and M. Niño-Zarazúa (2016): "What works to improve the quality of student learning in developing countries?" *International Journal of Educational Development*, 48, 53–65, aid, Education Policy, and Development.

Mbiti, I., K. Muralidharan, M. Romero, Y. Schipper, C. Manda, and R. Rajani (2019): "Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania," *The Quarterly Journal of Economics*, 134, 1627–1673.

McCoy, D. C., H. Yoshikawa, K. M. Ziol-Guest, G. J. Duncan, H. S. Schindler, K. Magnuson, R. Yang, A. Koepp, and J. P. Shonkoff (2017): "Impacts of Early Childhood Education on Medium- and Long-Term Educational Outcomes," *Educational Researcher*, 46, 474–487.

McKenzie, D. (2012): "Beyond Baseline and Follow-up: The Case for More T in Experiments," *Journal of Development Economics*, 99, 210–221.

Noble, K. G., K. Magnuson, L. A. Gennetian, G. J. Duncan, H. Yoshikawa, N. A. Fox, and S. Halpern-Meekin (2021): "Baby's first years: design of a randomized controlled trial of poverty reduction in the United States," *Pediatrics*, 148.

Phillips, D. A. and J. P. Shonkoff (2000): "From neurons to neighborhoods: The science of early childhood development," .

Pianta, R. C., K. M. La Paro, and B. K. Hamre (2008): *Classroom Assessment Scoring System (CLASS) Manual, Pre-K*, Baltimore, MD: Paul H. Brookes Publishing Co.

Poudyal, N., M. Banskota, and D. Khadka (2019): "Disability in Nepal," Report, UNICEF.

Pritchett, L. (2013): *The Rebirth of Education: Schooling Ain't Learning*, Washington, DC: CGD Books.

Rubio-Codina, M., O. Attanasio, C. Meghir, N. Varela, and S. Grantham-McGregor (2015): "The socioeconomic gradient of child development: Cross-sectional evidence from children 6–42 months in Bogota," *Journal of Human Resources*, 50, 464–483.

Shonkoff, J. P. and P. A. Fisher (2013): "Rethinking evidence-based practice and two-generation programs to create the future of early childhood policy," *Development and Psychopathology*, 25, 1635–1653.

Spier, E., K. Kamto, A. Molotsky, A. Rahman, N. Hossain, Z. Nahar, and H. Khondker (2020): "Bangladesh early years preschool program impact evaluation," *World Bank, Washington, DC*.

Walker, S., S. Chang, M. Vera-Hernández, and S. Grantham-Mcgregor (2011): "Early Childhood Stimulation Benefits Adult Competence and Reduces Violent Behavior," *Pediatrics*, 127, 849–57.

Weaver, J., S. Sukhtankar, P. Niehaus, and K. Muralidharan (2025): "Maternal Cash Transfers for Gender Equity and Child Development: Experimental Evidence from India," *Working Paper*.

Weiland, C. and H. Yoshikawa (2013): "Impacts of a Prekindergarten Program on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills," *Child Development*, 84, 2112–2130.

Winichakul, K. P., G. Lezama, P. Mustafi, M. Lepper, A. Wilson, D. Danz, and L. Vesterlund (2024): "The effect of experimenter demand on inference," .

Wolf, S., J. L. Aber, J. R. Behrman, and E. Tsinigo (2019): "Experimental impacts of the Quality Preschool for Ghana interventions on teacher professional well-being, classroom quality, and children's school readiness," *Journal of Research on Educational Effectiveness*, 12, 10–37.

World Bank (2020): "Nepal Early Childhood Education Diagnostic," Tech. rep., Washington, D.C.

World Bank and UNICEF (2020): "Costing Study on Early Childhood Education and Development (ECED) in Nepal: A Case for Investment in ECED," Tech. rep., World Bank, Washington, DC.

Yousafzai, A. K., M. A. Rasheed, A. Rizvi, R. Armstrong, and Z. A. Bhutta (2014): "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," *The Lancet*, 384, 1282–1293.

Tables

Table 1: Effect on Learning Environment and Practices Index

	Learning Environment & Practices Index
	(1)
Panel A: Pooled Treatment Arms	
Any Treatment	0.91***
	(0.18)
Observations	201
R^2	0.27
Panel B: Combining Training and Help	
Training Only	0.97***
	(0.23)
Training + Any Help	0.87***
	(0.20)
Observations	201
R^2	0.27
Panel C: Individual Treatment Arms	
Training Only	0.97***
	(0.23)
Training + Helper	1.14***
	(0.24)
Training + Facilitator	0.60***
	(0.23)
Observations	201
R^2	0.29
Helper = Facilitator (p-value)	0.03

Notes: The sample comprises all schools that participated in the intervention. The table reports treatment effects on the *Learning Environment and Practices Index* (see Appendix Section E.2). This index aggregates 20 checklist items capturing lesson planning, pedagogical practices, physical infrastructure, and organization of structured learning areas. Each item was coded 1 if present and 0 if absent; totals are standardized relative to the control group. All regressions include municipality fixed effects. *Any Treatment* equals one if a school received *Training Only*, *Training + Helper*, or *Training + Facilitator*, and zero otherwise. Standard errors are clustered at the school level and reported in parentheses. ***, **, * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 2: Effects on Teacher–Student Interaction Quality

	Positive Interaction Index	Negative Interaction Index
	(1)	(2)
Panel A: Pooled Treatment Arms		
Any Treatment	0.49***	-0.57***
	(0.17)	(0.15)
Observations	3,004	3,004
R^2	0.15	0.20
Panel B: Combining Training and Help		
Training Only	0.55***	-0.59***
	(0.21)	(0.17)
Training + Any Help	0.46**	-0.56***
	(0.18)	(0.16)
Observations	3,004	3,004
R^2	0.15	0.20
Panel C: Individual Treatment Arms		
Training Only	0.55***	-0.59***
	(0.21)	(0.17)
Training + Helper	0.44**	-0.53***
	(0.19)	(0.18)
Training + Facilitator	0.48**	-0.58***
	(0.20)	(0.17)
Observations	3,004	3,004
R^2	0.15	0.20
Helper = Facilitator (p-value)	0.84	0.71

Notes: The table reports results on teacher–student interactions estimated from regression equation (9). The outcome variables are the *Positive Interaction Index* and the *Negative Interaction Index* (see Appendix Section E.2), each measured as the frequency of interactions observed across 20 structured rounds. Positive interactions include reading to a group, teacher or student singing, circle-time activities, back-and-forth discussions, engaging in pretend play, telling children what to do, calmly negotiating peer conflict, and smiling. Negative interactions capture negative discipline, yelling, not engaging with children, talking with another adult, or being out of the classroom. A higher score reflects greater frequency of the specified type of interaction. All regressions include municipality fixed effects. *Any Treatment* equals one if a school received *Training Only*, *Training + Helper*, or *Training + Facilitator*, and zero otherwise. Standard errors are clustered at the school level and reported in parentheses. The data comes from the Endline Survey, which includes all 3,004 student respondents. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 3: Treatment Effects on Child Development and Caregiver Outcomes

	ASQ Index	PESL Index	Caregiver Stress Index
	(1)	(2)	(3)
Panel A: Pooled Treatment Arms			
Any Treatment	0.15***	0.13**	0.03
	(0.05)	(0.05)	(0.05)
BH-Corrected P-Values	[0.04]	[0.04]	[0.64]
Observations	3,004	3,004	3,004
R^2	0.12	0.08	0.36
Panel B: Combining Training and Help			
Training Only	0.21***	0.11*	-0.01
	(0.06)	(0.06)	(0.06)
BH-Corrected P-Values	[0.00]	[0.13]	[0.86]
Training + Any Help	0.11**	0.13**	0.05
	(0.06)	(0.05)	(0.05)
BH-Corrected P-Values	[0.11]	[0.04]	[0.41]
Observations	3,004	3,004	3,004
R^2	0.12	0.08	0.36
Panel C: Individual Treatment Arms			
Training Only	0.21***	0.11*	-0.01
	(0.06)	(0.06)	(0.06)
BH-Corrected P-Values	[0.00]	[0.13]	[0.86]
Training + Helper	0.16**	0.16***	-0.01
	(0.07)	(0.06)	(0.06)
BH-Corrected P-Values	[0.04]	[0.04]	[0.86]
Training + Facilitator	0.08	0.11*	0.11*
	(0.06)	(0.06)	(0.06)
BH-Corrected P-Values	[0.27]	[0.13]	[0.12]
Observations	3,004	3,004	3,004
R^2	0.12	0.08	0.36
Helper = Facilitator (p-value)	0.18	0.44	0.05

Notes: The table presents results for the main outcomes estimated using regression equation (9). Each outcome variable is a standardized score relative to the control group: the *ASQ Index* (Ages and Stages Questionnaire, a measure of child development), the *PESL Index* (Practices of Early Stimulation and Learning, capturing parental engagement), and the *Caregiver Stress Index*. All regressions include municipality fixed effects. *Any Treatment* equals one if a school received *Training Only*, *Training + Helper*, or *Training + Facilitator*, and zero otherwise. Standard errors are clustered at the school level and reported in parentheses. The data comes from the Endline Survey, which includes all 3,004 student respondents. P-values were also adjusted for multiple hypothesis testing using the False Discovery Rate (FDR) procedure Benjamini and Hochberg (1995). These adjusted p-values are reported in brackets under the standard errors. Raw p-values were first ranked in ascending order, then each was multiplied by the total number of hypotheses and divided by its rank. To ensure consistency, adjusted p-values were monotonically corrected by replacing each with the minimum of its own value and all larger-ranked values. The final adjusted p-values therefore increase weakly with rank and represent the standard Benjamini–Hochberg FDR correction. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 4: Effects on Caregiver Beliefs and Attitudes

	Attitude to Discipline	Attitude to Disability Index	Confidence Disability Index
	(1)	(2)	(3)
Panel A: Pooled Treatment Arms			
Any Treatment	0.00	0.09*	0.08
	(0.03)	(0.05)	(0.06)
Observations	3,004	3,004	3,003
R^2	0.06	0.10	0.09
Panel B: Combining Training and Help			
Training Only	-0.00	0.07	0.01
	(0.03)	(0.06)	(0.07)
Training + Any Help	0.00	0.10*	0.11*
	(0.03)	(0.05)	(0.06)
Observations	3,004	3,004	3,003
R^2	0.06	0.10	0.09
Panel C: Individual Treatment Arms			
Training Only	-0.00	0.07	0.01
	(0.03)	(0.06)	(0.07)
Training + Helper	-0.03	0.14**	0.13*
	(0.03)	(0.06)	(0.06)
Training + Facilitator	0.03	0.05	0.10
	(0.03)	(0.06)	(0.07)
Observations	3,004	3,004	3,003
R^2	0.06	0.11	0.09
Helper = Facilitator (p-value)	0.08	0.14	0.62

Notes: The table presents results on caregiver beliefs estimated using regression equation (9). The outcome variable *Attitude to Discipline* is a dummy equal to 1 if the parent considers it necessary or normal to spank a child for discipline. The outcome variable *Attitude to Disability Index* is a standardized index measuring caregiver attitudes toward disability, constructed as the additive sum of responses to whether the caregiver believes that children with disabilities should be separated from their parents, whether schools should accommodate children with disabilities, whether disabilities can be helped, and whether children with disabilities deserve access to educational opportunities. The outcome variable *Confidence Disability Index* measures the caregiver's confidence in supporting children with disabilities, based on whether they believe they have the necessary knowledge and skills and whether they feel able to provide such support. All regressions include municipality fixed effects. *Any Treatment* equals one if a school received *Training Only*, *Training + Helper*, or *Training + Facilitator*, and zero otherwise. Standard errors are clustered at the school level and reported in parentheses. The data comes from the Endline Survey, which includes all 3,004 student respondents. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 5: Effect on Main Outcomes by Baseline PESL Scores

	Above Median PESL			Below Median PESL		
	ASQ Index	PESL Index	Caregiver Stress Index	ASQ Index	PESL Index	Caregiver Stress Index
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Pooled Treatment Arms						
Any Treatment	0.05	0.14**	-0.06	0.22***	0.09	0.12**
	(0.08)	(0.06)	(0.06)	(0.08)	(0.06)	(0.05)
Observations	1,506	1,506	1,506	1,497	1,497	1,497
R^2	0.10	0.08	0.37	0.16	0.11	0.37
Panel B: Combining Training and Help						
Training Only	0.12	0.15**	-0.14**	0.28***	0.10	0.11
	(0.09)	(0.07)	(0.06)	(0.09)	(0.07)	(0.06)
Training + Any Help	0.01	0.14**	-0.02	0.19**	0.08	0.13**
	(0.08)	(0.06)	(0.06)	(0.08)	(0.07)	(0.06)
Observations	1,506	1,506	1,506	1,497	1,497	1,497
R^2	0.10	0.08	0.37	0.16	0.11	0.37
Panel C: Individual Treatment Arms						
Training Only	0.12	0.15**	-0.14**	0.28***	0.10	0.11
	(0.09)	(0.07)	(0.06)	(0.09)	(0.07)	(0.06)
Training + Helper	0.06	0.20***	-0.09	0.24***	0.10	0.07
	(0.10)	(0.07)	(0.07)	(0.09)	(0.08)	(0.07)
Training + Facilitator	-0.03	0.08	0.03	0.13	0.05	0.20***
	(0.09)	(0.07)	(0.07)	(0.09)	(0.07)	(0.07)
Observations	1,506	1,506	1,506	1,497	1,497	1,497
R^2	0.10	0.09	0.37	0.16	0.11	0.37
Helper = Facilitator (p-value)	0.37	0.13	0.08	0.17	0.54	0.13

Notes: The table presents results for the main outcomes by subgroups defined according to baseline parental engagement levels. Each outcome variable is a standardized score relative to the control group: the *ASQ Index* (Ages and Stages Questionnaire, a measure of child development), the *PESL Index* (Practices of Early Stimulation and Learning, capturing parental engagement), and the *Caregiver Stress Index*. All regressions include municipality fixed effects. *Any Treatment* equals one if a school received *Training Only*, *Training + Helper*, or *Training + Facilitator*, and zero otherwise. Panel A reports results for those with above-median baseline PESL scores, while Panel B reports results for those with below-median scores. Standard errors are clustered at the school level and reported in parentheses. The data comes from the Baseline and Endline Surveys, covering all students who participated in the Endline Survey except one child with missing baseline PESL, for a final sample of 3,003 respondents. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 6: Heterogeneous Impacts by Baseline Outcome Levels

	Child Development and Caregiver Outcomes		
	ASQ Index	PESL Index	Stress Index
	(1)	(2)	(3)
Training Only	0.31***	0.03	-0.11
	(0.12)	(0.09)	(0.08)
Training + Helper	0.22*	0.02	-0.06
	(0.12)	(0.09)	(0.09)
Training + Facilitator	0.17	-0.03	0.02
	(0.12)	(0.09)	(0.08)
Training Only × Quartile 2	-0.15	0.11	0.14
	(0.12)	(0.10)	(0.10)
Training Only × Quartile 3	-0.23*	0.08	0.18*
	(0.13)	(0.14)	(0.11)
Training Only × Quartile 4	-0.11	0.39***	0.15
	(0.13)	(0.14)	(0.13)
Training + Helper × Quartile 2	-0.04	0.16	0.08
	(0.13)	(0.11)	(0.11)
Training + Helper × Quartile 3	-0.17	0.07	0.03
	(0.14)	(0.13)	(0.11)
Training + Helper × Quartile 4	-0.13	0.30**	0.11
	(0.14)	(0.15)	(0.12)
Training + Facilitator × Quartile 2	-0.05	0.17*	0.04
	(0.13)	(0.10)	(0.11)
Training + Facilitator × Quartile 3	-0.15	-0.05	0.14
	(0.14)	(0.14)	(0.11)
Training + Facilitator × Quartile 4	-0.05	0.24*	0.19
	(0.14)	(0.13)	(0.13)
Quartile 2	0.40***	0.35***	0.09
	(0.10)	(0.08)	(0.06)
Quartile 3	0.63***	0.82***	0.05
	(0.10)	(0.10)	(0.07)
Quartile 4	0.72***	0.75***	0.08
	(0.10)	(0.09)	(0.09)
F-test of interaction terms (p-value)	0.73	0.15	0.68
Observations	3,004	3,003	2,987
R^2	0.30	0.21	0.37

Notes: The table reports heterogeneous treatment effects of the intervention across quartiles of the baseline distribution of each outcome. Each column corresponds to an OLS regression of the endline outcome—the *ASQ Index* (Ages and Stages Questionnaire, a measure of child development), the *PESL Index* (Practices of Early Stimulation and Learning, capturing parental engagement), or the *Caregiver Stress Index*—on treatment assignment, quartile indicators, and their interactions. All regressions include municipality fixed effects. The coefficients on the treatment indicators represent impacts for children in the first quartile of the baseline distribution, while the interaction terms capture differential impacts for children in higher quartiles. The F-test at the bottom reports whether the set of interaction terms jointly equals zero. *Any Treatment* equals one if a school received *Training Only*, *Training + Helper*, or *Training + Facilitator*, and zero otherwise. Standard errors are clustered at the school level and reported in parentheses. The data comes from the Endline Survey, which includes all 3,004 student respondents. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 7: Effects on Parent Education Session Attendance

	Attendance	Attendance Consistency
	(1)	(2)
Training + Helper	1.24*	-0.02
	(0.70)	(0.03)
Training + Facilitator	0.07	-0.02
	(0.72)	(0.02)
Session Number	-0.03*	0.00
	(0.01)	(0.00)
Observations	2,600	2,600
R^2	0.37	0.27

Notes: The table presents results on Parent Education session attendance. The outcome variable is the number of parents attending each session in column (1), and the within-school coefficient of variation in session attendance in column (2). All regressions include municipality fixed effects. *Training Only* is the reference category. *Training + Helper* equals one if a school was assigned to the training plus classroom helper arm, and zero otherwise. *Training + Facilitator* equals one if a school was assigned to the training plus outside classroom helper (facilitator) arm, and zero otherwise. The data come from Parent Education Session Attendance Records, covering 130 treatment schools that implemented parent education sessions and for which attendance data were collected, totaling 2,600 session-level observations. Of the 150 treatment schools, 12 did not implement Parent Education sessions (9 had not sent teachers for Parent Education training, and 3 had facilitators who opted out), and attendance data were not collected for 8 additional schools. Standard errors are clustered at the school level and reported in parentheses. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 8: Complementarity Effects Across Child and Caregiver Outcomes

	ASQ Index	PESL Index	Caregiver Stress Index
	(1)	(2)	(3)
Any Treatment	0.09	0.14*	0.14*
	(0.12)	(0.08)	(0.08)
High PESL	0.07	0.64***	0.07
	(0.06)	(0.07)	(0.06)
High LEP	0.12	0.19	0.21
	(0.33)	(0.16)	(0.18)
Any Treatment × High PESL	-0.09	-0.04	-0.15*
	(0.09)	(0.10)	(0.08)
Any Treatment × High LEP	0.08	-0.24	-0.23
	(0.33)	(0.16)	(0.18)
High PESL × High LEP	-0.69***	-0.27	-0.13
	(0.22)	(0.25)	(0.17)
Any Treatment × High PESL × High LEP	0.53**	0.36	0.09
	(0.23)	(0.27)	(0.18)
Observations	3,003	3,003	3,003
R^2	0.13	0.19	0.36

Notes: Each outcome variable is a standardized score relative to the control group for the aggregate Ages and Stages Questionnaire (ASQ), the Practice of Early Stimulation and Learning (PESL), and the Caregiver Stress Indexes. All regressions include municipality fixed effects. *Any Treatment* equals one if a school received *Training Only*, *Training + Helper*, or *Training + Facilitator*, and zero otherwise. *High PESL* equals one if baseline PESL is above the median. *High LEP* (Learning Environment and Practices Index) equals one if predicted LEP improvement is above the median. The data comes from the Baseline Survey, Learning Environment and Practices Checklist, and the Endline Survey covering all students who participated in the Endline Survey except one child with missing baseline PESL, for a final sample of 3,003 respondents. Standard errors are clustered at the school level and reported in parentheses. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 9: Benchmarking Program Impacts and Costs

Study	Country	Main Outcomes	Effect Size (SD)	Program Type/Cost-Effectiveness
A. Teacher Professional Development Center-Based Quality Reforms				
Nepal ECD (this study)	Nepal	Ages and Stages Questionnaire (ASQ); Practices of Early Stimulation and Learning (PESL)	+0.13–0.20 (ASQ); +0.11–0.20 (PESL)	Teacher training + parent engagement; modest cost; no formal benefit–cost ratio (BCR) yet
Andrew et al. (2024)	Colombia	Cognitive latent factor	+0.16 overall; +0.30 poorest	Teacher professional development (Hogares Infantiles); $35 upfront + $13 annually per child
Wolf et al. (2019)	Ghana	School readiness; teacher well-being	+0.11 (cognitive); +0.18 (socio-emotional)	Quality preschool reform; BCR 11–103:1
Gallego et al. (2021)	Peru	Cognitive test scores	+0.19	Pedagogical reform in preschools; BCR 29–88
Brinkman et al. (2017)	Indonesia	Socio-emotional skills	+0.16	ECE expansion in rural areas; cost-effective at scale
Spier et al. (2020)	Bangladesh	Cognitive and socio-emotional	+0.25 (cognitive); +0.37 (socio-emotional)	Preschool initiative; BCR 3–29
Coffey et al. (2017)	India	Field-based math tasks; socio-emotional	+0.09 (cognitive); +0.17 (socio-emotional)	Preschool/classroom enrichment; BCR 6–54
B. Parenting Groups/Home-Visiting				
Justino et al. (2023)	Rwanda	ASQ indices; parenting practices	+0.30–0.38 (ASQ); +0.47–0.62 (parent–child activities)	Parenting group model (First Steps); $2 per caregiver per session
Andrew et al. (2020)	India (urban)	Bayley-III developmental indices	+0.30 overall; +0.35 cognitive	Home-visiting program; $148 per child per year
C. Financial Access/Scholarship Subsidies				
Dean and Jayachandran (2020)	India	Cognitive composite; school attendance	+0.80 at end of kindergarten	Kindergarten scholarship subsidy; cost not reported
D. Cash Transfers				
López Bóo and Creamer (2019)	Honduras	ASQ (cognitive)	+0.13	Conditional cash transfer; $500 per household per year
E. State Capacity/System Reform				
Ganimian et al. (2024)	India (Tamil Nadu)	Math, language, executive function, nutrition	+0.29–0.46 (tests); +0.10–0.11 (nutrition, HH tests)	ICDS facilitator model; BCR 13:1–22:1
F. Access/Quality Expansion with Null Impacts				
Blimpo et al. (2022)	Gambia	Motor and language/hearing (MDAT)	Null	Expansion/quality trial; low cost, no impacts

Notes: ASQ = Ages and Stages Questionnaire. PESL = Practices of Early Stimulation and Learning. MDAT = Malawi Developmental Assessment Tool. ICDS = Integrated Child Development Services. BCR = Benefit–Cost Ratio. Effect sizes reported in standard deviations (SD). Costs expressed in 2020 US dollars where available.

Table 10: Program Costs and Unit Costs (Aug 2022–Feb 2023)

A. Cost Summary	
Remuneration	USD 2,508.79 [NPR 316,108]
Program costs	USD 41,539.68 [NPR 5,234,998]
Indirect costs (8%)	USD 3,523.88 [NPR 444,009]
Grand total	**USD 47,572.35** [**NPR 5,994,116**]
B. Unit costs	
Centers covered (number)	150
Cost per center	USD 317.15 [NPR 39,961]
Per-child cost[b] (12/15/20)	26.43/21.14/15.86 [NPR 3,330/2,664/1,998]

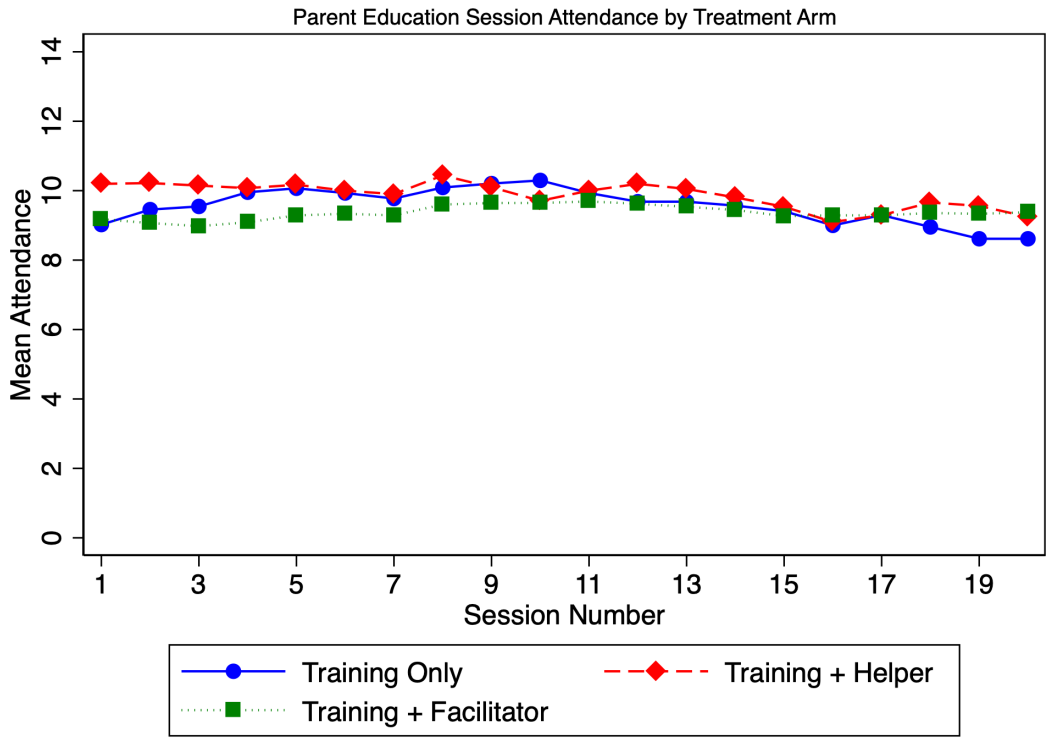
[a] Exchange rate: 1 USD = 126 NPR (financial proposal). NPR shown in brackets.

[b] Terms of Reference caps Parenting Education group size at 20 per center. Per-child cost = (total cost ÷ 150 centers) ÷ assumed enrollment.

Notes: Costs cover twenty-session Parenting Education, remuneration for 50 facilitators and 50 caregivers, coaching/mentoring, and monitoring/supervision.

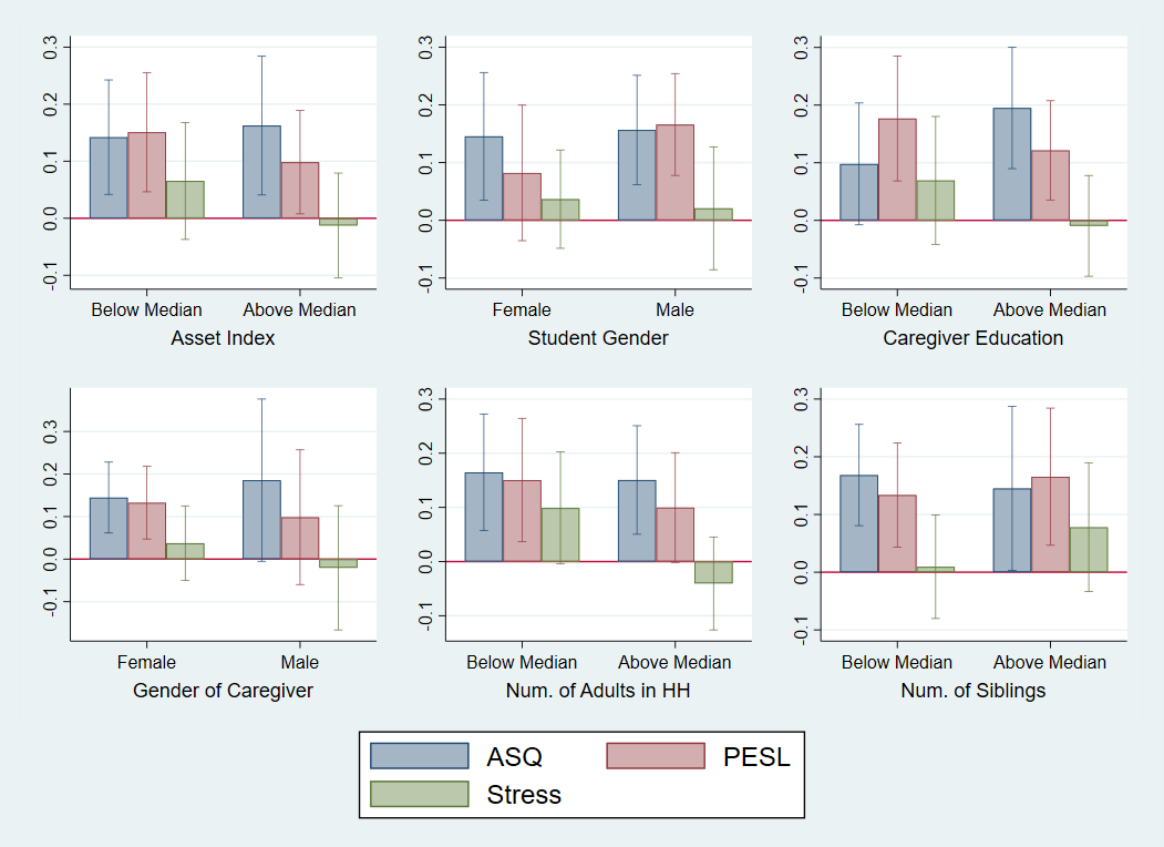
Figures

Figure 2: Parent Education Session Attendance by Treatment Arm



Notes: The figure shows the mean number of parents attending each parent education session by treatment arm. The sample includes 130 treatment schools that implemented parent education sessions and for which attendance data were collected. The horizontal axis represents the session number (1–20), and the vertical axis shows mean attendance. The blue line represents schools in the *Training Only* arm, the red line represents schools in the *Training + Helper* arm, and the green line represents schools in the *Training + Facilitator* arm. Out of 150 treatment schools, 12 schools did not implement parent education sessions (9 had not sent teachers for PE training, and 3 had facilitators who opted out), and attendance data were not collected for 8 additional schools. *Source:* Parent education session data.

Figure 3: Heterogeneity by Child and Household Characteristics



Notes: This figure shows heterogeneity in the impacts of the pooled treatments on the three main outcomes, disaggregated by key baseline child and household characteristics. Each panel contrasts estimated treatment effects for subgroups below and above the median, with bars denoting point estimates and error bars indicating 90% confidence intervals.

Appendices

Table of Contents

A Appendix Tables

Table A1: Balance Table

	Overall	Control	Training Only	Training + Helper	Training + Facilitator	Training Only vs. Control	Training + Helper vs. Control	Training + Facilitator vs. Control
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Demographics								
Child's Age (months)	47.644	48.131	48.095	47.781	46.570	-0.035	-0.350	-1.560
	(12.000)	(12.550)	(11.872)	(12.220)	(11.279)	(1.100)	(1.178)	(1.073)
Child's Gender (female)	0.503	0.495	0.499	0.519	0.500	0.005	0.024	0.005
	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.024)	(0.024)	(0.022)
Number of Adults in HH	3.135	3.115	3.007	3.233	3.189	-0.108	0.118	0.074
	(1.533)	(1.433)	(1.442)	(1.581)	(1.661)	(0.093)	(0.095)	(0.130)
Number of Siblings	1.253	1.185	1.338	1.225	1.265	0.153	0.039	0.079
	(1.130)	(1.039)	(1.189)	(1.154)	(1.129)	(0.099)	(0.090)	(0.104)
Asset Index	7.800	7.803	7.721	7.836	7.841	-0.082	0.033	0.038
	(3.315)	(3.319)	(3.241)	(3.501)	(3.203)	(0.306)	(0.313)	(0.286)
Caregiver's Gender (female)	0.835	0.816	0.847	0.833	0.845	0.031	0.017	0.029
	(0.371)	(0.388)	(0.360)	(0.373)	(0.362)	(0.025)	(0.027)	(0.025)
Caregiver's Education	2.834	2.898	2.777	2.810	2.849	-0.121	-0.089	-0.049
	(1.374)	(1.330)	(1.389)	(1.399)	(1.379)	(0.118)	(0.129)	(0.120)
Panel B: Developmental Outcomes and Caregiver Attitude								
Baseline ASQ Score	0.036	0.000	0.108	0.099	-0.061	0.108	0.099	-0.061
	(0.950)	(0.999)	(0.895)	(0.919)	(0.983)	(0.078)	(0.082)	(0.078)
Confidence towards Disability Index	-0.005	0.000	0.071	-0.052	-0.041	0.071	-0.052	-0.041
	(1.033)	(1.000)	(1.101)	(0.980)	(1.042)	(0.078)	(0.073)	(0.081)
Attitude towards Disability Index	-0.096	0.000	-0.076	-0.134	-0.175	-0.076	-0.134*	-0.175**
	(0.929)	(1.000)	(0.917)	(0.903)	(0.887)	(0.069)	(0.075)	(0.068)
Baseline Communication Index	0.044	0.000	0.106	0.116	-0.043	0.106	0.116	-0.043
	(0.950)	(0.999)	(0.878)	(0.930)	(0.977)	(0.078)	(0.084)	(0.085)
Problem Solving Index	-0.018	0.000	0.018	-0.015	-0.076	0.018	-0.015	-0.076
	(0.997)	(0.999)	(0.990)	(0.972)	(1.024)	(0.077)	(0.077)	(0.072)
Baseline PESL Index	0.037	0.000	-0.040	0.043	0.146	-0.040	0.043	0.146**
	(0.989)	(1.000)	(0.985)	(0.972)	(0.990)	(0.061)	(0.072)	(0.070)
Baseline Stress Index	0.066	0.000	-0.009	0.116	0.160	-0.009	0.116	0.160
	(1.073)	(1.000)	(1.059)	(1.077)	(1.145)	(0.091)	(0.098)	(0.099)
Respondent thinks child has Disability	0.096	0.164	0.024	0.164	0.034	-0.141	0.000	-0.130
	(2.559)	(3.580)	(0.152)	(3.679)	(0.182)	(0.127)	(0.186)	(0.127)
Observations	3,004	766	757	725	756			

Notes: The table reports descriptive statistics and balance tests, showing group means and mean differences between treatment and control schools. Columns (1)–(5) present means for the *Overall* sample, the *Control* group, the *Training Only* group, the *Training + Helper* group, and the *Training + Facilitator* group, respectively. Columns (6)–(8) report differences in means between each treatment group and the control group. Standard errors clustered at the school level are reported in parentheses. The analysis sample comprises 3,004 student respondents from the Endline Survey. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table A2: Treatment Effects on Child Development and Caregiver Outcomes with Baseline Controls

	ASQ Index	PESL Index	Caregiver Stress Index
	(1)	(2)	(3)
Panel A: Pooled Treatment Arms			
Any Treatment	0.13***	0.11**	0.03
	(0.05)	(0.05)	(0.05)
Observations	3,004	3,003	2,987
Panel B: Combining Training and Help			
Training Only	0.16***	0.14**	-0.01
	(0.05)	(0.06)	(0.06)
Training plus any help	0.10*	0.09**	0.05
	(0.05)	(0.05)	(0.05)
Observations	3004	3003	2987
Panel C: Individual Treatment Arms			
Training Only	0.17***	0.13**	-0.01
	(0.05)	(0.06)	(0.06)
Training + Helper	0.12*	0.14**	-0.01
	(0.06)	(0.06)	(0.06)
Training + Facilitator	0.10	0.06	0.10*
	(0.06)	(0.05)	(0.06)
Observations	3004	3003	2987
Baseline Control	Yes	Yes	Yes

Notes: The table reports results for the main outcomes estimated using regression equation (9), with baseline controls included for each outcome. Outcomes are standardized relative to the control group: the *ASQ Index* (Ages and Stages Questionnaire, a measure of child development), the *PESL Index* (Practices of Early Stimulation and Learning, capturing parental engagement), and the *Caregiver Stress Index*. Baseline controls correspond to the same measure at baseline (e.g., baseline ASQ score for ASQ outcomes, baseline PESL score for PESL outcomes, and baseline Stress score for Stress outcomes). All regressions include municipality fixed effects. *Any Treatment* equals one if a school received *Training Only*, *Training + Helper*, or *Training + Facilitator*, and zero otherwise. The analysis sample comprises 3,004 student respondents from the Endline Survey. Standard errors are clustered at the school level and reported in parentheses. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table A3: Treatment Effects on Child Development across ASQ Index Sub-Components

	ASQ	Communication	Gross Motor	Fine Motor	Problem-solving	Personal-social
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Pooled Treatment Arms						
Any Treatment	0.13*	0.15**	0.09*	0.12**	0.10	0.08
	(0.07)	(0.06)	(0.04)	(0.06)	(0.07)	(0.06)
Observations	3,004	1,455	3,004	3,004	3,004	3,004
Panel B: Combining Training and Help						
Training Only	0.20**	0.22***	0.13***	0.16**	0.17**	0.14*
	(0.08)	(0.07)	(0.05)	(0.07)	(0.08)	(0.07)
Training plus any help	0.09	0.12*	0.06	0.10	0.06	0.05
	(0.07)	(0.07)	(0.05)	(0.06)	(0.08)	(0.07)
Observations	3004	1455	3004	3004	3004	3004
Panel C: Individual Treatment Arms						
Training Only	0.20**	0.22***	0.13***	0.16**	0.17**	0.14*
	(0.08)	(0.07)	(0.05)	(0.07)	(0.08)	(0.07)
Training plus Helper	0.15*	0.17**	0.02	0.13	0.13	0.12
	(0.08)	(0.07)	(0.06)	(0.08)	(0.09)	(0.08)
Training plus Facilitator	0.04	0.07	0.10**	0.08	-0.01	-0.01
	(0.08)	(0.08)	(0.05)	(0.07)	(0.08)	(0.08)
Observations	3004	1455	3004	3004	3004	3004

Notes: The table reports results for the main outcomes estimated using regression equation (9), with ASQ sub-components. Outcomes are standardized relative to the control group: the *ASQ* (Ages and Stages Questionnaire, a measure of child development) is an aggregated score across its sub-components presented in columns (2) through (6). All regressions include municipality fixed effects. *Any Treatment* equals one if a school received *Training Only*, *Training + Helper*, or *Training + Facilitator*, and zero otherwise. The analysis sample comprises 3,004 student respondents from the Endline Survey. Standard errors are clustered at the school level and reported in parentheses. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

Table A4: Attrition

Predictor Variable	Coefficient	P-value
	(1)	(2)
Panel A: Differential Attrition (Treatment vs. Control)		
Training Only	0.029	0.20
	(0.023)	
Training + Helper	0.022	0.25
	(0.019)	
Training + Facilitator	0.006	0.52
	(0.010)	
Panel B: Baseline Predictors of Attrition		
Age of the Child	-0.000	0.42
	(0.000)	
Gender of Child (Female = 1)	-0.010*	0.09
	(0.006)	
Number of Adults in HH	-0.000	0.99
	(0.002)	
Number of Children under 18	0.003	0.42
	(0.004)	
Asset Index	-0.003**	0.01
	(0.001)	
Gender of Caregiver (Female = 1)	-0.015	0.12
	(0.010)	
Education of Caregiver	-0.004*	0.07
	(0.002)	
Baseline PESL Index	-0.003	0.40
	(0.003)	
Baseline Stress Index	-0.003	0.44
	(0.003)	
Baseline ASQ Index	-0.003	0.54
	(0.004)	

Notes: The table reports attrition results from the Endline Survey. Panel A presents coefficients from OLS regressions of an attrition indicator, equal to one if a student was not observed at endline and zero otherwise, on treatment group indicators. Panel B reports coefficients from OLS regressions of the same attrition indicator on baseline child and household characteristics. All regressions include municipality fixed effects. The *ASQ Index* refers to the Ages and Stages Questionnaire (a measure of child development), the *PESL Index* refers to the Practices of Early Stimulation and Learning (capturing parental engagement), and the *Caregiver Stress Index* measures self-reported stress among primary caregivers. Standard errors are clustered at the school level and reported in parentheses, with corresponding p-values shown in Column (2).

Table A5: Lee Bounds Estimates of Treatment Effects

Outcome	Lower Bound [95% Conf. Interval]	Upper Bound [95% Conf. Interval]
	(1)	(2)
ASQ Index	0.066	0.229
	[-0.022, 0.153]	[0.136, 0.322]
PESL Index	-0.116	0.121
	[-0.189, -0.043]	[0.042, 0.199]
Stress Index	-0.119	0.051
	[-0.220, -0.019]	[-0.049, 0.151]

Notes: The table reports upper and lower bounds on treatment effects for the *ASQ Index* (Ages and Stages Questionnaire, a measure of child development), the *PESL Index* (Practices of Early Stimulation and Learning, capturing parental engagement), and the *Caregiver Stress Index*, accounting for attrition from the Endline Survey. Bounds are estimated following Lee (2009). Ninety-five percent confidence intervals are reported in brackets beneath each estimate.

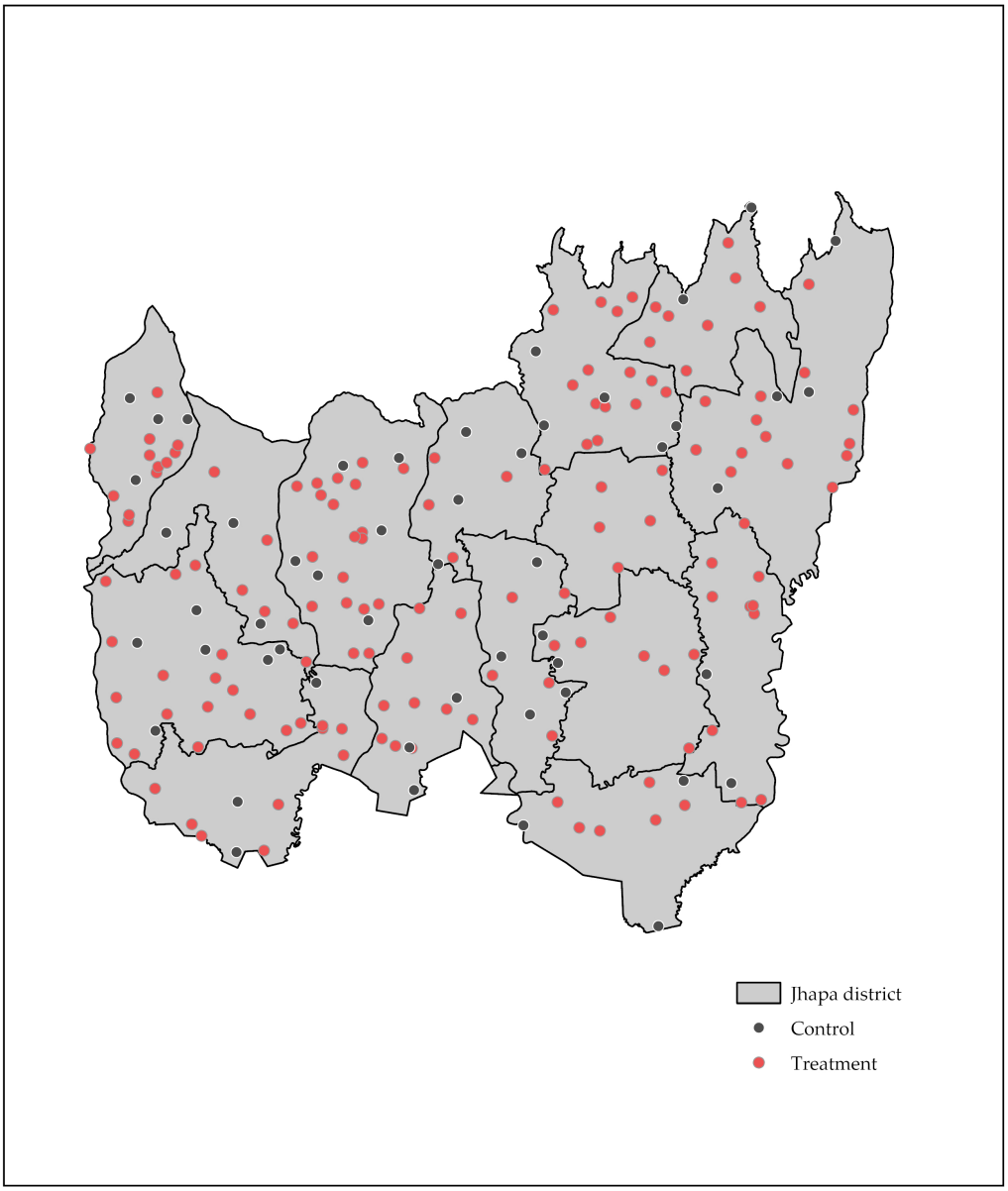
Table A6: Non-Compliance in Parental Education Sessions

Baseline Characteristic	Coefficient	P-value
	(1)	(2)
Age of the Child	0.894	0.42
	(1.105)	
Gender of Child (Female = 1)	0.031	0.30
	(0.030)	
Number of Adults in HH	0.313**	0.03
	(0.143)	
Number of Siblings	0.102	0.28
	(0.094)	
Asset Index	0.023	0.94
	(0.297)	
Gender of Caregiver (Female = 1)	-0.067*	0.05
	(0.035)	
Education of Caregiver	-0.356**	0.01
	(0.140)	
Baseline ASQ Index	0.031	0.67
	(0.073)	
Baseline PESL Index	0.019	0.79
	(0.072)	
Baseline Stress Index	-0.012	0.92
	(0.125)	

Notes: The table reports associations between non-compliance in parental education (PE) sessions and baseline child and household characteristics. Each row reports coefficients from regressing the corresponding baseline characteristic on an indicator variable equal to one if a student's school did not conduct PE sessions and zero otherwise. Coefficients are estimated using OLS with municipality fixed effects. The *ASQ Index* refers to the Ages and Stages Questionnaire (a measure of child development), the *PESL Index* refers to the Practices of Early Stimulation and Learning (capturing parental engagement), and the *Caregiver Stress Index* measures self-reported stress among primary caregivers. Standard errors are clustered at the school level and reported in parentheses, with corresponding p-values shown in Column (2).

B Appendix Figures

Figure A1: Map of Distribution of Treatment and Control Schools across Jhapa District



Notes: This figure plots the location of the 51 *Control* and 150 *Treatment* schools in this study. *Treatment* schools include 50 schools from each of the three treatment arms (*Training Only*, *Training + Helper*, and *Training + Facilitator*).

C Teacher Education Training Sessions

C.1 Overview

This appendix provides a summary of the 10 day education training sessions for teachers and facilitators of the parental education sessions.

C.2 Session Details

No.	Topic and Session Details
1	**Conceptual Understanding of Early Childhood Development** On Day I of the Early Childhood Development (ECD) Training of Trainers, teachers engaged in a comprehensive learning journey exploring the foundations of child development. They began with introductions and a pre-test, then delved into understanding ECD concepts, exploring its importance from multiple perspectives including brain development and early learning. Throughout the day, the teachers focused on the major domains of child development (physical, social, emotional, cognitive, and language development). The teachers participated in interactive activities such as group presentations, brainstorming sessions, and a debate on biological versus environmental factors influencing child growth. They examined child development domains (physical, social, emotional, cognitive, and language), identified ECD guiding principles, and concluded the day by reflecting on the benefits of quality ECD experiences for children, families, and society through group work and a collaborative feedback circle.
2	**Child Growth and Development** On Day II of the Early Childhood Development training, teachers explored the comprehensive landscape of child growth and development. They began by discussing the benefits of ECD, then examined developmental milestones from prenatal stages through early childhood, covering key developmental periods from birth to 96 months. The session progressed to an in-depth exploration of childhood disabilities, focusing on understanding different types of disabilities, their causes, and preventive strategies. Teachers engaged in discussions about inclusive education, learning how to create supportive environments for children with diverse needs. They also analyzed child behavior and disorders, developing skills to recognize, understand, and address behavioral challenges. The day concluded with a reflective session where participants wrote diary notes, encouraging personal growth and deeper understanding of the complexities of early childhood development.

No.	Topic and Session Details
3	**ECD Policies, Child Health, Nutrition, Care** On Day III, teachers explored ECD policies, child rights, health, and nutrition through an interactive and analytical approach. The day began with participants sharing personal stories and reviewing developmental milestones, then progressed to analyzing the situation of children in local contexts, examining their basic needs and rights. They thoroughly discussed constitutional provisions, child rights based on UNCRC, and the holistic development of children, with a particular focus on the Nepalese context. The session covered critical aspects such as health, nutritional, and protection needs, and participants reviewed existing acts and provisions. A significant portion of the day was dedicated to understanding the ECD National Strategy 2077, the Nurturing Care Framework, and exploring the roles of local authorities, parents, and stakeholders in ECD implementation. Teachers also examined development plans, including the School Sector Development Plan and Sustainable Development Goals, concluding with a reflection on the federal government's role in early childhood development across different administrative levels.
4	**ECD Programs** On Day IV, teachers explored comprehensive ECD program management and implementation strategies. The day began with participants reflecting on their previous assignments and discussing the role of government in ECD, then transitioned into an in-depth examination of various ECD program models including parenting programs, home-based, community-based, and school-based initiatives. The session covered critical aspects of ECD center management, including daily operations, physical and learning environments, and the roles of teachers, parents, and community members. Teachers also participated in interactive activities such as role-plays focusing on WASH (Water, Sanitation, and Hygiene) management, resource mobilization techniques, and exploring coordination and communication strategies. The day concluded with participants identifying opportunities for quality ECD management and evaluating the day's learnings.
5	**ECD Curriculum** On Day V of the Early Childhood Development training, teachers engaged in an exploration of the ECD national curriculum, focusing on its fundamental components, implementation strategies, and holistic child development approach. The day began with reflections on previous management learnings and then delved deeply into the 2061 National Curriculum, examining its goals, objectives, learning environments, instructional strategies, and approaches to inclusive education. Participants analyzed core competencies, developmental skills, and early learning frameworks, with particular emphasis on understanding holistic child development. The session included interactive activities such as paired discussions on integrated thematic learning, debates on school readiness, and collaborative creation of theme-based web charts linking developmental domains and learning concepts. Teachers explored practical aspects of ECD curriculum implementation, including daily routine activities, early literacy and math concepts, and strategies for creating inclusive and developmentally appropriate learning environments. The day concluded with participants preparing weekly and daily theme-based plans, reflecting a hands-on approach to understanding and applying ECD curriculum principles.

No.	Topic and Session Details
6	**Early Learning and Development** On Day VI, teachers focused on in-depth learning assessment, play-based pedagogical approaches, and practical implementation strategies for early childhood education. The day began with participants reflecting on previous curriculum learnings and then transitioned into exploring various assessment techniques, observation methods, and documentation strategies for tracking children's developmental progress. Teachers engaged in interactive sessions discussing developmentally appropriate assessment practices, learning how to create meaningful documentation that captures children's holistic growth across different developmental domains. The session included hands-on activities such as designing observation tools, practicing documentation techniques, and understanding how to use assessment data to inform teaching strategies. Participants examined play-based learning approaches, exploring how different types of play contribute to children's cognitive, social, emotional, and physical development. The day concluded with collaborative work on creating assessment frameworks and play-based learning strategies.
7	**Instruction and Learning** On Day VII, teachers dove into child-centered instruction and learning approaches, focusing on creating inclusive, engaging, and developmentally appropriate educational experiences. Teachers extensively examined child-centered instruction strategies, discussing early learning principles, sensitive learning periods, and effective approaches to educational engagement. A significant portion of the day was dedicated to hands-on group activities exploring various child-centered learning techniques, including play, songs, rhymes, drama, role-play, art, craft, and storytelling. Participants also focused on critical aspects of inclusive education, particularly strategies for engaging children with disabilities and understanding different types of learning disabilities. The session emphasized practical skill development, with teachers working collaboratively to create theme-based learning materials, songs, stories, and activities that promote holistic child development.
7	**Child Learning and Development Assessment** Day VII of the Early Childhood Development training was broken into two parts. In the second half of the day, teachers explored comprehensive child assessment strategies, focusing on continuous and holistic methods of evaluating children's development and learning progress. The day began with participants displaying their theme-based learning materials and then transitioned into an in-depth examination of continuous assessment techniques, emphasizing the importance of ongoing, contextual evaluation methods. Teachers learned about various assessment approaches, including direct and indirect methods such as observation, interaction, questioning, and performance analysis, with a particular focus on assessing progress across physical, cognitive, emotional, social, and language domains. Participants engaged in practical sessions on developing assessment tools, learning how to record and interpret children's developmental status through methods like child intake forms, monitoring and evaluation documents, and comprehensive child portfolios. The day concluded with participants preparing individual child case reports, reflecting a thorough and empathetic approach to understanding and documenting children's learning journeys.

No.	Topic and Session Details
8	**Teacher's Performance Development (TPD) for ECD Facilitators** On Day VIII, teachers focused on professional development for ECD facilitators, exploring comprehensive strategies for enhancing teaching performance and competence. The day began with participants sharing case reports from previous assessments and then delved into the critical aspects of Teachers Professional Development (TPD), examining its conceptual frameworks, importance, and practical implementation. Participants engaged with detailed TPD curriculum models specifically designed for ECD facilitators, reviewing resource materials and understanding the nuanced approaches to professional growth. The session included presentations from HCRD (Human Capital and Resource Development) authorities, who provided insights into performance evaluation methods, test types, and assessment tools for educators. Teachers learned about various evaluation techniques, report preparation strategies, and documentation practices essential for continuous professional improvement.
9	**Model Training Curriculum** On Day IX of the Early Childhood Development training, teachers engaged in a comprehensive exploration of the Teachers Professional Development (TPD) Model Training Curriculum, focusing on creating and implementing robust professional development frameworks for ECD facilitators. The day was structured around developing a detailed 10-day facilitator training program, with participants meticulously preparing session contents and contextual facilitation plans. Teachers worked collaboratively to develop five-day project plans, reviewing and presenting ECD TPD documents to ensure comprehensive and adaptable professional development strategies. The session included practical simulation activities, where participants practiced demonstration teaching and facilitation techniques, gaining hands-on experience in implementing training contents. Participants also focused on developing training implementation guidelines, exploring coordination methods, reporting techniques, and various assessment approaches. The day concluded with a comprehensive evaluation of the entire Training of Trainers (TOT) program, reflecting on the learning journey and synthesizing the key insights gained throughout the multi-day professional development experience.
10	**Project work, field based learning and Evaluation and Assessment** On Day X of the Early Childhood Development Training of Trainers, teachers engaged in a comprehensive culmination of their learning journey, focusing on practical application, project implementation, and final assessment of their acquired skills and knowledge. The day began with a detailed exploration of facilitator evaluation criteria, including marking rules and grading procedures, emphasizing the importance of rigorous professional assessment. Participants developed practical project work plans for ECD centers, creating actionable strategies to apply their newfound knowledge and skills. Teachers collaboratively prepared comprehensive action plans covering facilitator training, self-study, coordination, communication, and report preparation, demonstrating their ability to translate theoretical learning into practical implementation. The session included a multi-faceted evaluation process examining participants' attendance, assignments, written tests, project practices, and reports. The training concluded with feedback collection, a review of the entire 10-day learning experience, and a closing ceremony that celebrated the participants' professional growth and commitment to early childhood development.

D Participatory Education Training Sessions

D.1 Overview

This appendix provides a summary of the 20 participatory education training sessions, outlining key topics and learning objectives.

D.2 Session Details

No.	Topic and Session Details
1	**Introduction and Importance of Early Child Development** Introduces the critical importance of early childhood (pregnancy to age 8) for physical, cognitive, emotional, and social development. Key points include rapid brain growth, building a foundation for lifelong learning, and parents' role in providing nurturing care, safety, and equal opportunities. Discussion topics include perceptions of early childhood and challenges, with a focus on the responsibility of parents to foster healthy development through active participation and positive role modeling.
2	**Role of Father in Early Child Development** Defines the role of the father in early childhood development and shared parental responsibilities because a nurturing father-child relationship is beneficial to the well-being of a child. Fathers should provide prenatal support, emotional bonding, be actively involved in supporting mother and child. Discussion focuses on challenges and ways in which fathers can stay engaged.
3	**Early Child Development Programs and Responsible Sections and Institutions** This sections informs parents about local government and non-government programs that support early childhood development. These include parenting education, home-based development, entry-level childcare, and early childhood programs focused on children's physical, cognitive, and emotional growth. Key takeaways for parents include identifying available programs and ensuring that essential needs such as vaccination, healthcare, nutrition, clean drinking water, security, and early education are met.
4	**Care of Mothers, Children During Pregnancy, and Newborn Infants** Provides a comprehensive guide for caring for mother and child from pregnancy to early infancy, covering maternal and child health, family support, preparation for birth, and early newborn needs. Parents learn about fetal development, nutrition, safe delivery practices, early infant care, and developmental milestones. The session also addresses postpartum recovery and emotional well-being, ensuring parents feel confident in their journey from pregnancy to parenthood.

No.	Topic and Session Details
5	**Stimulation In Early Childhood** Emphasizes the importance of sensory and kinesthetic stimulation in brain development. Parents learn how visual, auditory, tactile, gustatory, and kinesthetic stimulation promotes brain development and learning. The discussion covers the benefits of stimulation, the risks of its absence, and the caregivers' role in creating a environment that fosters cognitive, emotional, and physical development.
6	**Components of Early Child Development** Details five key aspects of child development: physical, cognitive, linguistic, emotional, and social. Parents learn all aspects are interrelated and that nutrition, stimulation, health services, protection, care, and learning opportunities encourage holistic development. Parental responsibilities include providing proper health and education, as well as fostering a supportive environment through consistent engagement and role modeling. Discussion explores challenges and how to best support development.
7	**Early Childhood Development Stages and Characteristics** Outlines developmental phases of early childhood, as well as milestones and age-appropriate activities that align with each stage. Parents are taught to identify each stage and what factors contribute to well-rounded growth. This session encourages discussion on observed child behaviors, ways to accelerate development, and importance of engagement, protection, and proper care in early childhood.
8	**Early Detection of Children with Special Needs** Parents learn how to identify disabilities in infants and children, along with the importance of early detection and appropriate support. Learning objectives include familiarity with different types of disabilities, their causes, preventative measures, and indicators for early diagnosis. Preventive measures include maternal care, vaccinations, and safe childbirth practices. Parents and guardians are encouraged to focus on their children's abilities rather than their limitations, as it is important to create an inclusive environment where children with disabilities receive equal opportunities for development. Parents of children with disabilities are encouraged seek available resources, including disability identity cards and support services.
9	**Causes and Prevention of Child Malnutrition** Malnutrition is taught as an imbalance between the essential elements required by the body that can be identified as acute malnutrition (wasting), chronic malnutrition (stunting), underweight, or micronutrient deficiencies. Parents learn its causes, symptoms, and prevention strategies. Causes include hereditary factors, lack of awareness, insufficient maternal nutrition, food insecurity, vitamin and mineral deficiencies, and frequent illness. Prevention methods involve maternal and child nutrition, exclusive breastfeeding, regular health check-ups, hygiene, and utilizing government nutrition programs. Parents play a key role in providing balanced diets, and monitoring growth.

No.	Topic and Session Details
10	**Child Health and Care** Parents learn the importance of nurture and timely treatment in protecting children from illnesses. This involves ability to identify symptoms, their causes, preventive measures, and provide appropriate care—including emotional support. Common childhood diseases are fever, pneumonia, diarrhea, and infections. Home remedies and medical interventions for various illnesses are discussed, stressing the need for attentive caregiving and prompt medical consultation when necessary.
11	**Prevention of Harmful Contagious Diseases In Children** Brings awareness to child-fatal contagious diseases and their prevention through full immunization. Diseases such as tetanus, tuberculosis, diphtheria, polio, and measles, along with their symptoms, transmission methods, and available vaccines are discussed. Parents are encouraged to follow immunization schedules, recognize early symptoms, and seek immediate medical care when necessary. The government provides free vaccines to all children, so parents must get their children vaccinated.
12	**Child Behavior and Positive Discipline;** **Protection of Children from Violence, Abuse, and Punishment** Children develop habits and behavior by observing their parents and environment. Parents learn their importance as role models and reinforcers of good habits through praise, consistency, and positive engagement. Harsh discipline and neglect negatively impact personality development, while encouragement and structured support help children manage emotions and behavior effectively. Violence, abuse, and discrimination have lasting detrimental effects and are violations of children's rights. A nurturing, non-discriminatory environment with clear expectations, encouragement, and legal protections is essential for healthy child development.
13	**Child Safety, Protection from Accidents, and Learning in Emergencies** Creating a safe environment and emergency preparedness at home and in the community protects children from accidents, ensures their safety and continued development. Children are naturally curious and may not foresee dangers so childproofing and risk through stories and demonstrations are essential. Preventative measures include securing hazardous objects and keeping them out of reach, teaching road safety, and supervising child activities. In emergencies, knowing basic first aid—such as treating burns, cuts, and fractures—can be lifesaving. Disaster preparedness involves identifying hazards, planning emergency responses, storing essential supplies, and practicing evacuation drills.
14	**Importance of Early Childhood Learning** Introduces the importance of early childhood learning through play and experience-based methods. Babies learn from birth by engaging with their environment using their senses, imitation, problem-solving, and exploration. These early experience significantly impact brain development due to rapid formation of neural connections in the first three years. Age-appropriate sports and activities are suggested for cognitive and social development. The role of parents is to provide a stress-free, nurturing environment, engage in the play, and encourage curiosity through storytelling and conversation. Discussion includes children's learning processes and suitable play activities.

No.	Topic and Session Details
15	**Child-Centered Learning Activities** Children learn best when provided with opportunities and an environment suited to their age, interests, and abilities. This session focuses on child-centered learning that is safe, healthy, inclusive, and stimulating. Child-centered activities include sports, storytelling, music, dance, drawing, and nature observation. These help develop creativity, communication skills, confidence, problem-solving abilities, and leadership. Parents' role is to create spaces for play and ensuring learning aligns with a child's natural curiosity.
16	**Utilization of Local Learning Materials** Discusses the use and availability of homemade toys and locally sourced learning materials in a child's development. Homemade toys crafted from readily available resources are cost-effective and culturally relevant and promote creativity, emotional bonding, and cognitive skills. Locally made toys like dolls, puppets, soft balls, and pictorial storybooks aid in language development, motor skills, and social learning. Parents are responsible for ensuring these materials are safe, age-appropriate, and accessible while creating a supportive play environment.
17	**Early Literacy and Numeracy Skills in Children** Explains the importance of early exposure to language and numeracy skills through everyday activities rather than formal instruction before school. Children develop communication, literacy, and numeracy skills naturally through interactions, storytelling, counting household objects, and recognizing shapes in their surroundings. Early literacy elements include awareness of print and signs, listening, speaking, and recognizing letters, while early numeracy focuses on measurement, sorting, and geometric concepts. Parents play a crucial role in fostering these skills by integrating learning into daily routines, reading to children, and encouraging exploration through play.
18	**Child Development Center and Pre-Primary Classes** Child Development Centers (CDCs) provide early learning experiences that are accessible for all children before formal schooling. Parents are instructed to enroll children above three years and below five years of age at CDCs where they can participate in pre-primary classes and learn in a safe and stimulating environment through activities like storytelling, music, and creative play. Parents can reinforce education received at CDCs by talking to their children about lessons and focusing on holistic development, not just educational progress.
19,20	**Disability Modules** Raise awareness about childhood disabilities—their causes, types, and ways to support children effectively. Interactive methods such as storytelling and simple games are used to help caregivers understand the functional challenges children may face and how to address them with empathy and inclusion. Sessions emphasize the importance of early identification, reducing stigma, and creating supportive environments at home and in the community. Real-life examples are used to spark discussion and reflection, encouraging caregivers to embrace inclusive practices. Participants also share key definitions from Nepal's disability framework and guide families on how to seek relevant services and support.

E Outcome Variables

E.1 Overview

This appendix provides details on the pre-specified outcomes in the pre-analysis plan registered at RCT ID: AEARCTR-0011318.

E.2 Secondary Outcomes: Learning Environment & Practices and Teacher-Student Interaction

Outcome	Definition
Teacher–Student Interaction	*Response Options: 1 if Yes, 0 if No.* *Note:* There are two indices: (i) Positive index *The sum of responses to B1, B2, B3, B4, B5, B6, B7, B9, and B10 across 20 rounds of observation.* (ii) Negative index *The sum of responses to B8, B11, B12, B13, and B14 across 20 rounds of observation.* 1. Reading to a small/large group. 2. Teachers or Students singing. 3. Circle time: Child leading a whole class activity. 4. Back and forth discussion/conversation in a circle time. 5. Back and forth discussion/interaction/conversation with a small group or one child. 6. Engaging in pretend play with children/child. 7. Telling child or children what to do (not back-and-forth discussion). 8. Negative discipline. 9. Calmly negotiating conflict between peers. 10. Smiling/positive affect. 11. Negative affect/yelling. 12. Teacher is sitting at her desk or not engaged with children. 13. Teacher is talking with another adult. 14. Teacher out of the classroom.
Learning Environment and Practices Index	*Coded: 1 if Yes, 0 if No, discarded if N/A* *The index will be the additive sum of the responses* 1. Teacher has prepared yearly, monthly and daily plans. 2. Teachers follows standard daily time table (as learnt during the training). 3. Stories, poems and songs are regularly used in ECD center. 4. ECD center and classroom are regularly cleaned. 5. Access to safe drinking water. 6. Access to Child friendly washroom and toilets. 7. Child-friendly, clean (free from dust); adequate space classroom; Proper seating arrangement in ECD center (Use of carpet, pre fab, cushions, round tables etc. 8. Six learning areas management and arrangement of learning materials accordingly. 9. Language learning areas arrangement and management.

10. Science learning areas arrangement and management.
11. Play/Act/role-play skills related learning areas arrangement and management.
12. Creativity learning areas arrangement and management.
13. Math learning areas arrangement and management.
14. Construction/Materials development learning areas arrangement and management.
15. Use of learning and play materials and conducts activities accordingly.
16. Teacher takes daily attendance of all children (attendance registers maintained).
17. Teacher also conducts appropriate and disable-friendly activities for children with disability.
18. Use of ECD Kit Box in children's daily activities.
19. Regular observation/monitoring/supervision and guidance by HT.
20. Teachers have maintained records on students' progress (as per the prescribed formats).

E.3 Child Primary Outcomes: Ages and Stages Questionnaire Score

Note: *Each response in this subsection will be scored in accordance with the standard methodology used for scoring ASQ in the literature. Aggregate score in this section will be the sum of individual scores for all these questions.*

Aggregate ASQ score is the sum of scores in all five subsections: Communication, Gross Motor, Fine Motor, Problem- solving, and Personal-social.

Subsection of ASQ (36 month)	**Definition**
Communication	1. When you ask your child to point to her nose, eyes, hair, feet, ears, and so forth, does she correctly point to at least seven body parts? 2. Does your child make sentences that are three or four words long? 3. Without giving your child help by pointing or using gestures, ask him to "put the book on the table" and put the shoe under the chair". Does your child carry out both of these directions correctly? 4. When looking at a picture book, does your child tell you what is happening or what action is taking place in the picture (for example, "barking","running", "eating" or "crying"?) 5. Show your child how a zipper on a coat moves up and down, and say, "See, this goes up and down". Put the zipper to the middle and ask our child to move the zipper down. Return the zipper to the middle and ask your child to move the zipper up. Do this several times, placing the zipper in the middle before asking you child to move it up or down. Does your child consistently move the zipper up when you say "up" and down when you say "down"?
Gross Motor	1. Without holding onto anything for support, does your child kick a ball by swinging his leg forward? 2. Does your child jump with both feet leaving the floor at the same time? 3. Does your child walk up stairs, using only one foot on each stair?

4. Does your child stand on one foot for about 1 second without holding onto anything?
5. While standing, does your child throw a ball overhand by raising his arm to shoulder height and throwing the ball forward?
6. Does your child jump forward at least 6 inches with both feet leaving the ground at the same time?

Fine Motor

1. After your child watches you draw a line from the top of the paper to the bottom with a pencil, crayon, or pen, ask her to make a line like yours. Do not let your child trace your line. Does your child copy you by drawing a single line in a vertical direction?
2. Can your child string small items such as beads, macaroni, or pasta "wagon wheels" onto a string or shoelace?
3. After your child watches you draw a single circle, ask him to make a circle like yours. Do not let him trace your circle. Does your child copy you by drawing a circle?
4. After your child watches you draw a line from one side of the paper to the other side, ask her to make a line like yours. Do not let your child trace your line. Does your child copy you by drawing a single line in a horizontal direction?
5. Does your child try to cut paper with child-safe scissors? He does not need to cut the paper but must get the blades to open and close while holding the paper with the other hand.
6. When drawing, does your child hold a pencil, crayon, or pen between her fingers and thumb like an adult does?

Problem-Solving

1. While your child watches, line up four objects like blocks or cars in a row. Does your child copy or imitate you and line up four objects in a row?
2. If your child wants something he cannot reach, does he find a chair or box to stand on to reach it (for example, to get a toy on a counter or to "help" you in the kitchen?
3. When you point to the figure and ask your child, "What is this?" does our child say a word that means a person or something similar? Please write your child's response.
4. When you say, "Say 'seven three'", does your child repeat just the two numbers in the same order? Do not repeat the numbers. If necessary, try another pair of numbers and say, "Say 'eight two'".
5. Show your child how to make a bridge with blocks, boxes, or cans, like the example. Does our child copy you by making one like it?
6. When you say, "Say 'five eight three'", does our child repeat just the three numbers in the same order? Do not repeat the numbers. If necessary, try another series of numbers and say. "Say 'six nine two'".

Personal-Social

1. Does your child use a spoon to feed herself with little spilling?
2. Does your child push a little wagon, stroller, or toy on wheels, steering it around objects and backing out of corners if he cannot turn?
3. When your child is looking in a mirror and you ask, "Who is in the mirror?" does she say either "me" or her own name?
4. Does your child put on a coat, jacket, or shirt by himself?
5. Using these exact words, ask your child, "Are you a girl or a boy?" Does our child answer correctly?
6. Does your child take turns by waiting while another child or adult takes a turn?

Subsection of ASQ (48 month)	Definition
Communication	1. Does your child name at least three items from a common category? 2. Does your child answer the following questions? (Mark "sometimes" if your child answers only one question.) a. "What do you do when you are hungry?" Please write your child's response. b. "What do you do when you are tired?" Please write your child's response. 3. Does your child tell you at least two things about common objects? 4. Does your child use endings of words, such as "-s," "-ed," and "-ing"? 5. Without your giving help by pointing or repeating, does your child follow three directions that are unrelated to one another? Give all three directions before your child starts. 6. Does your child use all of the words in a sentence (for example, "a," "the," "am," "is," and "are") to make complete sentences, such as "I am going to the park," or "Is there a toy to play with?" or "Are you coming, too?"
Gross Motor	1. Does your child catch a large ball with both hands? 2. Does your child climb the rungs of a ladder of a playground slide and slide down without help? 3. While standing, does your child throw a ball overhand in the direction of a person standing at least 6 feet away? To throw overhand, your child must raise his arm to shoulder height and throw the ball forward. 4. Does your child hop up and down on either the right or left foot at least one time without losing her balance or falling? 5. Does your child jumpforward a distance of 20 inches from astanding position, starting with his feet together? 6. Without holding onto anything, does your child stand on one foot for at least 5 seconds without losing her balance and putting her foot down?
Fine Motor	1. Does your child put together a five- to seven-piece interlocking puzzle? 2. Using child-safe scissors, does your child cut a paper in half on a more or less straight line, making the blades go up and down? 3. Using the shapes below to look at, does your child copy at least three shapes onto a large piece of paper using a pencil, crayon, or pen, without tracing? 4. Does your child unbutton one or more buttons? 5. Does your child draw pictures of people that have at least three of the following features: head, eyes, nose, mouth, neck, hair, trunk, arms, hands, legs, or feet? 6. Does your child color mostlywithin the lines in a coloring book or within the lines of a 2-inch circle that you draw?
Problem-Solving	1. When you say, "Say 'five eight three,'" does your child repeat just the three numbers in the same order? Do not repeat the numbers. If necessary, try another series of numbers and say, "Say 'six nine two.'" 2. When asked, "Which circle is the smallest?" does your child point to the smallest circle? 3. Without your giving help by pointing, does your child follow three different directions using the words "under," "between," and "middle"?

	4. When shown objects and asked, "What color is this?" does your child name five different colors, like red, blue, yellow, orange, black, white, or pink? 5. Does your child dress up and "play-act," pretending to be someone or something else? 6. If you place five objects in front of your child, can he count them by saying, "one, two, three, four, five," in order?
Personal-Social	1. Does your child serve herself, taking food from one container to another using utensils? 2. Does your child tell you at least four of the following? 3. Does your child wash his hands using soap and water and dry off with a towel without help? 4. Does your child tell you the names of two or more playmates, not including brothers and sisters? 5. Does your child brush her teeth by putting toothpaste on the toothbrush and brushing all of her teeth without help? 6. Does your child dress or undress himself without help (except for snaps, buttons, and zippers)?

Subsection of ASQ (60 month)	**Definition**
Communication	1. Without your giving help by pointing or repeating directions, does your child follow three directions that are unrelated to one another? Give all three directions before your child starts. 2. Does your child use four- and five-word sentences? For example, does your child say, "I want the car"? Please write an example. 3. When talking about something that already happened, does your child use words that end in "-ed," such as "walked," "jumped," or "played"? Ask your child questions, such as "How did you get to the store?" ("We walked.") "What did you do at your friend's house?" ("We played.") Please write an example. 4. Does your child use comparison words, such as "heavier," "stronger," or "shorter"? Ask your child questions, such as "A car is big, but a bus is ..." (bigger); "A cat is heavy, but a man is ..." (heavier); "A TV is small, but a book is ..." (smaller). Please write an example. 5. Does your child answer the following questions? (Mark "sometimes" if your child answers only one question.) a. "What do you do when you are hungry?" Please write your child's response here: b. "What do you do when you are tired?" Please write your child's response. 6. Does your child repeat the sentences shown below back to you, without any mistakes? a. Jane hides her shoes for Maria to find. b. Al read the blue book under his bed.
Gross Motor	1. While standing, does your child throw a ball overhand in the direction of a person standing at least 6 feet away? To throw overhand, your child must raise his arm to shoulder height and throw the ball forward. 2. Does your child catch a large ball with both hands?

3. Without holding onto anything, does your child stand on one foot for at least 5 seconds without losing her balance and putting her foot down?
4. Does your child walk on his tiptoes for 15 feet (about the length of a large car)?
5. Does your child hop forward on one foot for a distance of 4–6 feet without putting down the other foot?
6. Does your child skip using alternating feet?

Fine Motor

1. Ask your child to trace on the line below with a pencil. Does your child trace on the line without going off the line more than two times?
2. Ask your child to draw a picture of a person on a blank sheet of paper. You may ask your child, "Draw a picture of a girl or a boy."
3. Draw a line across a piece of paper. Using child-safe scissors, does your child cut the paper in half on a more or less straight line, making the blades go up and down?
4. Using the shapes below to look at, does your child copy the shapes in the space below without tracing?
5. Using the letters below to look at, does your child copy the letters without tracing? Cover up all of the letters except the letter being copied.
6. Print your child's first name. Can your child copy the letters? The letters may be large, backward, or reversed.

Problem Solving

1. When asked, "Which circle is the smallest?" does your child point to the smallest circle?
2. When shown objects and asked, "What color is this?" does your child name five different colors, like red, blue, yellow, orange, black, white, or pink?
3. Does your child count up to 15 without making mistakes?
4. Does your child finish the following sentences using a word that means the opposite of the word that is italicized? For example: "A rock is hard, and a pillow is soft." Please write your child's responses.

   a. A cow is *big* and a mouse is ...

   b. Ice is *cold*, and fire is ...

   c. We see stars at *night*, and we see the sun during the ...

   d. When I throw the ball *up*, it comes ...
5. Does your child know the names of numbers?
6. Does your child name at least four letters in her name? Point to the letters and ask, "What letter is this?" (Point to the letters out of order.)

Personal-Social

1. Can your child serve himself, taking food from one container to another, using utensils?
2. Does your child wash her hands and face using soap and water and dry off with a towel without help?
3. Does your child tell you at least four of the following? Please mark the items your child knows.
4. Does your child dress and undress himself, including buttoning medium-size buttons and zipping front zippers?
5. Does your child use the toilet by herself? (She goes to the bathroom, sits on the toilet, wipes, and flushes.)
6. Does your child usually take turns and share with other children?

E.4 Caregiver Primary Outcomes: Index of Stress and Practice of Early Stimulation and Learning

Outcome	Definition
Parent (Caregiver) Stress Index	*Response Options: 1- Strongly disagree; 5-Strongly agree.* 1. My child rarely does things that make me feel good. 2. Sometimes I feel my child doesn't like me and doesn't want to be close to me. 3. My child smiles at me much less than expected. 4. When I do things for my child, I get the feeling that my efforts are not appreciated very much. 5. I expected to have closer and warmer feelings for my child than I do, and this bothers me. 6. Sometimes my child does things that bother me just to be mean. 7. I feel that I am ... (*Response Options: 1-A very good parent; 5-Not very good at being a parent.*)
Practice of early stimulation and learning (Index)	*Response Options: 1 if Yes, 0 if No; The index will be the additive sum of the responses.* 1. Read books or looked at picture books with the child. 2. Told stories to the child. 3. Sang songs to or with the child (including lullabies)? 4. Took the child outside the home? 5. Played with the child. 6. Named, counted, or drew things for or with the child.

E.5 Care Givers' Secondary Outcomes

Outcome	Definition	Coding
Attitude to disciplining the Child	Do you think it's necessary or/and normal to spank to discipline your child?	0 - No, 1 - Yes
Intensity of discipline	In the past 14 days, how often have you spanked to discipline your child?	0 - Never, 1 - Once, 2 - Twice, 3 - More than twice
Intensity of early stimulation and learning (index)	How often did you engage in this activity with the child?	The index will be the additive sum of the responses. 1 - Only several times (1-3 days a week), 2 - Often (4-5 days a week), 3 - Nearly every day (6 days a week)

Attitude to Disability	1. Do you believe children with disabilities should be separated from their parents? 2. Do you think schools should accommodate children with disabilities? 3. Do you think disabilities can be helped? 4. I believe that children with disabilities deserve access to educational opportunities	The index will be the additive sum of the responses. 1: 0 - Yes, 1 - No 2 & 3: 1 - Yes, 0 - No 4: 5 - Strongly Agree, 4 - Agree, 3 - Not sure, 2 - Disagree, 1 - Strongly Disagree
Feeling confident to help with disability	1. I have the necessary knowledge and skills to support children with disabilities 2. I feel confident in my ability to support children with disabilities.	The index will be the additive sum of the responses. 5 - Strongly Agree, 4 - Agree, 3 - Not sure, 2 - Disagree, 1 - Strongly Disagree

E.6 Testing for Complementarities Between Home and School Environments

We formally test whether improvements in classroom quality and parental engagement operate as substitutes or complements in producing child development outcomes. To avoid potential endogeneity concerns, we use predicted classroom quality as a proxy for implementation intensity.

E.6.1 Empirical Strategy

We employ a two-stage approach. First, we predict classroom quality improvements using treatment assignment:

$$\text{ClassFacilities}_s = \alpha_0 + \alpha_1 \text{T}_s + \gamma_m + \nu_s \tag{10}$$

where ClassFacilities_s is the classroom facilities index for school s, T_s indicates treatment assignment, γ_m represents municipality fixed effects, and ν_s is the error term. We obtain predicted values $\widehat{\text{ClassFacilities}}_s$ and generate a binary indicator HighClass_s for schools with above-median predicted improvements.

Second, we estimate a model with triple interaction terms:

$$\begin{aligned} Y_{is} = \beta_0 &+ \beta_1 \text{T}_s + \beta_2 \text{HighPESL}_{is} + \beta_3 \text{HighClass}_s + \beta_4 (\text{T}_s \times \text{HighPESL}_{is}) \\ &+ \beta_5 (\text{T}_s \times \text{HighClass}_s) + \beta_6 (\text{HighPESL}_{is} \times \text{HighClass}_s) \\ &+ \beta_7 (\text{T}_s \times \text{HighPESL}_{is} \times \text{HighClass}_s) + \gamma_m + \epsilon_{is} \end{aligned} \tag{11}$$

where Y_{is} represents outcomes for child i in school s, HighPESL_{is} indicates above-median baseline parental engagement, and β_7 captures the complementarity effect.

Table 8 presents results for child development (ASQ), parental engagement (PESL), and caregiver stress. For child development, the coefficient on High PESL × High LEP in control schools is -0.69 ($p < 0.01$), indicating substitution. The triple interaction coefficient is 0.53 ($p < 0.05$), indicating that treatment shifts the relationship toward complementarity. Combined, these coefficients suggest that in treatment schools, the interaction becomes -0.69 + 0.53 = -0.16, substantially less negative than in controls.

For parental engagement, baseline PESL is the dominant predictor (0.64, $p < 0.01$), and the triple

interaction is small and insignificant (0.36, $p > 0.10$). For caregiver stress, the triple interaction is also small and insignificant (0.09, $p > 0.10$). These null results for secondary outcomes strengthen confidence that the complementarity finding for child development reflects a genuine mechanism rather than spurious correlation.