

RESEARCH

Open Access



Effects of an early childhood father engagement program in Rohingya camps and host community in Cox's Bazar, Bangladesh: a cluster randomized controlled trial

Yeshim Iqbal^{1*}, Hirokazu Yoshikawa², Duja Michael², Sneha Bolisetty³, Sakila Yesmin⁴, Ashraf Uddin Mian⁵, Maung Ting Nyeu⁶, Sangyoo Lee⁷, A. Brooks Bowden⁸, Kate Schwartz², Kuri Chisim⁴, A. R. M. Mehrab Ali⁹, Sadia Sumaia Chowdhury⁹, Anika Alam¹⁰ and Jere R. Behrman⁸

Abstract

Background This study evaluates a 6-month early childhood father engagement intervention delivered in-person to fathers in the Rohingya camps and surrounding host communities in Cox's Bazar, Bangladesh. The intervention is an added component to an existing intervention that works with mothers of 0–3-year-old children to improve positive outcomes of child development and targets fathers with children below 3 years, with objectives to promote fathers' wellbeing by improving their emotional literacy, encouraging fathers to strengthen relationships with their spouses and children, and encouraging responsive and stimulating caregiving practices among fathers.

Methods We used a cluster-randomized controlled trial design (total $N = 2002$ fathers, 786 in the Rohingya camps and 1216 in the host communities) to assess the impact of the program on fathers' parenting and engagement with family (reported by both fathers and mothers), fathers' mental health, and child development.

Results We find that the program has a positive impact on the fathers' parenting (father reported) and engagement with family (father and mother reported), as well as on father-reported child social-emotional development, compared to the mother-only program. While the program did not have a main effect on directly assessed child development outcomes, we do find that baseline mother-reported stimulation with child and mother health moderate the impact on father-reported child social-emotional development in the camp community (larger positive impacts for fathers from households with lower baseline mother-reported stimulating behaviors and health). We also find that child gender and mother-reported stimulating behaviors moderate impacts on multiple mother-reported child developmental outcomes in both host and camp communities (treatment effects on these outcomes were negative for boys and positive but weak for girls, and larger positive impacts on mother-reported motor development were found among mothers from households who reported lower stimulation at baseline). Additionally, child age moderates the impact on fathers' collaboration with the mother (reported by the mother), such that the treatment effect is larger for mothers of older children.

Conclusions Our findings demonstrate the potential of father-focused programs as powerful caregiver interventions.

*Correspondence:

Yeshim Iqbal

yeshim.iqbal@dukekunshan.edu.cn

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



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Trial registration Pre-registered with REES (13880.1v2); retrospectively registered at ISRCTN (ISRCTN13510859).

Keywords Early childhood development, Father engagement, Rohingya, Bangladesh

Background

Throughout early childhood, fathers' engagement with children positively influences child development, with fathers' extent and quality of involvement with their children predicting positive development across multiple domains [1, 2]. Most existing research on father engagement and parenting interventions occurs in western, educated, industrialized, rich democracies [3], highlighting the gap in the knowledge base on fathering and father-focused interventions in LMICs or humanitarian contexts. Research on fathers who have been forcibly displaced is crucial: over 110 million individuals are displaced worldwide [4]. Understanding fathers' experiences can provide insights into how this very large population copes with uncertainty, trauma, and displacement. These families in settings of forced displacement may experience changes in traditional gender roles and family and community structures, and understanding their experiences can provide insights into newly developing family dynamics and childcare practices.

Father involvement early in their children's development plays a crucial role in later development. Research in high-income countries, including recent meta-analyses, shows that prenatal and early childhood father involvement with children predicts a range of cognitive, language, and social-emotional outcomes. Prenatal involvement of fathers with pregnant mothers, for example, can predict engagement with children post-birth [1], and research in infancy finds that fathers' involvement correlates with children's sense of security at later ages [5]. Fathers' involvement in early childhood is associated with higher levels of cognitive development [6–8], early numeracy [9], and pro-social skills [10]. Father attachment relationships also play a role in child social competence and emotion regulation [11]. Fathers may also indirectly affect children through their communication and collaboration with the mother [12]. The research in LMICs is comparatively sparse but also suggests positive associations of early father involvement with early literacy [13] and socio-emotional development [14], and a global analysis found father-child stimulation to be associated with higher levels of mother-reported child development [15].

Several dimensions of fathers' wellbeing and involvement with their families are important in predicting child development outcomes. The first is *father engagement* (i.e., direct interaction) through caregiving and stimulating activity/play with children. Two reviews, examining

father-child play [16] and stimulating and challenging activities [17] indicate positive associations of this kind of involvement with children's cognitive, psychosocial, and physical development. Similarly, a review of longitudinal studies [18] provides evidence for fathers' direct interaction with the child as having effects on children's social, behavioral, and psychological outcomes. A second important dimension is the *relationship between fathers and mothers*; in rural Kenya, fathers' support to mothers was found to be positively associated with children's outcomes, with fathers' influence on children partially mediated by maternal outcomes [19]. Third, some recent research has shown associations of *fathers' beliefs about their role and the importance of play* with child outcomes [20]. One study in Brazil showed associations between fathers' child socialization beliefs and parenting processes [21]. The final dimension is *fathers' depression and anxiety*. A large population-based study showed that depression in fathers in the postnatal period is associated with negative emotional and behavioral outcomes in children aged 3–5 years; fathers' depression is also associated with an increased risk of disharmony in partner relationships, reported by both fathers and their partners [22]. Additionally, using a longitudinal household survey in China, Jeong and Li [23] found that paternal depression was negatively associated with children's socioemotional development.

A systematic review [24, 25] focusing on early childhood “father-inclusive” interventions (those involving male caregivers as program participants and focused on improving these caregivers' nurturing care for young children) in LMICs found family-wide benefits of these interventions for women, children, and fathers. This review found 44 studies that directly engaged fathers/male caregivers in LMIC contexts. Six of the 19 studies that directly assessed paternal outcomes found positive impacts, including on stimulation, attachment, and various childcare activities [24]. Two of these six programs also led to reductions in fathers' physical punishment of their children [26, 27].

A few experimental studies have evaluated interventions developed specifically for fathers in LMIC contexts. Abimpaye et al. [28] examined a program in Rwanda that focused on health and nutrition, responsive and stimulating parenting, and book sharing for mothers and fathers of children aged 6–24 months. Fathers and mothers in treatment groups reported engaging in significantly more learning/play, nurturing/care, and positive discipline

behaviors than parents in a control group; however, this study did not include child outcomes. Betancourt et al. [29] examined an intervention that covered early stimulation, play, nutrition, hygiene, responsive parenting, and nonviolent interactions for both male and female Rwandan caregivers. The program included a 12–16 week home-visit program for families experiencing severe poverty with at least one child of 6–36 months. Results showed improvement of caregiving practices, including parent–child interactions and stimulation, nutrition, care seeking, and reduced violence, but child development outcomes were not reported. Jeong et al. [30] evaluated a 12-month program for couples with children below 18 months in Tanzania. Villages were assigned to one of five groups: a nutrition intervention for mothers; a nutrition intervention for couples; a nutrition plus parenting intervention for mothers; a nutrition plus parenting intervention for couples; or a standard-of-care control. The mother-only model involved peer groups of mothers, while for couples, there were separate peer groups for mothers and for fathers. Results indicated that compared to conditions with mothers only, engaging fathers improved paternal stimulation but did not improve the child development outcomes (cognitive and language development) more than engaging mothers only.

There are a few takeaways from these studies and the meta-analysis by Jeong et al. [31]. First, involving fathers can lead to positive outcomes for fathers and mothers, but this is somewhat more inconsistent (and assessed more infrequently) for child development outcomes.

Second, Jeong et al. [25] note, importantly, that programming almost always incorporates fathers into existing interventions intended for mothers, rather than designing them specifically for fathers themselves. Among the 23 programs including both mothers and fathers, only two interventions tailored the delivery for fathers [25]. Third, most father-directed programming focused on involving fathers in reproductive/prenatal care (programs conducted in Bangladesh, Brazil [32], Burkina Faso [33], Kenya [34], Nepal [35], Korea [36] and Rwanda [27]) or on reducing intimate partner violence or child maltreatment (e.g., in Tanzania [37], Colombia [38], and China [39]). Very few programs assess other components such as father–child stimulation or warmth. Finally, few of the programs for fathers reviewed were implemented in a context of forced displacement or migration. In these contexts, families undergo rapid changes in and exposure to different social norms. In the Rohingya context, for instance, for those residing in camps in Bangladesh, official employment is not permitted, resulting in an increase in men's available time and presence in the household. These changes have significant implications for how fathers interact with their families, making these

settings especially pertinent for father-based interventions. As of yet, no studies have focused on interventions for fathers in the Rohingya population.

Due to statelessness, discrimination, and violence, the Rohingya community is considered one of the most persecuted minorities in the world. Originally residing in northern Rakhine State, Myanmar [40], over several decades, many fled the region to Bangladesh due to violence [41], with the largest exodus occurring in the months following August 2017. This influx to Cox's Bazar district included displacement and migration of approximately 750,000 people, including more than 400,000 children, and occurred following a military crackdown in Myanmar. Camps in Cox's Bazar, Bangladesh are now home to over 900,000 forcibly displaced Rohingya people [42].

In addition to stresses and changes that displaced populations go through, host communities also encounter stresses where a large influx of refugees occurs. The host population living alongside the Rohingya in Cox's Bazar district has one of the highest rates of poverty in Bangladesh; this community has become increasingly vulnerable due to competition for scarce opportunities with the Rohingya [43]. This population is majority Bengali and Muslim but the region is also home to several different ethnic and religious minorities. Most of the population speaks Chittagonian Bengali, a dialect thought to have 70–90% overlap with the Rohingya dialect [43].

The present study is an impact evaluation of an intervention for fathers of 0–3-year-old children in the Rohingya camps (one of the largest groups of forcibly displaced people in the world) [44] and surrounding host communities in Cox's Bazar District, Bangladesh. We conducted a cluster randomized trial to examine the added impacts of the 6-month program for fathers, administered through home visits and group meetings, on fathers' parenting and engagement with the family, father mental health, and children's development, in comparison to families receiving an intervention for mothers only. The intervention we examine here is a 6-month parenting intervention for fathers delivered in person by male volunteers, through weekly home visits and monthly group sessions, in the Rohingya camps and surrounding host communities in Cox's Bazar. The program was developed and implemented by BRAC-IED and Sesame Workshop as part of the Humanitarian Play Labs, a key component of the larger "Play to Learn" program funded by the LEGO Foundation and aiming to establish play-based early childhood development as an essential component of humanitarian response. We from New York University Global TIES (when this research was being conducted) served as the independent research partner and conducted an evaluation of the father-focused program. The program was designed

as a complement to an existing program for mothers of 0–3-year-old children. The impact evaluation employed a cluster randomized control trial in which the caseloads of 375 mother volunteers (a total of 2,002 families) were randomized such that families received the mothers' intervention only in the control condition and received both the fathers' and mothers' intervention in the treatment condition.

All the primary (1–3) and exploratory (4–5) research questions, together with the analytic approach, are defined below as per the trial protocol (see Supplementary material 1).

1. What is the added impact of a parenting intervention, delivered by male volunteers and consisting of home visits (every week) and group sessions (every 3 weeks) to fathers of children aged 0–3, on fathers' parenting and engagement with family, in comparison to families who only receive a parenting intervention targeting mothers?
2. What is the added impact of this program on fathers' mental health (depression and anxiety)?
3. What is the added impact of this program on child development (cognitive, language, motor, and social-emotional domains)?
4. Are effects of the intervention on hypothesized outcomes in Research Questions 1, 2, and 3 moderated by camp/host community residence or baseline father education, household resources, mother health, child age/gender, or mother report of mother–child stimulating behaviors?
5. What resources are delivered through the 6-month father engagement model and what is the cost of delivering the program? What is the additional cost of the father engagement model relative to the mother program?

Methods

The trial was pre-registered at the Registry of Efficacy and Effectiveness (<https://sreereg.icpsr.umich.edu/sreereg/subEntry/20180/pdf?action=view>; Aug 9, 2023) and retrospectively registered at ISRCTN (<https://www.isrctn.com/ISRCTN13510859>, Nov 20, 2025). The paper reports according to CONSORT (CONSolidated Standards of Reporting Trials) guidelines [45]; please see Supplementary material 2 for the CONSORT checklist.

Intervention

The program is a father-focused parenting intervention component added to an existing intervention for mothers of 0–3-year-old children to improve child development outcomes. The father intervention was originally designed for the Rohingya camp context and

then expanded to the host community, in Cox's Bazar Sadar, Ramu, Ukhiya and Teknaf regions. In both control and treatment conditions, the mothers in the families received a weekly group-based mothers' intervention, conducted by female volunteers. Each hour-long session included objectives to increase caregiver roles in child development, provide an opportunity to share child-rearing experiences, create supportive networks, promote play, and recognize the importance of mental health (BRAC & Sesame Workshop. Humanitarian Play Lab Parenting Curriculum Under Child Protection: Home Based Parenting: Fathers' Engagement Model for 0–3 years old children and their father. [Manual]. 2022, unpublished). In addition to the mothers' intervention, the fathers in treatment group families received a weekly fathers' parenting intervention, delivered by male volunteers. The objectives of the curriculum for fathers were to promote fathers' wellbeing by improving their emotional literacy, encourage fathers to strengthen relationships with their spouses and children, and encourage responsive caregiving practices among fathers (BRAC & Sesame Workshop. Humanitarian Play Lab Parenting Curriculum Under Child Protection: Home Based Parenting: Fathers' Engagement Model for 0–3 years old children and their father. [Manual]. 2022, unpublished). While many of the objectives of the fathers' and mothers' interventions were similar (strengthening relationships, the importance of play, and attention to wellbeing), the modality of delivery differed based on fathers' daily routines. For instance, during the one-on-one sessions (which did not take place for mothers), father volunteers delivered content in relation to these topics, taught the participating father activities to do with the child (involving the child when present; though the child's presence was not a required component), gave the father opportunities to discuss anything on his mind, and helped connect to other services as needed. Importantly, the weekly sessions were scheduled based on the father's availability; three sessions each month were home visits, and one each month occurred in group format. This format was vital to allowing fathers to participate (due to the flexible one-on-one scheduling) while also providing a group component.

The intervention, designed by BRAC-IED with support from Sesame Workshop, was developed in close consultation with the Rohingya through an extensive co-creation process; content, scheduling, and modality were piloted and developed through multiple iterations of testing and formative research (Yesmin S, Begum T, Gani O, Islam MJ, Das B, Rahman A, & Ruma UR. Formative study on fathers' engagement model in the camp. BRAC Institute for Educational Development. 2022, unpublished). BRAC-IED also monitored fidelity of program delivery, session quality, and competencies

of father volunteers, primarily through observation of ongoing sessions and using a simple fidelity instrument developed for this purpose. Feedback on these sessions was provided through regular meetings with father volunteers and in refresher trainings. As a separate process from BRAC-IED's monitoring of implementation, our team from NYU-TIES designed and used an observational measure of quality of the father home-visit sessions, which was collected approximately 4 months into the intervention; findings from these measures have been detailed elsewhere (Friere S, Ramachandran A, Iqbal Y. Implementation in Unequal Contexts: How Setting Shapes Implementation in Early Childhood Home-Visits for Rohingya and Bangladeshi fathers, unpublished) [46] (Iqbal, Y., Yoshikawa, H., Michael, D., Bolisetty, S., Schwartz, K., Yesmin, S., Mian, A., Nyeu, M., Sunny, S., Ramachandran, A., Hilgendorf, D., Tiwari, P., Barrios, S.F., Chisim, K., Mur Salin, M. S., Hossain, N., Zaman, S.S., Ali, A. R. M. M., Chowdhury, S., Wuermli, A. Impact Evaluation of a Father Engagement Model in the Home Visiting Humanitarian Play Lab (HPL) Program in Rohingya Camps and Host Community in Cox's Bazar, Bangladesh. 2024, unpublished).

Although the Rohingya camp and Cox's Bazar host communities received the same intervention curriculum and delivery formats, there were differences in how it was implemented. First, the camp community participants received a booklet of images reinforcing the curriculum activities; the host community participants did not. Second, the original design of the curriculum assumed a 10-month intervention for mothers. The camp community mothers received this 10-month intervention, but this material was condensed into 6 months for host community mothers. Third, the mothers' intervention had existed in the camps for several years, with the fathers' component being added and evaluated for this study; however, in the host community, parents had not been receiving any intervention at all. In the camp community, the mothers' intervention was already running (with, on average, four or 5 months of the 10-month session having been implemented) when the fathers' intervention began. In the host community, the mothers' and fathers' intervention started simultaneously.

The parenting interventions were provided by trained volunteers who received a stipend for their work (in many humanitarian contexts due to regulations related to formal employment, the use of the term "volunteer" combined with a stipend conforms to local employment regulations). Men (referred to as "Father volunteers") provided the intervention to fathers and women (Mother volunteers) provided the intervention to mothers. All volunteers were selected from the communities that the intervention families belonged to. They received

a 5-day training on the curriculum, followed by a daylong monthly refresher training. All volunteers worked under the guidance and supervision of Bangladeshi BRAC-IED-trained "paracounselors". Paracounselors served as the primary supervisor for father and mother volunteers working with participants, in addition to providing psycho-social support for complicated cases, helping the volunteers manage stress, and understanding regulations for protection of both the volunteer and participant (BRAC & Sesame Workshop. Fathers' Engagement Model: Basic Training Guideline: Father Volunteer - Humanitarian Play Lab - 0–3 Years, unpublished).

The beneficiaries were divided into groups based on the volunteers' geographical locations. Each "geographical group" had approximately 10–14 fathers, and each father volunteer served 4 groups, conducting individual home visits each week with all fathers from these 4 groups. Once a month, each group of fathers met together. One-on-one home visit sessions were 30 min long, and in the case of group sessions, 45 min to 1 h. One-on-one sessions provided a space for the volunteer to deliver the content to fathers in a focused way and for fathers to share, in a comparatively private setting, whatever they wanted to discuss. The group sessions allowed fathers to build relationships with one another and the volunteer to set group norms.

Piloting

All research protocols used in this study were extensively piloted [47, 48] for appropriateness, cultural relevance, and ease of implementation, and reviewed by all relevant members of the research and program implementation teams. The supplementary research protocol provides additional detail regarding procedures for the study [see Supplementary material 1].

Procedure

Enumerators and enumerator training

Baseline survey training and data collection were conducted with 80 enumerators. The endline survey also had a full team of 80 enumerators (60% returners from baseline), working in both camp and host communities. Approximately 70% of these enumerators were women. Most of them were in the process of obtaining their Bachelors' degree, with their ages ranging mostly from 21 to 24 years. All of them were fluent in Standard and Chittagonian Bengali. During data collection, enumerators were accompanied by a team of 11 field supervisors, with each supervisor therefore being responsible for 7–8 enumerators.

Survey training (for all measures collected from the father and mother) was approximately 10 days long, covering the basic theory and concepts behind the

constructs being assessed, the specifics of the measures and response types themselves, the consenting process, protocols and logistics for field operation, and usage of the device and software. One important aspect of enumerator training in this context is the function it serves to clarify language use in conducting the surveys. Rohingya, the primary language of the Rohingya population, is 70–90% mutually intelligible with Chittagonian Bengali, the dialect of Bengali spoken in Cox's Bazar and the native language of the enumerators. Neither the Rohingya dialect nor Chittagonian Bengali has a written script and are therefore translated in the moment by the enumerator reading the survey text (which is in Standard Bengali). In training, enumerators repeatedly practiced this process out loud, clarifying and standardizing the spoken language to be used and, where necessary, adding transliterated words to the surveys.

The training for administering the directly assessed child development outcome, using the Bayley-4 [49], involved 27 enumerators, selected for training based on their experience (such as having worked with children in a learning setting) and willingness to commit to a month-long training followed by data collection. A 5-week training was then conducted, divided into four parts: (a) theoretical training on the cognitive and language domain of the assessment, (b) introduction of the Bayley-4 assessment tool to practice administering the tool item by item, (c) mock practice in pairs among the enumerators with part or whole assessment and (d) hands-on full test practice with children at the training venue. By the end of the hands-on training, a reliability test was conducted as a measure of ensuring quality in which the gold-standard trainers parallel scored at least three Bayley assessments with each enumerator. The aim was to secure more than 85% match in scoring of the trainer and the enumerator (91.07% was achieved). The results of the parallel scoring helped select the final 18 enumerators who met quality standards to conduct the final data collection.

Sampling strategy

Multi-level power analyses indicated that we need a total sample of 2000 to be sufficiently powered (minimum detectable effect size 0.17) to answer our confirmatory research questions. Our sampling and randomization strategy made use of the infrastructure and provision of the mothers' intervention that both treatment and control groups were receiving.

Baseline data collection and randomization

For both the camp and host communities, sampling of fathers for baseline data collection occurred from

databases provided by BRAC-IED of families eligible for inclusion in the study. For the camps, this database included (1) families involved with BRAC-IED programming who had children below 24 months; (2) pregnant mothers who would age into the intervention (among the Rohingya virtually all mothers are married; there are no single-parent families with young children). For the host communities, the database included (1) households geographically proximal to locations with BRAC-IED-provided early childhood development (ECD) programming having at least one child under the age of 24 months; (2) fathers reported as planning to reside with the family for at least the next 8 months (the Rohingya in the camps are restricted in their mobility, while the host community is not). Across both communities, because of difficulties locating fathers and tight timelines for data collection, families that had fathers available were prioritized for inclusion into the baseline sample. Although this likely biases the representativeness of our sample towards fathers who were less employed, employed near home, or otherwise near their residences during hours of data collection, this was the only feasible way of completing data collection in this context. From these initial databases, mother volunteers' existing caseloads were used to develop the baseline sample and were the unit of randomization. In the camps, we sampled 250 randomly selected mother volunteers (out of 350 total mother volunteers) and 1 household from each of the 4 groups the mother volunteer supervised, leading to a total of 1000 focal children's fathers and mothers across 27 (of 32) camps. In the host community, we sampled all 125 mother volunteers, and 4 households from each of the 4 groups a volunteer supervised, leading to a total of 2000 focal children's fathers and mothers. The host community had twice the necessary sample for the study at baseline; this was a risk mitigation strategy in case restrictions from local authorities prohibited endline data collection in the camps as a non-essential activity in this humanitarian context. We conducted randomization at the mother volunteer caseload level, with each mother volunteer (and her caseload) randomly allocated to treatment or control; husbands of the mothers assigned to these mother volunteers in the treatment group received the father intervention. In the camps, the caseloads of 125 of 250 mother volunteers were allocated to treatment; here we performed constrained randomization to ensure covariate balance between arms on all potential confounders. Constrained randomization [50] involves generating many possible allocation schemes and calculating a balance score that assesses covariate imbalance for each of those schemes. It then limits the randomization space to a pre-specified percentage of candidate

allocations before randomly selecting one scheme to implement. In the host community, the caseloads of 63 of 125 mother volunteers were assigned to treatment. Here, given the large baseline sample and the high number of sampled households per mother volunteer, we performed simple randomization. We ran balance checks for both camp and host samples on all collected mother volunteer characteristics (age, education, tenure in the organization) as well as household demographics (mother and father age, education, literacy, and mother health) using the standardized difference approach. All variables had a standardized difference of less than 0.08, suggesting that randomization was successful. Camp baseline data collection was conducted from September 26, 2022, to October 13, 2022. Host community baseline data collection was conducted from December 3 to 19, 2022. Participants were compensated for their time (fathers and mothers were compensated separately for their individual surveys, as well as for a third visit for the direct child assessment) at approximately 200 BDT for the surveys and 300 BDT for the child assessment (~ 2 or 3 USD respectively). This took the form of cash in the host community and in-kind (noodles/biscuits) in the camps, as per camp regulations.

Endline data collection and attrition

After completion of the 6-month intervention, we collected endline data in person both in the camp and host communities (July and August 2023). In the camps, two of the Rohingya camps (a total of 16 father volunteers and 124 households that had been included in baseline) received the intervention on a substantially delayed schedule; these camps were removed from the sample to avoid delaying completion of endline data collection. Besides those 124 households that were dropped from the sample due to “loss of eligibility”, an additional 90 households had attrited from the camp sample because the fathers could not be located for data collection at endline (10.3% of the total sample from the camps that were slated for inclusion at endline). Ultimately, we surveyed 786 of the 876 (~ 90%) of families from baseline that were slated for inclusion at endline.

For the host community, we had a larger sample at baseline than we required at endline. We therefore randomly selected 2 households per group (rather than the 4 that were in baseline), then added an additional 124 given that we had lost 124 from the camp sample, to create a sample of 1124 that we attempted to collect data for. Of these 1124 households, we successfully carried out the endline father survey with 940 (~ 84%). We then replaced

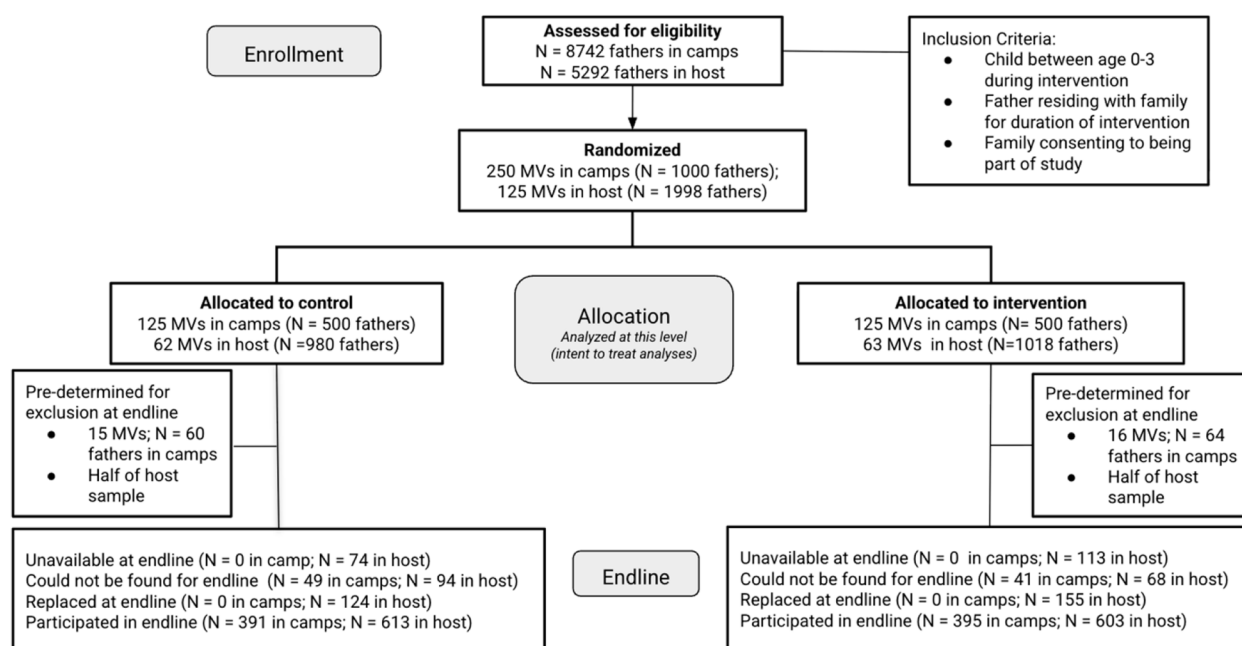


Fig. 1 CONSORT flow diagram of recruitment, allocation and follow-up. Notes. Diagram is based on numbers for completed fathers; mothers and child numbers differ slightly. MV = Mother Volunteer. ‘Pre-determined for exclusion’ includes families: (1) from camps that received the intervention on a delayed schedule and were dropped from the study (camp 8w and 22; a total of 124 families), (2) the half of the host sample that were not included in endline. ‘Unavailable’ refers to participants whose families were found but could not participate for some reason, e.g., the father had moved away, was ill, or was living away from his family due to seasonal labor. ‘Could not be found’ refers to families who could not be tracked down again after baseline

the 184 we could not reach with other host households that we had baseline data for, and added a further 276 households (from the baseline sample that had not originally been selected for inclusion) to have a slightly larger host sample. (We chose to oversample host households, above and beyond these replacements to offset some loss of power due to the attrition of camp households.)

Ultimately, we ended up with 786 families from the camps and 1216 from the host. We ran attrition tests in both the camp and eligible host samples, and we found no evidence of differential attrition (the treatment indicator does not explain attrition in either group, with or without controlling for all baseline variables used in the analyses). We also found no evidence that the characteristics of those who attrited in the host community were statistically significantly different from those that were added to the sample across any of the covariates.

Figure 1 summarizes the recruitment, randomization, and sample sizes of treatment and control groups in the study.

Measures

For both baseline and endline, only measures that were absolutely required of the father were asked of him to keep the survey as short as possible (given the challenges of their unpredictable work schedules). A larger range of child development outcomes and basic information about the household and environment were therefore reported by the mother only.

Outcomes—parenting and engagement with family

Engagement with child and relationship with wife. We used a set of items about the fathers' engagement with the child, communication with the wife, and involvement in daily household activities. Many of these items were drawn from the Father Involvement Questionnaire (FIQ), which was developed for a study on paternal involvement amongst refugee families in Lebanon [51]; some were adapted from a qualitative study on fathers' roles/coparenting in Pakistan [52]. Both fathers and mothers answered these items, with the father responding about himself and the mother responding about her husband. For fathers, the prompt for the item was, "How often in the last week have you/helped [child name] do the following," while for mothers the prompt was about how often her husband had done the following. Participants answered on a 5-point scale ranging from "not at all" to "every day of the week." Baseline psychometrics indicated a six-factor substructure: (1) collaboration with wife about child (8 items, example: "I share my views about our child with my wife"), (2) physical support for wife (4 items, example: "I physically helped my wife with something when she was tired"), (3) warmth/play (3

items, example: "I hug and cuddle my child"), (4) harsh discipline (3 items, example: "I hit [child name] when disciplining"), (5) Responding to child needs and ensuring safety (7 items, example: "I intervene when my child is about to do something dangerous" and (6) communication with wife (4 items, example: "I speak to my wife in a respectful manner"). This factor structure held for both fathers' and mothers' responses and was confirmed at endline using separate confirmatory factor analyses (CFAs) for father and mother-reports. We run and report these outcomes as mean scores (we also ran the factor scores as outcomes. The coefficients on the treatment effects and the significance levels are comparable using both methods. We chose to report mean scores to allow for more straightforward comparisons with other studies).

Beliefs about family engagement, fathering and play (father reported). Participants answered 16 items (developed by our research team) on a 5-point scale ranging from "not at all important" to "extremely important"; example items include: "how important is it for a father to cuddle with the child", "how important is it for a father to talk to his child about feelings", and "how important is it for a father to talk to other fathers about parenting". Baseline psychometrics did not indicate a multi-factor structure emerging from these items; we therefore used these items as one scale (mean score); $\alpha = 0.84$.

Caregiver-child stimulating behaviors (mother and father reported). Fifteen items assessing caregiver-child stimulating behaviors with the child were developed for this study (example item: "In the last week, how often have you counted objects with your child") Items included general activities such as counting objects or naming colors, to more culturally specific activities including playing tuki or reciting kabbiya with the children. Responses were on a 5-point scale ranging from "not at all" to "every day of the week." This instrument was used with both mothers and fathers, with participants responding about themselves. The father-reported measure was used as an outcome and the mother-reported baseline measure as a moderator. Eleven of the 15 items held together well at baseline for both mother and father reports; these were asked again at endline. Items that had loadings < 0.4 at endline were dropped and we generate and use the mean scores for mother and father reports in the analysis (alphas are 0.73 and 0.68 respectively).

Outcomes—mental health

The PHQ-8 depression scale was used in both the mother and father surveys to assess depressive symptoms. This widely used measure has been validated in several contexts including with the Rohingya population [53]; it was also tested for validity and reliability during our pilot

process [47]. The prompt for all items was “over the last month, how often have you been bothered by any of the following things” (example item: trouble falling asleep, staying asleep, or sleeping too much) and participants answered on a 4-point scale (“not at all” to “almost every day”). In line with previous use of this scale, factor analysis indicated a 1-factor solution and supported use of mean scores ($\alpha=0.77$ for fathers, 0.75 for mothers at endline).

Additionally, the GAD-7 was used to assess generalized anxiety for fathers only at endline; this has also been used in several contexts worldwide, including Bangladesh [54]. The prompt for all items was “over the last 2 weeks, how often have you been bothered by the following problems” (example item: feeling nervous, anxious, or on edge). Participants answered on a 4-point scale ranging from “not at all” to “nearly every day”. In line with previous use of this scale, factor analysis indicated a 1-factor solution and supported use of mean scores ($\alpha=0.77$).

Outcomes—child development

To measure child development, we used the Caregiver-Reported Early Development Instrument (CREDI) and the Bayley Scales of Infant and Toddler Development (Bayley-4) at endline. The CREDI is a caregiver-reported, cross-culturally comparable, population-level measure of ECD designed to function across a variety of cultural, linguistic, and socioeconomic contexts. and has been validated in 17 low-, middle-, and high-income countries, including Bangladesh [55]. Participants answer a series of questions about their child with yes, no, or do not know responses. The questions are categorized chronologically by age with start/stop rules. The participant continues answering questions, going up in age, until the participant answers “no” or “don’t know” 5 times. Because of a mistake in how data collection took place, baseline CREDI data were unusable. At endline, mothers reported on the full long-form CREDI, while fathers reported only on the social-emotional (SEL) subscale. Overall and domain scores were obtained using the CREDI scoring app. The internal consistency of all CREDI domains/subscales—cognitive, motor, language, and social-emotional—show high Cronbach’s alphas, ranging from 0.96 to 0.98. We also found a strong correlation between the CREDI subscales of cognitive, language, motor, and social-emotional, with correlation coefficients ranging from 0.88 to 0.96, all significant at $p<0.001$. As recommended by the measure authors, the models estimating treatment effect on the overall CREDI and the domains use the inverse of the squared standard errors of the score as analytical weights to account for the uncertainty in producing those scores.

The Bayley Scales of Infant and Toddler Development 4 (Bayley-4) was used as the direct measure for child development. It is the most frequently used direct assessment for this age range in global parenting program evaluations [31]. In the present study, we measured only the cognitive and language development domains. The Bayley Scales had not been used with Rohingya children before, which called for rigorous piloting and adaptation of the tool, which we conducted prior to beginning the study [47]. We conducted analyses using the raw scores for the Cognitive, Expressive Language and Receptive Language scales, controlling for age. In both the camp and host communities, domain-level Bayley-4 mean scores showed variability and approximated normal distributions, reflecting good between-child discrimination. When.

separated into age-bands, the stratified data yielded increasing domain level scores by child age band for all domains, showing a strong linear developmental progression for each. We also found a strong correlation between the raw scores of the Bayley-4 cognitive scale, expressive language, and receptive language, with correlation coefficients ranging from 0.79 to 0.84, (all correlations significant at $p<0.001$). In addition, the internal consistency of cognitive and language items within each domain ranged from 0.88 to 0.90, indicating high reliability.

We chose to include two different measures of child development, one self-reported by caregivers and one directly assessed, as two of the principal ways that early childhood development has been assessed globally (e.g., the CREDI has been used extensively and is one of the bases for the recently developed WHO/UNICEF Global Scales for Early Development (GSED); the Bayley is the single most frequently used outcomes measure in LMICs for parenting intervention for 0–3 year olds [31]). In this sample, correlations between the Bayley and the CREDI were not significant, a discrepancy potentially stemming from the fundamental difference between caregiver-reported and directly assessed measures of child development; caregivers may be prioritizing or interpreting developmental milestones through a cultural lens that does not align with the Bayley framework.

Covariates

The covariates included, for both father and mother: age, school type (no school – reference category; public school; private school; religious school; NGO; or other), highest grade completed, literacy (no; yes—moderately; yes—well); for child: gender (1=female, 0=male), age in months; for mother only: general health (how they rated their overall health on a 5 point scale ranging from “very bad” to “very good”), whether or not she had suffered

serious injuries/illnesses in the last year, whether she was pregnant; the number of children under the age of 3 in the family, financial concerns (how often did they worry about having enough money in the last 30 days on a 5 point scale ranging from “never” to “always”), and the two factors from the *Perceived Refugee Environment Index (PREI)*: housing quality and household resources, described below. Of these covariates, child gender, child age, a father education index (see analytic approach for details), the two PREI factors, and mother health were also tested as moderators. Mother-reported stimulating behaviors, which was used also as a moderator, was only added as a covariate in the models testing that moderation.

The *Perceived Refugee Environment Index (PREI)*, was used to understand participants’ housing resources (collected from mothers at baseline only). Developed by Pluess and colleagues [56], this measure assesses household resources and social supports available to families. Participants are prompted to answer, on a 4-point scale, how often in the last 30 days they felt a specified item in their environment (e.g., access to clean drinking water whenever you need it). Psychometrics indicated a two-factor structure for the PREI items, specifically, subscales of “Household Resources” (9 items, e.g., “How often in the last 30 days did you feel that your household have enough income or resources to pay for basic needs like food?”) and “Housing Quality” (5 items focusing on how impervious the housing was to heat, rain, wind, e.g., “How often do you feel that your accommodation provides protection from rain?”).

within community, the magnitudes of covariate associations with outcomes could vary across communities and threaten the validity of our coefficients. To eliminate this threat, we did the following: (1) pooled the data from both communities into one dataset and created an indicator variable identifying whether a certain household is from the camp or host community; (2) interacted the community indicator variable with all the covariates in the model (this gives identical coefficients as running the two samples separately but it preserves power); (3) tested the significance of each interaction using a t-test. All interactions with a p -value greater than 0.1 were deemed not meaningfully different and dropped from the estimating equation for parsimony. To ensure a consistent set of interaction terms across all outcomes, we kept the interaction terms in all models (i.e., for each outcome) if it appeared with a p -value less than 0.1 in *any* of the outcomes. Ultimately, this led to keeping all the interaction terms.

The model controlled for covariates listed above and the baseline measure of each outcome when available (i.e., for all outcomes except the CREDI and the Bayley-4). The models in which the Bayley-4 was the outcome also controlled for enumerator effects. We accounted for clustering in both the camp and host sample at the mother volunteer level as well as at the geographical group level (where the “geographical groups consist” of the 10–14 families served together in one group by the mother volunteer). We used a mixed effects model, estimated through restricted maximum likelihood (REML), to identify the treatment effects for each outcome.

$$y_{ijk} = \beta_0 + \alpha tx_k + \gamma camp_k + \beta_{Ncov} N_{ijk} + \delta_{Mcamp} * covM_{ijk} + \zeta_k + \mu_{jk} + \varepsilon_{ijk}$$

Missing data

See previous section on Endline data collection for attrition information. Item level missingness was very low (<1%) on covariates; a complete case analysis would have resulted in 1974 for fathers and 1964 for mothers. To have a full analytic sample, we ran a single imputation (using nearest neighbor matching) for the covariates which had missing data (we did not impute any outcomes), leading to an analytic sample of 2002 for fathers and 2000 for mothers.

Analytic approach

Our study implements similar (but not identical) interventions in two separate communities (Rohingya camp and host). Caseloads from these two communities were randomized into treatment arms separately. Thus, while we have balance on potential confounders

Mixed effects model where

y_{ijk} is the outcome variable, measured at the individual level i within geographical group j and mother volunteer k .

α is the treatment effect.

tx is the treatment assignment with a subscript k given cluster randomization.

ζ_k is the random effect at the mother volunteer level.

μ_{jk} is the random effect at the geographical group level. and ε_{ijk} is the individual level error.

We organized the outcome variables into conceptual groupings as recommended by the Institute for Educational Sciences [57] and conducted p -value adjustments using false discovery rate (FDR) methods [58], to adjust for multiple hypothesis testing. The groupings are as follows: (1) father’s mental health (PHQ and GAD), (2) fathers’ parenting and engagement with family, reported by father (ensuring child’s safety and responding to

needs, collaboration with wife about child, physical support for wife, communication with wife, father-child warmth/play, father-child harsh discipline, father-child stimulating behaviors), (3) father's parenting and engagement with family, reported by mother (same sub-scales as previous point, without the stimulating behaviors scale), (4) directly assessed child development (Bayley-4), (5) mother-reported CREDI. Father-reported social-emotional subscale of CREDI and beliefs about fathering were considered separate constructs and not included in these adjustments.

To keep the number of moderation analyses tractable, we ran exploratory and confirmatory factor analyses on all the PREI items, the father education items, and the financial worry question to reduce dimensionality. The factor analysis confirmed the presence of three distinct dimensions: father education (which includes questions on type of school, maximum schooling level, and literacy), household resources including financial worry, and housing quality. Other tested moderators included child gender, child age, baseline mother stimulating behaviors, and baseline mother health. We first tested the moderation of camp/host on all the outcomes. We ran the rest of the listed moderators on the overall sample as well as

camp and host samples separately to avoid interpreting double interactions. Continuous moderators (economic status, child age, father education) were run as continuous and the binary (child gender) and ordinal (mother health) moderators were run as categorical to avoid assuming a linear relationship. All moderation analyses went through the same *p*-value adjustment process and in the same outcome groups as the outcome models, with adjustments conducted separately for the overall sample, host sample, and camp sample. Effect sizes were calculated by dividing the regression coefficient by the standard deviation of the outcome variable.

The data and code used for these analyses are publicly available at: <https://doi.org/10.7910/DVN/6OTJQC> [59]

Results

Both mothers' and fathers' attendance in the program were monitored by the implementing organization as the number of sessions attended out of the total number of sessions conducted for that caregiver (for the fathers, this was the number of one-on-one visits and group sessions with the father volunteer (out of a possible 3 one-on-one and 1 group session each month). On average, the camp fathers ($N=397$) attended 88% of the conducted

Table 1 Sample characteristics for camp and host communities

Mother and father characteristics	Camp	Host
Mean age (SD) (in years)		
Father	32.9 (8.76)	34.6 (7.72)
Mother	28.15 (5.73)	28.06 (6.63)
Age range (years)		
Father	18–76	18–66
Mother	15–57	14–52
Attended primary school or less		
Father	66%	59%
Mother	87%	38%
Unable to read		
Father	57%	28%
Mother	70%	16%
Mother's General Health Scale (very good, good, moderate, bad, very bad)	14%, 33%, 29%, 22%, 2%	6%, 37%, 29%, 23%, 5%
Mother reported financial worry: (always, very often, sometimes, rarely, never)	9%, 16%, 35%, 25%, 16%	10%, 17%, 33%, 21%, 19%
Focal child and household characteristics (reported by mothers)		
Mean age (SD) (in months)	25.34 (4.25)	21.65 (6.23)
Age range (months)	13–40	8–41
Gender split (% male)	51.85%	50.71%
# of children under 3y/o: (one child, two children, three children)	79.75%, 20%, 0.25%	93.17%, 6.83%, N/A
# of times children fell sick/month (once, twice, three or more)	31%, 33%, 36%	26%, 27%, 46.58%

The camp sample had 6 mothers over the age of 50. This seems unlikely given that these mothers all ostensibly have children under the age of 3; in some cases, enumerators surmised that the caregiver saying she was a mother was actually a grandmother, but this could not be established definitively during data collection; nor could we establish that this was an error. We are therefore including the age as it was collected

Table 2 Outcome means and SDs by treatment group and measure alphas

Outcome	Baseline					Endline				
	Cronbach's alpha	Control		Treatment		Cronbach's Alpha	Control		Treatment	
		Mean	Std. dev	Mean	Std. dev		Mean	Std. dev	Mean	Std. dev
FR PHQ	0.81	0.98	0.54	0.99	0.56	0.77	0.93	0.47	0.91	0.45
FR GAD						0.77	0.90	0.56	0.88	0.52
FR responding to child needs	0.75	2.41	0.91	2.45	0.89	0.74	2.53	0.78	2.62	0.79
FR collaboration about child	0.81	2.13	0.89	2.16	0.88	0.81	1.95	0.80	2.05	0.78
FR physical support for wife	0.68	2.05	0.99	2.09	0.98	0.66	1.91	0.92	2.06	0.91
FR communication w/wife	0.61	3.60	0.62	3.57	0.62	0.69	3.47	0.71	3.47	0.71
FR father-child harsh discipline	0.64	0.45	0.69	0.49	0.69	0.70	0.60	0.73	0.57	0.73
FR father-child warmth/play	0.69	3.37	0.84	3.42	0.75	0.69	3.24	0.81	3.28	0.80
FR father-child stimulation	0.84	1.17	0.87	1.20	0.87	0.73	1.53	0.86	1.77	0.87
Beliefs about fathering	0.82	3.32	0.54	3.26	0.56	0.84	3.34	0.53	3.42	0.50
MR responding to child needs	0.79	2.30	0.96	2.33	0.95	0.73	2.48	0.81	2.63	0.75
MR collaboration about child	0.83	1.86	0.90	1.88	0.92	0.78	1.83	0.76	1.94	0.75
MR physical support for wife	0.70	1.83	1.01	1.89	1.02	0.63	1.87	0.95	2.02	0.88
MR communication w/wife	0.70	3.51	0.80	3.52	0.76	0.64	3.43	0.76	3.44	0.71
MR father-child harsh discipline	0.64	0.40	0.69	0.40	0.65	0.68	0.70	0.79	0.73	0.83
MR father-child warmth/play	0.74	3.27	0.92	3.30	0.86	0.69	3.21	0.88	3.30	0.79

FE father reported outcomes, MR mother-reported outcomes

sessions and the host fathers ($N=601$) attended 89% of the conducted sessions. The mean number of sessions conducted was 21.5 (min=20; max=22). As for mothers' groups, the average number of sessions conducted was 21.8 (min=19; max=23). In the camp community ($N=806$), mothers attended on average 91% of the conducted sessions, while in the host community ($N=1233$), they attended at a rate of 93%.

Camp and host family characteristics are summarized in Table 1. Means and SDs of the baseline and endline levels of all outcome variables are presented in Table 2. Below, we display the results by research question. Impact estimates are reported in Table 3 by conceptual groupings of variables. We first report the results for a pooled sample, combining host and camp communities (the columns on the left of Table 3). Following this, we report the findings of the exploratory research questions, emphasizing first the moderating effect of camp/host residence.

Table 1 shows that mean ages of caregivers are similar across camp and host communities, with fathers tending to be slightly older and with a much greater range of ages (maximum age of a father in the camps reported as 76) than mothers. Literacy and schooling were lower in the camp population (57% of fathers unable to read) than in the host population (27%). The focal child's

gender was evenly split and had mean ages of approximately 2 years old. Mother health, child health, and mother's financial concerns look similar across camp and host.

Impact on fathers' parenting and engagement with family

For the father-reported parenting and engagement outcomes, there were positive, statistically significant impacts of the treatment on responding to child needs ($\beta=0.09$, $p=0.03$, effect size (ES)=0.12, 95% confidence interval (CI) for ES=0.02–0.22), collaboration about child ($\beta=0.11$, $p=0.03$, ES=0.14, 95% CI=0.03–0.25), physical support for wife ($\beta=0.13$, $p=0.01$, ES=0.14, 95% CI=0.04–0.25), father-child stimulation ($\beta=0.23$, $p<0.001$, ES=0.23, 95% CI=0.13–0.34) and beliefs about fathering ($\beta=0.11$, $p=0.002$ (non-adjusted), ES=0.21, 95% CI=0.08–0.34). For the mother-reported outcomes, there was also a statistically significant impact of the treatment on responding to child needs ($\beta=0.13$, $p<0.001$, ES=0.17, 95% CI=0.06–0.27), and physical support for wife ($\beta=0.12$, $p=0.03$, ES=0.13, 95% CI=0.03–0.23) subscales. These results represent the weighted average of the impacts in the host and camp communities.

Table 3 Treatment effects on overall sample and camp and host sub-sample

Family of variable	Outcome	Overall effect (camp and host)			Camp treatment effect			Host treatment effect			Are camp and host significantly different?	
		β	Adj p	Effect size	β	Adj p	Effect size	β	Adj p	Effect size	β	Adj p
1	FR responding to child needs	0.07	0.03	0.12	0.09	0.02	0.17	0.05	0.58	0.08	− 0.05	1.00
1	FR collaboration about child	0.10	0.02	0.15	0.16	0.01	0.24	0.04	0.58	0.06	− 0.12	0.79
1	FR physical support for wife	0.08	0.02	0.15	0.11	0.01	0.21	0.05	0.58	0.09	− 0.06	0.92
1	FR communication w/wife	0.01	0.33	0.03	0.02	0.09	0.09	− 0.01	0.90	− 0.03	− 0.03	0.92
1	FR father-child harsh discipline	− 0.03	0.22	− 0.05	− 0.15	0.01	− 0.23	0.09	0.46	0.13	0.24	0.01
1	FR father-child warmth/play	0.04	0.12	0.07	0.05	0.09	0.09	0.0	0.58	0.06	− 0.02	1.00
1	FR father-child stimulation	0.20	0.001	0.23	0.23	0.01	0.23	0.18	0.04	0.21	− 0.05	1.00
–	Beliefs about fathering	0.11	0.002	0.21	0.18	0.00	0.36	0.02	0.71	0.04	− 0.16	0.02
2	MR Responding to child needs	0.09	0.02	0.15	0.04	1.00	0.08	0.14	0.01	0.22	0.09	0.78
2	MR Collaboration about child	0.07	0.06	0.11	0.04	1.00	0.06	0.11	0.04	0.16	0.07	0.78
2	MR Physical support for wife	0.07	0.02	0.16	0.05	1.00	0.11	0.10	0.01	0.21	0.06	0.78
2	MR Communication w/wife	0.05	0.36	0.02	− 0.01	1.00	− 0.031	0.02	0.19	0.07	0.03	0.78
2	MR Father-child harsh discipline	0.01	0.36	0.01	− 0.004	1.00	− 0.01	0.02	0.28	0.03	0.03	0.86
2	MR Father-child warmth/play	0.07	0.05	0.12	0.01	1.00	0.03	0.10	0.01	0.19	0.08	0.78
3	FR PHQ	− 0.03	0.40	− 0.07	− 0.06	0.09	− 0.13	− 0.01	1.00	− 0.02	0.05	0.38
3	FR GAD	− 0.03	0.40	− 0.06	− 0.08	0.09	− 0.14	0.001	1.00	0.02	0.09	0.38
4	Bayley cognitive	0.07	1.00	0.03	0.14	0.50	0.06	0.02	1.00	0.01	− 0.11	0.37
4	Bayley receptive language	0.002	1.00	0.00	− 0.08	0.61	− 0.03	0.07	1.00	0.03	0.15	0.37
4	Bayley expressive language	− 0.06	1.00	− 0.02	− 0.32	0.26	− 0.13	0.15	1.00	0.06	0.48	0.13
4	Bayley language	− 0.05	1.00	− 0.01	− 0.40	0.39	− 0.09	0.22	1.00	0.05	0.63	0.23
5	MR CREDI cognitive	− 0.04	1.00	− 0.03	− 0.04	1.00	− 0.04	− 0.04	1.00	− 0.03	0.00	1.00
5	MR CREDI language	− 0.13	0.10	− 0.08	− 0.19	0.04	− 0.13	− 0.06	1.00	− 0.03	0.14	1.00
5	MR CREDI motor	− 0.09	1.00	− 0.06	− 0.09	1.00	− 0.07	− 0.09	1.00	− 0.06	0.01	1.00
5	MR CREDI SEL	− 0.03	1.00	− 0.02	0.00	1.00	0.00	− 0.05	1.00	− 0.03	− 0.05	1.00
–	FR CREDI SEL	0.14	0.01	0.12	0.16	0.04	0.18	0.11	0.14	0.09	− 0.04	0.79

Multiple hypothesis adjustment took place within groupings of variables: 1) Father Reported Parenting & Engagement with Family 2) Mother Reported Father Parenting & Engagement with Family; 3) Father-reported mental health 4) Directly Assessed Child Development (Bayley-4) 5) Mother-reported child development (CREDI)

Beliefs about Fathering and Father-reported CREDI SEL were not included in these groupings; these p -values are unadjusted

Table 4 Moderation effects remaining significant after FDR adjustment

Moderator	Outcome	Community	Interaction β	Adjusted p
Child age	Collaboration w/wife	Host	0.02	0.00
Child gender	MR CREDI Cognitive	Host	0.25	0.03
	MR CREDI Language	Host	0.26	0.03
	MR CREDI Motor	Host	0.29	0.03
	MR CREDI SEL	Host	0.26	0.03
MR child stimulation at baseline	MR CREDI Motor	Camp	− 0.30	0.04
	FR CREDI SEL	Camp	− 0.19	0.01
Mother health	FR CREDI SEL	Camp	− 0.18	0.00

FR father-reported, MR mother-reported, SEL social-emotional learning

Impact on fathers' mental health (depression and anxiety)

We found no statistically significant impacts on father depressive or anxiety symptoms, measured through father-reported PHQ-8 and GAD across the pooled camp and host sample, though the coefficients on PHQ-8 and GAD in the camp approach significance (PHQ: $\beta = -0.06$, $p = 0.08$, $ES = 0.14$, 95% $CI = -0.29$ – 0.01 ; GAD: $\beta = -0.08$, $p = 0.08$, $ES = 0.15$, 95% $CI = -0.30$ to 0.00).

Impact on child development

We found no statistically significant impacts on cognitive or language domains of direct assessment (Bayley-4) or mother-reported child development (CREDI). However, we did find a positive impact on father-reported social-emotional development (sub-scale of CREDI; $\beta = 0.14$, $p = 0.01$, $ES = 0.12$, 95% $CI = 0.03$ – 0.21).

Moderation of effects by family characteristics

All moderation effects were adjusted for multiple hypothesis testing; only the results that remain statistically significant after this are discussed here. The results of moderation analyses for camp and host community residence (i.e., analyses run on restricted camp and host subsamples) are shown in Table 3, along with the impacts on the overall sample described above for comparison purposes. The last column of the table indicates whether the camp and host communities are significantly different from each other. The magnitude and p -value of other moderation effects (by child age, child gender, mother-reported stimulating behaviors with child at baseline, and mother health) that remain significant after the FDR adjustment are reported in Table 4, along with the community where the relationship was present.

Moderation of impacts by camp/host community residence

The interaction term of the camp/host indicator and the treatment effect was not statistically significant for any of the mother-reported outcomes or the Bayley-4 scales, and was only significant for two of the father-reported outcomes: harsh discipline ($\beta = 0.30$, $p < 0.001$; note that mean levels of harsh discipline in the camp community are significantly higher) and beliefs about fathering ($\beta = 0.16$, $p = 0.02$) (see last two columns of Table 3). The program impact on reduced father-reported harsh discipline was stronger in the camp community ($\beta = -0.20$, $p < 0.001$, $ES = 0.27$, 95% $CI = -0.43$ – -0.12) than the host ($\beta = 0.09$, $p = 0.43$, $ES = 0.13$, 95% $CI = -0.03$ – 0.29). Similarly, the impact of the program on beliefs about fathering was stronger in the camp community ($\beta = 0.18$, $p < 0.001$, $ES = 0.35$, 95% $CI = 0.18$ – 0.53) than in the host community ($\beta = 0.02$, $p = 0.72$, $ES = 0.04$, 95% $CI = -0.16$ – 0.23). Although we cannot reject the null

hypothesis that the treatment effects are similar across camp and host community for all other outcomes—likely due to being underpowered to test moderation (our study was only sufficiently powered to detect main effects, we therefore may not have sufficient statistical power to detect all small possible subgroup effects)—we note that the magnitudes of the coefficients for father-reported outcomes are almost invariably greater in the camp sample than the host sample.

Moderation of child outcome impacts

Statistically significant moderation effects other than camp/host residence that withstand the multiple testing adjustments are presented in Table 4. While we find no evidence of any moderation of direct child assessment impacts (Bayley), mother-reported stimulating behaviors and mother health moderate the impact on father-reported child SEL (CREDI) in the camp community, while child gender and mother-reported stimulating behaviors moderate impacts on mother-reported CREDI outcomes in the host and camp community, respectively. Child gender (1 = female) is a significant moderator of all mother-reported CREDI outcomes in the host community ($\beta_s = 0.25$ – 0.29 , $p_s = 0.03$), whereby the treatment effects on these items were negative for boys ($\beta_s = -0.17$ – -0.23 , $ES_s = -0.11$ – -0.15), and positive but weak for girls ($\beta_s = 0.06$ – 0.08 , $ES_s = 0.04$ – 0.07). Baseline values of mother stimulating behaviors in the household moderate the impact on the motor domain of the CREDI (mother-reported) in the camp community, whereby larger positive impacts were found among mothers from households who reported lower baseline levels of stimulation. Mother-reported stimulating behaviors and mother health also moderate the impacts on the father-reported SEL CREDI domain, whereby larger positive impacts were found among fathers from households who reported lower baseline levels of mother stimulation and health.

Moderation of father engagement and mental health impacts

We find no evidence of moderation of impacts on father engagement with the child and family or father mental health outcomes in either community that survive multiple testing adjustments except for child age moderating the mother-reported “collaboration with wife about child” outcome in the host community, whereby larger positive impacts were found among mothers of older children than younger ones.

Resources delivered through program and cost of delivery

Using the ingredients method [60], we examined the resources required for delivering both the mother and

father 6-month program, with the goal of estimating the cost of the program in total as well as the relative cost of the father engagement aspect of the program relative to the control condition. This second, relative estimate captures the economic costs that correspond to the estimated effectiveness of the father engagement program. The total economic cost of delivering the 6-month program to the 998 families in treated households was approximately 553,740 USD (as of 2023; 60,911,400 BDT), which averages to 550 USD (60,500 BDT) per household. About 27% of costs were borne by BRAC-IED in co-creating the father curriculum; 71% of costs were borne by BRAC-IED for program delivery; and 2% of the costs reflect the time caregivers contributed to participate in the program. Future implementation in the Rohingya context would bear lower costs for program delivery since the content is already developed and there is a structure to support the program delivery model. The added cost of the 6-month father engagement component relative to the mother program was 347,460 USD (38,220,600 BDT), or about 350 USD (38,500 BDT) per household.

Discussion

To our knowledge, this study is the only randomized controlled trial to date of a program focusing on father engagement with young children in a LMIC humanitarian context. The role of fathers in children's lives is well established in the literature [1, 5–15] but interventions involving them are comparatively lagging, and especially so in humanitarian contexts. Not a single published study has focused on father engagement or parenting among fathers of young children among the Rohingya, and the program presented here is the first to focus on fathers of 0–3-year-olds for this population. This study is therefore filling a conspicuous gap and provides valuable information, through both the findings and methods, on the potential of father engagement programs for early childhood development. It also provides valuable insights for practitioners working to design and provide services in the Rohingya camps and surrounding communities, who are faced with limited resources with which to improve outcomes for both children and caregivers in a protracted crisis of the largest forcibly displaced population in the world.

The program itself has several features that positively distinguish it from many existing programs intended to involve fathers, and that follow recommendations from Jeong et al. [25]. for the future of fathering interventions in LMICs. These include a curriculum designed deliberately for fathers, flexible scheduling, and a mixed home visit/group component. In addition, our study included

both Rohingya and host community samples, providing evidence from these two contrasting populations.

The program was well received by families, with fathers attending, on average, 89% of sessions in both host and camp communities (Iqbal, Y., Yoshikawa, H., Michael, D., Bolisetty, S., Schwartz, K., Yesmin, S., Mian, A., Nyeu, M., Sunny, S., Ramachandran, A., Hilgendorf, D., Tiwari, P., Barrios, S.F., Chisim, K., Mur Salin, M. S., Hossain, N., Zaman, S.S., Ali, A. R. M. M., Chowdhury, S., Wuermli, A. Impact Evaluation of a Father Engagement Model in the Home Visiting Humanitarian Play Lab (HPL) Program in Rohingya Camps and Host Community in Cox's Bazar, Bangladesh. 2024, unpublished). It had small positive impacts on fathers' parenting and family engagement (both father and mother-reported) and beliefs about fathering (father-reported). Specifically, the program increased fathers' reports of responding to child needs, collaboration with their wives about the child, and physically supporting the wife (note that during the 0–2 period physical care of the infant is developmentally prominent). In addition, father-reported stimulating behaviors with the child increased, as did levels of beliefs in the importance of play, fathering, and father engagement. Among mother-reported outcomes, the program increased fathers' responding to child needs, physical support of the mother, and fathers' warmth and play with the child. Although some parenting interventions have examined couples-related outcomes such as emotional support (Iqbal, Y., Yoshikawa, H., Michael, D., Bolisetty, S., Schwartz, K., Yesmin, S., Mian, A., Nyeu, M., Sunny, S., Ramachandran, A., Hilgendorf, D., Tiwari, P., Barrios, S.F., Chisim, K., Mur Salin, M. S., Hossain, N., Zaman, S.S., Ali, A. R. M. M., Chowdhury, S., Wuermli, A. Impact Evaluation of a Father Engagement Model in the Home Visiting Humanitarian Play Lab (HPL) Program in Rohingya Camps and Host Community in Cox's Bazar, Bangladesh. 2024, unpublished) or intimate partner violence (e.g., Ashburn et al. [26]), this is one of the few to examine collaboration with the mother specifically about the child.

The impact of the program on mothers' reports of her husband's behavior is of particular interest, given that one of the reasons the fathers' engagement component was developed by the implementing organization was that mothers repeatedly communicated to staff that an intervention should be provided for their husbands as well as for them (some staff also noted that several fathers also mentioned that they saw their wives receiving an intervention and wanted one for themselves). Notably, few studies on father-related interventions use both father- and mother- reports of father parenting. We found small, positive impacts specifically on mother reports of fathers' responding to child needs and

physical support for his wife (i.e., the mother). In addition to these findings, mothers in our qualitative data (Iqbal, Y., Sunny, S., Bolisetty, S. Caregiver and provider experiences of a father engagement program in the Rohingya camps and host community in Cox's Bazar, Bangladesh: A Qualitative Study, unpublished) whose husbands participated in the father program described less quarreling and greater peace in the household. Although these outcomes are sufficient in and of themselves as beneficial for the family, they also hold promise for potential longer-term benefits for children: we know from the literature that mother-reported father involvement has been found to be predictive of positive outcomes for children in LMICs, even after controlling for effects of mother involvement [23, 61].

The positive impact on fathers' beliefs about play, fathering and father engagement is also novel. Few studies have examined parent beliefs about play in an LMIC context (for an exception, see example in China [20]; one program in South Africa focused on father engagement and gender equality led to a shift in fathers' beliefs about fathering [62]). To our knowledge this is one of the first evaluations showing a positive impact on fathers' beliefs about play, which have been linked to positive child behavior in studies in high-income countries [63].

There was however no main effect on fathers' depressive and anxiety symptoms; given the serious mental health concerns for fathers in displaced populations [64], mental health was an important target of the program. It is possible that the most important effect of the program on fathers' emotions and emotional regulation was not on depression or anxiety but on management of anger, which was a central aspect of the curriculum. The one quantitative measure from our surveys that acts as a proxy for anger responses is the scale of harsh discipline (which showed significant reductions among the Rohingya). A scale on anger management, or expression of anger, would have been appropriate to include. However, we did not understand the extent to which this may have been a key outcome until we engaged in post-program discussions with program participants and father volunteers. One potential mechanism that fathers in both the camp and host community mentioned was more frequent sharing of their feelings with other fathers, their wives, or personnel associated with the intervention as a method of self-care. However, with the program only being 6 months long, such new mechanisms might take much longer before they bring about measurable change in participants' depression or anxiety levels. This may have implications for content development and delivery in the future: for instance, an iteration that meets less frequently but goes on for longer may be more successful at meeting the program's goals for certain outcomes.

The program also did not have a main effect on directly assessed cognitive or language development or mother-reported developmental outcomes; it did, however, have a main effect on father reported social-emotional development (a subscale of the CREDI). This finding might be driven by the father's new warm engagement with the child, and the father's perception of the child's behavioral response to this. However, it is also possible that fathers in the treatment condition overreport developmental milestones, given that the program may have sensitized them to observe their children's development.

Overall, for father-reported outcomes, the magnitudes of the estimated impacts of the program on fathers were stronger for the Rohingya than for the host community, particularly for increased responding to child needs, collaboration about child, and physical support for the wife, reduction in harsh discipline, and improvements in beliefs about fathering. There are three possible explanations for this. First, the intervention was originally intended for the camp community; it was implemented in the host community rapidly as the evaluation component was affected by changing research regulations. The visual flip book, for instance, was not developed for the host community but may have been useful in reinforcing fathering messages for the Rohingya. Second, there may have been less presence of implementer staff support in the host community, a potential by-product of the fact that the host community is much more spread out geographically than the camps, and that the camps have several other programs, such as those directed towards older children, being run by the same implementers. Finally, due to restrictions on employment options, Rohingya fathers were generally more available than host fathers. In combination, these factors might explain the program having a stronger impact on fathers in the camp community.

Among other moderators, father education and economic status do not moderate the impact of the program on any of our outcomes. We find that the baseline level of mother-reported stimulating behaviors with the child moderates impacts on both mother-reported child motor development and father-reported social-emotional development of the child in the camp community. Larger positive effects in children's motor development were found for mothers who reported lower baselines of child stimulation than those who have higher baseline levels. This suggests that the program helps children in households that have lower baseline stimulation levels catch up with those who have higher baseline levels. Similarly, in households where the mother reported lower baseline levels of child stimulation, larger positive effects were found on father-reported social-emotional learning due

to the program. These moderation findings represent some corroboration of mothers' and fathers' report of child development and are in line with general findings in the literature that positive effects of parent-directed programs are often stronger for more disadvantaged families [31].

Child age does not moderate the treatment impact on any of our outcomes except for mother-reported "collaboration with wife about the child". The magnitude of the interaction suggests that the treatment effect, on average, is positive for children who are older, but can be negative for younger children.

Finally, gender moderated the program's impact on mother-reported CREDI scores, such that the program had positive effects on all domains of girls' but not boys' development as reported by mothers in the host community. Meta-analyses of early childhood program evaluations do not find a clear pattern of gender moderation of effects on children, in either LMICs or high-income countries [31, 63], but our early piloting work suggests that aspirations for children may differ based on gender in our contexts from early ages [65], with expectations that children adopt traditional roles driven by religious beliefs.

The moderation effects observed for CREDI outcomes but not directly assessed outcomes through the Bayley likely reflect these measures' differing methodologies: caregiver-reported measures like the CREDI are potentially capturing development as perceived within everyday contexts, making them more sensitive to household or demographic characteristics.

There are several limitations of this study. At baseline recruitment, we selected a convenience sample of fathers (those who were partnered with program-participating mothers and who were available for research participation; and in the host community, those who were resident with the family for at least eight upcoming months). We are limited in our ability to generalize to the Rohingya camp and host communities as a whole. We consider a great *strength* of this research to be that it consists of families from both the forcibly displaced and host populations, which gives us considerable insights into how to consider early childhood interventions in humanitarian settings. However, this came with certain limitations. Due to time restrictions, we conducted in-depth piloting of our measures in the camp community, not in the host community, meaning that the measures may have been better suited to the camp community families. Baseline psychometrics, which guided how we conceptualized our measures for endline, were conducted on host and camp samples combined, but after mothers in the camps had received approximately 3 months of time in the program (host mothers had not

started receiving the intervention yet). Another limitation is a potential threat to internal validity: enumerators were likely often aware of the treatment status of the households they were visiting. Although the survey was designed to keep enumerators and supervisors blind to condition, treatment condition may have been revealed because there was no way to conduct endline data collection without father volunteers helping to locate the household. Finally, we did not have baseline measures of either the CREDI or the Bayley-4, so we could not control for them in the analyses, which would have likely provided more precise point estimates of impact for these outcomes. However, all our balance checks on a large range of other covariates suggested successful randomization in both communities.

Conclusions

The present study, evaluating an early childhood father engagement intervention delivered in person to fathers in the Rohingya camps and surrounding host communities in Cox's Bazar, Bangladesh, adds valuable evidence to our understanding of interventions for fathers in global contexts. As a protracted crisis with a forcibly displaced population, the Rohingya and host communities in Cox's Bazar represented an especially challenging context in which to deliver such an intervention and study its impacts. Given this difficulty, however, we can state that even in difficult contexts, attention to certain characteristics can yield positive results; specifically, a curriculum geared towards fathers, delivery by male volunteers from the cultural and linguistic group, flexible scheduling and both dyadic and group modalities. The study also provides evidence on how interventions impact families from both forcibly displaced and host populations. Our findings demonstrate the potential of father-focused programs as powerful ECD caregiver interventions, which continue to largely focus on mothers. In humanitarian contexts such as the one in our study, comprising tightly knit families and communities living near one another, focusing on fathers as a target for intervention has intriguing possibilities for the families and communities at large.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s44263-026-00266-x>.

Supplementary Material 1. Research Protocol, Impact Evaluation of a Father Engagement Model in Rohingya Camps and Host Community in Cox's Bazar, Bangladesh.

Supplementary Material 2. CONSORT Checklist for "Effects of Father Engagement Impact Evaluation".

Acknowledgements

This work was produced using resources and staff expertise from NYU Global TIES for Children in collaboration with Sesame Workshop, BRAC, icddr, and ARCED Foundation, as part of the Play to Learn program generously funded by the LEGO Foundation.

Authors' contributions

Y.I., H.Y., and S.B. contributed to the design and implementation of the study, project planning, data collection, data interpretation and manuscript writing. S.Y. and K.C. contributed to design, implementation, and project planning. A.U.M., A.R.M.M.A., and S.S.C. contributed to design, implementation, project planning, and data collection. D.M. conducted the analyses and contributed to interpretation and manuscript writing. M.T.N. and K.S. conducted the analyses and contributed to data interpretation. S.L., A.B.B., A.A., & J.R.B. contributed to the cost analyses for the paper. All authors read and approved the final manuscript.

Funding

This evaluation was conducted as part of the Play to Learn program funded by the LEGO Foundation.

Data availability

The datasets generated and analysed during the current study are publicly available at the Harvard Dataverse at: <https://doi.org/10.7910/DVN/6OTJQC> [59]

Declarations

Ethics approval and consent to participate

All research protocols used in this study were approved by the Institutional Review Board (IRB) of the Institute of Health Economics, Dhaka University (an IRB based in Bangladesh; reference number IHE/IRB/DU/49/2022/Final). In addition, this research was approved by the Office of the Refugee Relief and Repatriation Commissioner (RRRC), a necessary approval for all research and service activities conducted by NGOs or international organizations in the Rohingya camps (more information at <https://rrrc.gov.bd/>). This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki, which ensures respect for all individuals, safeguards their health and rights, and promotes ethical standards in research. Participant confidentiality was strictly maintained throughout the research, and all data were anonymized to ensure privacy. All participants were read a consent statement, and verbal consent was obtained (given most participants' inability to read/write).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Division of Social Sciences, Duke Kunshan University, Kunshan, China. ²Steinhardt School of Culture, Education, and Human Development, New York University, New York, NY, USA. ³Teachers College, Columbia University, New York, NY, USA. ⁴BRAC Institute of Educational Development, Dhaka, Bangladesh. ⁵Vancouver, Canada. ⁶The Gervitz Graduate School of Education, University of Santa Barbara, Santa Barbara, CA, USA. ⁷Monroe-Meyer Institute, University of Nebraska Medical Center, Omaha, NE, USA. ⁸Center for Benefit-Cost Studies of Education, University of Pennsylvania, Philadelphia, PA, USA. ⁹ARCED (Aureolin Research, Consultancy, and Expertise Development) Foundation, Dhaka, Bangladesh. ¹⁰Center for Educational Data Science and Innovation, University of Maryland, College Park, MD, USA.

Received: 16 October 2025 Accepted: 30 March 2026
Published online: 09 April 2026

References

1. Tamis-LeMonda CS, Kahana-Kalman R, Yoshikawa H. Father involvement in immigrant and ethnically diverse families from the prenatal period to the second year: prediction and mediating mechanisms. *Sex Roles*. 2009;60(7–8):496–509. <https://doi.org/10.1007/s11199-009-9593-9>.
2. Barker B, Iles JE, Ramchandani PG. Fathers, fathering and child psychopathology. *Curr Opin Psychol*. 2017 [cited 2023 Jul 10];15:87–92. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2352250X16301336>
3. Henrich J, Heine SJ, Norenzayan A. The weirdest people in the world? *Behav Brain Sci*. 2010 [cited 2023 Aug 2];33(2–3):61–83. Available from: https://www.cambridge.org/core/product/identifier/S0140525X0999152X/type/journal_article
4. UNHCR. unhcr.org. 2023. Refugee Data Finder. Available from: <https://www.unhcr.org/refugee-statistics/#:~:text=Low%2D%20and%20middle%2Dincome%20countries%20host%2076%20per%20cent%20of,per%20cent%20of%20the%20total.&text=Data%20on%20some%204.4%20million,was%20reported%20at%20end%2D2022>.
5. Grossmann K, Grossmann KE, Fremmer-Bombik E, Kindler H, Scheuerer-Engelsch H, Zimmermann AP. The uniqueness of the child-father attachment relationship: fathers' sensitive and challenging play as a pivotal variable in a 16-year longitudinal study. *Soc Develop*. 2002 [cited 2023 Jul 10];11(3):301–37. Available from: <https://onlinelibrary.wiley.com/doi/10.1111/1467-9507.00202>
6. Rollé L, Gullotta G, Trombetta T, Curti L, Gerino E, Brustia P, et al. Father involvement and cognitive development in early and middle childhood: a systematic review. *Front Psychol*. 2019 Oct 25 [cited 2025 Aug 2];10:2405. Available from: <https://www.frontiersin.org/article/https://doi.org/10.3389/fpsyg.2019.02405/full>
7. Cano T, Perales F, Baxter J. A matter of time: father involvement and child cognitive outcomes. *J Marriage Fam*. 2019 [cited 2023 Jul 10];81(1):164–84. Available from: <https://onlinelibrary.wiley.com/doi/10.1111/jomf.12532>
8. Sethna V, Perry E, Domoney J, Iles J, Psychogiou L, Rowbotham NEL, et al. Father-child interactions at 3 months and 24 months: contributions to children's cognitive development AT 24 MONTHS: Fathers' and Children's Cognitive Development. *Infant Ment Health J*. 2017 [cited 2023 Jul 10];38(3):378–90. Available from: <https://onlinelibrary.wiley.com/doi/10.1002/imhj.21642>
9. Frongillo EA, Kulkarni S, Basnet S, De Castro F. Family care behaviors and early childhood development in low- and middle-income countries. *J Child Fam Stud*. 2017;26(11):3036–44. <https://doi.org/10.1007/s10826-017-0816-3>.
10. McWayne C, Downer JT, Campos R, Harris RD. Father involvement during early childhood and its association with children's early learning: a meta-analysis. *Early Educ Develop*. 2013 [cited 2023 Jul 10];24(6):898–922. Available from: <http://www.tandfonline.com/doi/abs/10.1080/10409289.2013.746932>
11. Leidy MS, Schofield TJ, Parke RD. Fathers' contributions to children's social development. In: handbook of father involvement: multidisciplinary perspectives. Second. New York: Routledge; 2012. p. 151–67.
12. Edin K, Nelson TJ. Doing the best I can: fatherhood in the inner city. first paperback printing 2014. University of California Press; 2014.
13. Roopnarine JL, Dede Yildirim E. Fathers' cognitive engagement and preschoolers' literacy skills in three ethnic groups in Trinidad and Tobago. *J Black Psychol*. 2019 [cited 2023 Jul 10];45(6–7):494–517. Available from: <http://journals.sagepub.com/doi/10.1177/0095798419887067>
14. Maselko J, Hagaman AK, Bates LM, Bhalotra S, Biroli P, Gallis JA, et al. Father involvement in the first year of life: associations with maternal mental health and child development outcomes in rural Pakistan. *Social Science & Medicine*. 2019 [cited 2025 Aug 2];237:112421. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0277953619304137>
15. Jeong J, McCoy DC, Yousafzai AK, Salhi C, Fink G. Paternal stimulation and early child development in low- and middle-income countries. *Pediatrics*. 2016 Oct 1 [cited 2023 Jul 10];138(4):e20161357. Available from: <https://publications.aap.org/pediatrics/article/138/4/e20161357/52343/Paternal-Stimulation-and-Early-Child-Development>

16. Amodia-Bidakowska A, Laverty C, Ramchandani PG. Father-child play: a systematic review of its frequency, characteristics and potential impact on children's development. *Develop Rev.* 2020 [cited 2025 Aug 2];57:100924. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0273229720300307>
17. St.George JM, Wroe JK, Cashin ME. The concept and measurement of fathers' stimulating play: a review. *Attach Human Develop.* 2018 Nov 2 [cited 2025 Aug 2];20(6):634–58. Available from: <https://www.tandfonline.com/doi/full/10.1080/14616734.2018.1465106>
18. Sarkadi A, Kristiansson R, Oberklaid F, Bremberg S. Fathers' involvement and children's developmental outcomes: a systematic review of longitudinal studies. *Acta Paediatr.* 2008;97(2):153–8. <https://doi.org/10.1111/j.1651-2227.2007.00572.x>.
19. Luoto JE, Lopez Garcia I, Aboud FE, Singla DR, Fernald LCH, Pitchik HO, et al. Group-based parenting interventions to promote child development in rural Kenya: a multi-arm, cluster-randomised community effectiveness trial. *The Lancet Global Health.* 2021 [cited 2023 Jul 10];9(3):e309–19. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2214109X20304691>
20. Lin X, Li H. Parents' play beliefs and engagement in young children's play at home. In: *Working with Parents and Families in Early Childhood Education*. Routledge; 2019. p. 5–20.
21. De Mendonça JS, De Felipe RP, David VF. Socialization goals, familism, and interactional synchrony in low-income Brazilian mothers and fathers. *ADV RES SCI.* 2021;2(2):99–108. <https://doi.org/10.1007/s42844-021-00040-y>.
22. Ramchandani P, Stein A, Evans J, O'Connor TG. Paternal depression in the postnatal period and child development: a prospective population study. *The Lancet.* 2005 [cited 2025 Aug 2];365(9478):2201–5. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0140673605667785>
23. Jeong J, Li Z. The association between fathers' depression and children's socioemotional development: evidence from a longitudinal household survey in China. *Prev Sci.* 2020;21(5):672–80. <https://doi.org/10.1007/s11121-020-01117-3>.
24. Jeong J, Sullivan EF, McCann JK. Effectiveness of father-inclusive interventions on maternal, paternal, couples, and early child outcomes in low- and middle-income countries: a systematic review. *Soc Sci Med.* 2023;328: 115971.
25. Jeong J, Sullivan EF, McCann JK, McCoy DC, Yousafzai AK. Implementation characteristics of father-inclusive interventions in low- and middle-income countries: a systematic review. *Ann NY Acad Sci.* 2023;1520(1):34–52.
26. Ashburn K, Kerner B, Ojamege D, Lundgren R. Evaluation of the Responsible, Engaged, and Loving (REAL) fathers initiative on physical child punishment and intimate partner violence in Northern Uganda. *Prev Sci.* 2017;18(7):854–64. <https://doi.org/10.1007/s11121-016-0713-9>.
27. Doyle K, Levitov RG, Barker G, Bastian GG, Bingenheimer JB, Kazimbaya S, et al. Gender-transformative Bandedeho couples' intervention to promote male engagement in reproductive and maternal health and violence prevention in Rwanda: findings from a randomized controlled trial. *PLoS ONE.* 2018;13(4):e0192756. <https://doi.org/10.1371/journal.pone.0192756>.
28. Abimpaye M, Dusabe C, Pisani L, Nzabonimpa JP, Ashford R. Improving parenting practices and development for young children in Rwanda: results from a randomized control trial. *Int J Behav Dev.* 2020;44(3):205–15. <https://doi.org/10.1177/0165025419861173>.
29. Betancourt TS, Jensen SKG, Barnhart DA, Brennan RT, Murray SM, Yousafzai AK. Promoting parent-child relationships and preventing violence via home-visiting: a pre-post cluster randomised trial among Rwandan families linked to social protection programmes. *BMC Public Health.* 2020;20(1):621. <https://doi.org/10.1186/s12889-020-08693-7>.
30. Jeong J, Ahun MN, Gunaratna NS, Ambikapathi R, Mapendo F, Galvin L, et al. Effects of engaging fathers and bundling parenting and nutrition interventions on early child development and maternal and paternal parenting in Mara, Tanzania: a factorial cluster-randomized controlled trial. *Child Psychol Psychiatry.* 2023 Oct 6 [cited 2024 Feb 19];jcnp.13897. Available from: <https://acamh.onlinelibrary.wiley.com/doi/10.1111/jcpp.13897>
31. Jeong J, Franchett EE, Ramos De Oliveira CV, Rehmani K, Yousafzai AK. Parenting interventions to promote early child development in the first three years of life: a global systematic review and meta-analysis. *Persson* LÅ, editor. *PLoS Med.* 2021 May 10 [cited 2023 Jul 12];18(5):e1003602. Available from: <https://dx.plos.org/10.1371/journal.pmed.1003602>
32. Susin LRO, Giugliani ERJ. Inclusion of fathers in an intervention to promote breastfeeding: impact on breastfeeding rates. *J Hum Lact.* 2008;24(4):386–92 quiz 451–3.
33. Daniele MA, Ganaba R, Sarrassat S, Cousens S, Rossier C, Drabo S, et al. Involving male partners in maternity care in Burkina Faso: a randomized controlled trial. *Bull World Health Organ.* 2018 Jul 1 [cited 2023 Oct 15];96(7):450–61. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6022615/>
34. Dinga LA, Kiage BN, Kyallo FM. Effect of paternal education about complementary feeding of infants in Kisumu county, Kenya. *African J Food, Agricul, Nutr Develop.* 2018 [cited 2023 Oct 16];18(3):13702–16. Available from: <https://www.ajol.info/index.php/ajfand/article/view/181602>
35. Mullany BC, Becker S, Hindin MJ. The impact of including husbands in antenatal health education services on maternal health practices in urban Nepal: results from a randomized controlled trial. *Health Educ Res.* 2007;22(2):166–76.
36. Hong YJ, Yi SH. Actor and partner effects of mothers' and fathers' parenting stress and family cohesion and flexibility on warm parenting behavior. *Korean J Child Stud.* 2017 Apr 30 [cited 2025 Aug 2];38(2):37–50. Available from: <http://childstudies.org/journal/view.php?doi/10.5723/kjcs.2017.38.2.37>
37. Lachman J, Wamoyi J, Spreckelsen T, Wight D, Maganga J, Gardner F. Combining parenting and economic strengthening programmes to reduce violence against children: a cluster randomised controlled trial with predominantly male caregivers in rural Tanzania. *BMJ Global Health.* 2020 Jul 1 [cited 2023 Oct 13];5(7):e002349. Available from: <https://gh.bmj.com/content/5/7/e002349>
38. Skar AMS, Sherr L, Macedo A, Tetzchner SV, Fostervold KI. Evaluation of parenting interventions to prevent violence against children in Colombia: a randomized controlled trial. *J Interpers Violence.* 2017;36(1–2):NP1098–126. <https://doi.org/10.1177/0886260517736881>.
39. Su M, Ouyang YQ. Father's role in breastfeeding promotion: lessons from a quasi-experimental trial in China. *Breastfeed Med.* 2016 Apr [cited 2023 Oct 16];11(3):144–9. Available from: <https://www.liebertpub.com/doi/10.1089/bfm.2015.0144>
40. Kiragu E, Li Rosi A, Morris T. States of denial: a review of UNHCR's response to the protracted situation of stateless Rohingya refugees in Bangladesh. United Nations High Commissioner for Refugees Policy Development and Evaluation Service (PDES); 2011 Dec. Available from: <https://www.unhcr.org/media/states-denial-review-unhcrs-response-protracted-situation-stateless-rohingya-refugees>
41. Save the Children. [cited 2023 Oct 15]. The Rohingya crisis: explained. Available from: <https://www.savethechildren.org/us/what-we-do/emergency-response/rohingya-crisis>
42. UNHCR. Country - Bangladesh. [cited 2023 Sep 20]. Available from: <https://data.unhcr.org/en/country/bgd>
43. Sultana R, Parry AA, Hossain MR, Aktar B, Rashid SF. "We are invisible to them"—Identifying the most vulnerable groups in humanitarian crises during the COVID-19 pandemic: The case of Rohingyas and the Host communities of Cox's Bazar. *PLOS Global Public Health.* 2023 Jun 8 [cited 2023 Sep 14];3(6):e0000451. Available from: <https://journals.plos.org/globalpublichealth/article?id=10.1371/journal.pgph.0000451>
44. UNCHR. Emergency Appeal: Rohingya Emergency. 2025 [cited 2025 Aug 2]. Available from: <https://www.unhcr.org/emergencies/rohingya-emergency>
45. Hopewell S, Chan AW, Collins GS, Hróbjartsson A, Moher D, Schulz KF, Tunn R, Aggarwal R, Berkswits M, Berlin JA, Bhandari N, Butcher NJ, Campbell MK, Chidebe RCW, Elbourne D, Farmer A, Fergusson DA, Golub RM, Goodman SN, Hoffmann TC, Ioannidis JPA, Kahan BC, Knowles RL, Lamb SE, Lewis S, Loder E, Offringa M, Ravaud P, Richards DP, Rockhold FW, Schriger DL, Siegfried NL, Stanisiewska S, Taylor RS, Thabane L, Torgerson D, Vohra S, White IR, Boutron I. CONSORT 2025 statement: updated guideline for reporting randomised trials.
46. Iqbal, Y., Bolisetty, S., Ramachandran, A., Friere, S., Goodfriend, E., Sunny, S., Begum, T. Developing and implementing a measure of quality of home visit interactions for fathers: Rohingya camps and host communities in Cox's Bazar, Bangladesh [Internet]. Figshare; 2024 [cited 2026 Jan 25]. Available from: https://figshare.com/articles/preprint/Developing_and_Implementing_a_Measure_of_Quality_of_Home_Visit_Interactions_

- for_Fathers_the_Rohingya_Camps_and_Host_Communities_in_Cox_s_Bazar_Bangladesh/26097613/1?file=47234152
47. Iqbal Y, Bolisetty S, Alif A, Tiwari P, Hilgendorf D, et al. Testing measures of refugee camp environment, caregiver mental health, and child social-emotional development among the Rohingya in Cox's Bazar [Internet]. figshare; 2022 [cited 2025 Oct 14]. Available from: https://figshare.com/articles/preprint/Testing_measures_of_refugee_camp_environment_caregiver_mental_health_and_child_social-emotional_development_among_the_Rohingya_in_Cox_s_Bazar/20292117/1
48. Goodfriend E, Wuermli A, Hiott C, Trang K, Iqbal Y, & Castelyn, J. Delivering quality research in culturally dynamic, conflict-affected contexts: lessons from large-scale pilot research in Cox's Bazar [Internet]. Figshare; 2022 [cited 2026 Jan 25]. Available from: https://static1.squarespace.com/static/5fa0560c3a27c834f6a6b5ac/t/62da86b95359a965b88f922a/1658488511491/Delivering+Quality+Research_Final+July+2022-compressed+%281%29.pdf
49. Bayley Scales of Infant and Toddler Development | Fourth Edition. [cited 2023 Dec 10]. Available from: <https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Cognition-%26-Neuro/Bayley-Scales-of-Infant-and-Toddler-Development-%7C-Fourth-Edition/p/100001996.html>
50. Li F, Lokhnygina Y, Murray DM, Heagerty PJ, DeLong ER. An evaluation of constrained randomization for the design and analysis of group-randomized trials. *Stat Med*. 2016;35(10):1565–79. <https://doi.org/10.1002/sim.6813>.
51. Hein S, Bick J, Issa G, Aoude L, Maalouf C, Awar A, et al. Maternal perceptions of father involvement among refugee and disadvantaged families in Beirut, Lebanon. *PLoS ONE*. 2020;15(3): e0229670. <https://doi.org/10.1371/journal.pone.0229670>.
52. Jeong J, Siyal S, Fink G, McCoy DC, Yousafzai AK. "His mind will work better with both of us": a qualitative study on fathers' roles and coparenting of young children in rural Pakistan. *BMC Public Health*. 2018;18(1): 1274. <https://doi.org/10.1186/s12889-018-6143-9>.
53. Guglielmi S, Mitu K, Seager J. I just keep quiet': addressing the challenges of married Rohingya girls and creating opportunities for change. *Eur J Dev Res*. 2021;33(5):1232–51. <https://doi.org/10.1057/s41287-021-00437-6>.
54. Dhira TA, Rahman MA, Sarker AR, Mehareen J. Validity and reliability of the Generalized Anxiety Disorder-7 (GAD-7) among university students of Bangladesh. *PLoS One*. 2021;16(12):e0261590. <https://doi.org/10.1371/journal.pone.0261590>.
55. Waldman M, McCoy DC, Seiden J, Cuartas J, CREDI Field Team, Fink G. Validation of motor, cognitive, language, and socio-emotional subscales using the Caregiver Reported Early Development Instruments: an application of multidimensional item factor analysis. *Int J Behav Dev*. 2021;45(4):368–77. <https://doi.org/10.1177/01650254211005560>.
56. Smeeth D, McEwen FS, Popham CM, Karam EG, Fayyad J, Saab D. War exposure, post-traumatic stress symptoms and hair cortisol concentrations in Syrian refugee children. *Mol Psychiatry*. 2023;28(2):647–56. <https://doi.org/10.1038/s41380-022-01859-2>.
57. Schochet P. Technical Methods Report: Guidelines for Multiple Testing in Impact Evaluations. NCEE 2008-4018 National Center for Education Evaluation and Regional Assistance. 2008.
58. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann Statist*. 2001 Aug 1 [cited 2024 Feb 20];29(4). Available from: <https://projecteuclid.org/journals/annals-of-statistics/volume-29/issue-4/The-control-of-the-false-discovery-rate-in-multiple-testing/10.1214/aos/1013699998.full>
59. D Michael 2026 Replication Data for 'Impact Evaluation of a Father Engagement Model in the Home Visiting Humanitarian Play Lab (HPL) Program in Rohingya Camps and Host Community in Cox's Bazar, Bangladesh' Harvard Dataverse <https://doi.org/10.7910/DVN/60TJQC>
60. Levin HM, McEwan PJ, Belfield C, Bowden AB, Shand R. Economic evaluation in education: cost-effectiveness and benefit-cost analysis. 2455 Teller Road, Thousand Oaks California 91320: SAGE Publications, Inc; 2018 [cited 2025 Aug 4]. Available from: <https://methods.sagepub.com/book/economic-evaluation-in-education-3e>
61. Slaughter J, Nagoshi CT. Testing a contextual model of effects of father involvement on child behaviors. *Child Adolesc Soc Work J*. 2020 [cited 2025 Aug 4];37(5):547–56. Available from: <http://link.springer.com/10.1007/s10560-020-00649-5>
62. van den Berg W, Hendricks L, Hatcher A, Peacock D, Godana P, Dworkin S. 'One Man Can': shifts in fatherhood beliefs and parenting practices following a gender-transformative programme in Eastern Cape, South Africa. *Gend Dev*. 2013 Feb 1 [cited 2023 Dec 10];21(1):111–25. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3758761/>
63. Kroll ME, Carson C, Redshaw M, Quigley MA. Early father involvement and subsequent child behaviour at ages 3, 5 and 7 years: prospective analysis of the UK Millennium Cohort Study. *PLoS One*. 2016;11(9):e0162339. <https://doi.org/10.1371/journal.pone.0162339>.
64. Giallo R, Seymour M, Fogarty A, Hosking C, Williams LA, Cooklin A. Working out dads (WOD): a study protocol for a randomised controlled trial of a group-based peer support intervention for men experiencing mental health difficulties in early fatherhood. *BMC Psychiatry*. 2022;22(1):111. <https://doi.org/10.1186/s12888-022-03698-5>.
65. Zahra F, Sunny S, Bolisetty S, Rahim AK. Perceptions of children's development among Rohingya fathers in forced displacement: does gender of the child matter? [Under Review]

Publisher's Note

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