Implementation Quality of an Early Childhood Parenting Program in Colombia and Child Development

Raquel Bernal, PhD[a], María Lucía Gómez, BSc[b], Santiago Pérez-Cardona, MSc[a], Helen Baker-Henningham, PhD[b]

[a]Universidad de los Andes and CEDE, Colombia
[b]School of Human and Behavioral Sciences, Bangor University, UK

Correspondence:

Raquel Bernal: rbernal@uniandes.edu.co, Tel. +57 (601) 339 4949, Mailing address: Calle 19A No 1-37 Este. Bogotá, Colombia

Short title: Implementation Quality of a Parenting Program

Conflict of Interest Disclosure: The authors have no conflicts of interest relevant to this article to disclose.

Funding/Support: This study was funded by Grand Challenges Canada Grant 0462-03-10 and Fundacion Exito.

Role of Funder: The funders had no role in the data collection, data analyses, data interpretation, writing of the manuscript or the decision to submit for publication.

Clinical Trial Registration: The trial is registered at ISCTRN, number ISRCTN93757590.

Data Sharing Statement: Deidentified individual participant data (including data dictionaries) will be made available through the World Bank Open Data Depository within a year after publication.

Abbreviations:

- Bayley-III: Bayley Scales of Infant and Toddler Development, 3rd edition
- CI: confidence interval
- ECD: early childhood development
- FAMI: Family, Women, and Infancy Program in Colombia
- FCI: UNICEF Family Care Indicators
- IE: Indirect effect
- LMIC: low- and middle-income countries

Contributers' Statement:
Prof Bernal contributed to the conception and design of the study, contributed to data analyses and interpretation and reviewed the manuscript for important intellectual content.

Ms Gómez contributed to the design of the observational instrument, supported the adaptation of the curricula to the setting, trained and mentored intervention staff and critically reviewed the manuscript for important intellectual content.
Mr Pérez carried out the data analyses and contributed to the initial draft of the manuscript.
Prof Baker-Henningham contributed to the conception and design of the study, adapted the curricula and intervention to the setting, supported the training of the intervention team, contributed to the design of the observational instrument and to the interpretation of the data analysis and contributed to the initial draft of the manuscript.
All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Article Summary

Enhancing an existing early childhood parenting program in Colombia increased quality of implementation and quality predicted benefits to maternal engagement, parenting practices and child development.

What's Known on This Subject

Efficacy trials show that early childhood parenting programs benefit parental investment and child development in low- and middle-income countries. To promote and maintain quality at scale, reliable and valid measures of quality are needed and few measures have been tested.

What This Study Adds

Enhancing an early childhood parenting program through provision of structured curricula and training and support for frontline staff led to large benefits to the quality of parenting sessions in Colombia. Session quality was associated with child development and parental investment.

Abstract

OBJECTIVES:
We conducted a cluster-randomised trial of an enhancement to an existing parenting program in rural Colombia (called the FAMI-program), and found benefits to parenting practices and child development. In this study, we examine the effects of the enhancement on the quality of intervention implementation and examine associations between quality and child and maternal outcomes.

METHODS:
In Colombia, 340 FAMI-mothers in 87 towns were randomly assigned to quality enhancement through the provision of structured curricula, play materials, and training and supervision from professional tutors, or to control (no enhancement). Children younger than 12 months were enrolled (N=1460). A subsample of 150 FAMI-mothers (83 intervention, 67 control) in 29 towns (17 intervention, 12 control) participated in the assessment of the quality of group parenting sessions through independent observation. Child development and parenting practices were measured at endline (10.5 months after baseline).

RESULTS:
In ITT analyses, we found significant benefits of intervention for the observed quality of group sessions (1.67 SD (95% confidence interval (CI): 1.23 to 2.11). An SD increase in session quality predicted an increase in treatment mothers' attendance of 4.68 sessions (95% CI: 1.37 to 7.98). Session quality partially mediated the effect of the intervention on parental practices and child development.

CONCLUSIONS:
Enhancing an existing parenting program led to large benefits to the observed quality of intervention implementation. Quality was associated with increased maternal engagement, parenting practices and child development. The observational measure of quality has potential to promote and maintain quality at scale.

Introduction

There is a strong evidence base showing early childhood parenting programs benefit child development in low- and middle-income countries (LMIC).[1] The challenge now is to extend the reach of these programs.[2, 3] An essential component of scaling evidence-based interventions is to promote and sustain quality implementation.[4] Quality implementation of parenting programs encompasses structural attributes, including dosage and content, and process elements, which refer to how the intervention is delivered and the nature of the interactions between the facilitator, mothers and children.[5,6] To promote high-quality services at scale, we need measuring tools that are reliable, low-cost and associated with metrics of program success. While structural quality is relatively easy to measure using checklists and program records, few process quality measures have been validated in early childhood development (ECD) parenting programs in LMIC.[5,7,8] Furthermore, the available quality measures are mostly designed for home-visiting rather than group-based ECD parenting interventions.

In semi-urban and rural areas of Colombia, the Family, Women, and Infancy Program (FAMI-program) provides training and support for economically disadvantaged pregnant women and parents of children up to 2 years of age. The FAMI-program is delivered through group sessions held 2-to-4 times per month and monthly home visits by FAMI-mothers who are paraprofessional women from the local community. It is publicly funded and, on average, costs US$318 per child per year. We designed enhancements that included structured curricula, adapted from Reach-Up and Learn[9] and from a previous adaptation to the Colombian context,[10] and training and ongoing supervision for FAMI-mothers by tutors hired by the research team. In a cluster-randomised trial, we found benefits from these enhancements to child cognitive development (effect size (ES)=0.16)

and to stimulation in the home (ES=0.34).[11] In a complementary qualitative evaluation, participants reported that the techniques used to deliver the content (e.g., demonstration, practice, positive feedback) and the participatory and fun nature of the sessions promoted engagement and learning.[12]

In this study, we designed an observational measure of the process quality of group parenting sessions. We used the measure in a subsample of FAMI-mothers from the cluster-randomised trial to evaluate: 1) the effect of the intervention on session quality, and 2) associations between session quality and parent and child outcomes.

METHODS

Study Design and Participants

For the larger study,[11] we conducted a 2-arm cluster-randomized control trial in 3 districts in rural Colombia. A total of 87 towns participated in the study: 46 intervention, 41 control. Town was the unit of randomization to prevent contamination among FAMI-mothers. All FAMI-mothers within each town participated in the study for a total of 340 (Figure 1). The mean(SD) beneficiaries per FAMI-mother was 11.6(2.8), comprising 9.5(2.9) children younger than two and 2.1(1.7) pregnant women. Within each unit, we enrolled all children under twelve months of age at baseline in the evaluation sample to give a total of 1,456 children (Figure 1). We selected children under twelve months to maximize the potential time of exposure to our intervention before children outgrew the FAMI-program at age 2. At post-test, 319 (93.8%) FAMI-mothers (160 intervention, 159 control) and 1,335 children (91.4%) (628 intervention, 707 control) were evaluated (Figure 1).

For this study, we selected a subsample of towns to participate in the assessment of the quality of the group sessions through observation. The subsample was not randomly

selected, rather it was selected for logistical reasons and includes towns with more FAMI-mothers and those that permitted a shorter fieldwork route to optimize the number that could be included within the cost constraints of the study. The subsample was drawn from 29 out of 87 towns (17 intervention, 12 control). A total of 150 FAMI-mothers (83 intervention, 67 control) with 642 children in the evaluation sample (347 intervention and 295 control) were included in this sub-sample.

Participants were recruited into the study, and baseline measurements conducted between August and November 2014. Video recordings of parenting group sessions took place between July and December 2015, beginning 5-to-7 months after the start of intervention implementation. Post-test measurements were collected between April-to-July 2016. Written informed consent to participate in the study was obtained from FAMI-mothers and primary caregivers of participant children by survey staff at baseline, prior to the observational assessments, and at endline. No participants refused to participate. The study protocol was approved by Universidad de los Andes ethics committee (287/2014) and University College London ethics committee (2168/011). The trial registration number is ISRCTN93757590.

Intervention

The enhancement to the FAMI-program in the intervention group consisted of 4 main components: 1) two structured curricula: one for home visiting and one for group sessions, 2) developmentally appropriate and low-cost play materials (e.g., picture books, puzzles, home-made toys), 3) nutritional education and a food package, and 4) training and supervision by professional tutors trained by the research team. Tutors were responsible for an average of 5 towns and 19 FAMI-mothers and conducted an average of 3.5 weeks and

85 hours of training with the FAMI-mothers in each town. The training involved demonstration and practice in how to conduct the group sessions, how to engage mothers and children in play and language activities, toy-making, and how to promote sensitive, responsive parenting practices. Tutors also provided ongoing supportive supervision to FAMI-mothers through field visits, including attending 1 group session and 1 home visit every 4-to-6 weeks. The intervention lasted for an average of 10.4 months. Further details of the intervention are given in Appendix 1. FAMI-mothers in towns assigned to the control group continued with services as usual. We recorded attendance at group sessions in the intervention group only.

Measurements

Data were collected by an independent organization, IQuartil, with training from study researchers.

The primary outcome in this study is the process quality of the parenting sessions. Group parenting sessions were recorded using a camera on a tripod without a camera operator. Filming took place over 3 rounds with 4-to-6 weeks between each visit. Videos were coded by an independent masked observer using an observational schedule that combined counts of FAMI-mother's use of praise and efforts to promote mother's participation (7 items) with four rating scales: 1) demonstration (two items), 2) practice (three items), 3) atmosphere (seven items), and 4) fun and enjoyment (five items) (Table 1). The categories were designed to include the core delivery components of Reach-Up and Learn, adapted for the group setting and suitable for use with video recordings. All videos were coded over a 3-month period after post-test measurements were completed, when all videos were available. Training for the observer was conducted over 2 weeks: 1 week of

initial training followed by 1 week of inter-observer reliabilities. We randomly selected 15% of videos from each round of filming and conducted ongoing inter-observer reliabilities every week. Interobserver reliabilities (intraclass correlation coefficients) were mean(SD)=0.93(0.06), with a range of 0.86-1.0 (Webtable 1). All subscales had good internal consistency (Conbach's α mean(SD)=0.85(0.09), with a range of 0.69-0.97) (Webtable 2).

We also examined whether quality of the group sessions was associated with child development and parenting practices, two outcomes that showed significant benefits in our impact evaluation.[11] Child development was measured at post-test only using the Bayley Scales of Infant and Toddler Development, 3rd edition (Bayley-III).[13] We use a composite of child cognition, receptive and expressive language, and fine and gross motor development in the analyses.[11] We measured parenting practices at baseline and post-test using the UNICEF Family Care Indicators (FCI).[14] The FCI measures the variety of play materials in the home and the extent to which adults in the home engaged the child in play activities over the past three days.

Randomization and blinding

Towns were randomised before baseline assessments using a random number generator in Stata-13. Participants and intervention staff could not be masked to treatment status. The observer, testers and interviewers were masked to group assignment. However, the observer could have potentially inferred treatment status from activities during group sessions as intervention FAMI-mothers used a structured curriculum. In addition, mothers

may have talked about the intervention with testers/interviewers during endline assessments.

Statistical analysis

The observational sample consisted of 150 FAMI-mothers (83 intervention, 67 control) with at least 1 video recording. Minimum detectable effects were computed using an intracluster correlation coefficient of 0.25. With an average of 5 FAMI-mothers per town and 68 FAMI-mothers in each group, we could detect a difference in the quality of the group session of 0.70SD with 80% power at the 5% significance level.

For the analyses, we first present intention-to-treat (ITT) effects between the treatment and control group on the observed quality of group sessions. We calculated the average of the quality measures (i.e. sum of the count variables and four rating scales), pro-rated to 30 minutes of observation, across the number of observations available for each FAMI-mother. Exploratory factor analysis gave one factor (Webtable 3); factor scores were used in the analyses. SEs were clustered at the town level, and 2-sided p-values were calculated by using *t*-tests. We controlled for covariates to improve precision, in particular, baseline FAMI-mother years of experience, years of education, level of depressive symptoms, verbal ability, early childhood certificate, district fixed effects, and total number of videos. Missing covariates were replaced by sample means.

We then conducted a mediation analysis of the quality of group sessions on the impacts of the intervention on child development (Bayley-III) and parenting practices (FCI). We compared the total ITT effect on the outcome variable with the ITT effect when the mediator was included. We estimated these analyses at the child level, clustered SEs at the FAMI mother level, calculated 2-sided p-values using *t*-tests, controlled for the same

covariates as before, and included child's age sex, and tester fixed effects. We test the statistical significance of the indirect effect using Preacher and Hayes' approach.[15]

In supplementary analyses, we investigated whether session quality predicted child development and parenting practices in treatment and control groups separately.

Finally, we used a Poisson regression to estimate the association between participant attendance to group sessions and the quality of sessions in the treatment arm only. We present average marginal effects. We estimated these analyses at the child level, clustered SEs at the FAMI-mother level, calculated 2-sided p-values using *t*-tests, and controlled for the same covariates as before.

RESULTS

Analyses were conducted on all 150 FAMI-mothers included in the observational sample, and children with completed follow-up testing in the observational sample, with a total of 585 for the Bayley-III and 602 for parental practices in 29 towns (Figure 1). Losses were balanced across groups (see Webtable 4). The only differences between the observational sample and the total sample were a higher proportion of FAMI-mothers with an early childhood certificate (87% vs. 76%, p=0.01) and higher maternal education (9.00 years vs. 8.62, p=0.03) in the observational sample (Webtable 5). We control for these differences in the analyses. Eighteen FAMI-mothers were video-recorded once, 57 were recorded twice, and 75 were recorded three times, with a similar number of video recordings available per FAMI-mother across study groups (Table 2), and few differences in sample characteristics depending on the numbers of video observations conducted (Webtable 6). There were no significant differences in session quality of FAMI-mother

with 1, 2, and 3 videos and no differences across rounds for FAMI-mothers with 3 videos (Webtables 7 and 8).

Table 2 shows baseline characteristics across intervention and control groups in the observational sample. Only maternal verbal ability was significantly different across groups, with higher scores in the treatment group. We control for this in the analyses on child and maternal outcomes.

On average, each video recording was 36 minutes long with similar duration across groups. FAMI-mothers in the intervention group scored higher on all subscales (Table 2). In ITT analyses, we found that the intervention significantly improved the quality of group sessions with an effect size of 1.67 SD (95% CI: 1.23 to 2.11) (Table 3). The intervention had an effect of 0.27 SD (95% CI: 0.05 to 0.49) on child development (Bayley-III composite score) and 0.26 SD (95% CI: 0.08 to 0.45) on parental practices (FCI) in the observational sample (Table 3). After including session quality into the model, we found that session quality partially mediates the effect of the intervention on child development (Indirect Effect (IE): 0.12; 95% CI: -0.01 to 0.25) and parental practices (IE: 0.13; 95% CI: 0.00 to 0.25) (Table 3). When analysing treatment and control groups separately, associations between session quality and child and parent outcomes were evident in the control group only (Webtable 9).

Finally, an SD increase in the quality of the group parenting sessions predicted an increase on treatment mothers' attendance of 4.68 sessions (95% IC: 1.37 to 7.98) (Table 3). Mothers' attendance predicted child and maternal outcomes: for every 10 groups sessions attended, child Bayley test scores increased by 0.10 SD and parental practices increased by 0.04 SD (Webtable 10).

In Appendix 3, we present disaggregated analyses using the individual subscales.

DISCUSSION

In this study, we found that enhancing an existing government parenting program in Colombia (through provision of structured curricula, play materials, and training and supervision for program facilitators) led to significant benefits to the process quality of group parenting sessions measured through independent observation. The quality of the group parenting sessions partially mediated the effect of the intervention on parenting practices and child development. We also report a positive association between quality and treatment mothers' attendance at the group sessions; higher attendance was also associated with greater benefits to child development and parenting practices. We have previously reported that benefits to parenting practices mediated the effect of the intervention on child development.[11] These results suggest a pathway from high-quality implementation to maternal engagement to benefits to parenting practices, leading to benefits to children's development, which is consistent with mechanisms of action underpinning ECD parenting interventions.[17]

Previous studies have demonstrated that ongoing training and supervision improve the quality of implementation of ECD parenting programs over time, in both home-visiting programs[18] and group parenting sessions.[19] In this study, video recordings of group sessions were conducted after approximately 6 months of implementation, and even within this relatively short timeframe, we found large benefits to the quality of the sessions.

The findings that the group-session quality was associated with mothers' attendance, parenting practices, and child development provide empirical evidence for the importance of the behavioral techniques used in intervention delivery. These behaviors include using participatory, interactive methods, active learning techniques, making

sessions fun and promoting positive relationships. These factors have been identified as enablers to effective implementation in previous qualitative and implementation studies.[18,20-22] However, few studies have examined empirical associations between the quality of implementation of ECD programs and child and maternal outcomes in LMIC. In Kenya, higher quality implementation of group sessions, as rated by program supervisors, was associated with higher maternal attendance and higher levels of stimulation in the home. No associations were found with child development.[19] In Peru, observational assessments of the quality of home visits conducted within a large-scale ECD program were significantly associated with child development on the Ages and Stages Questionnaire (ES=0.15-0.25) but not on the Bayley test (ES=0.003-0.07).[7] In both studies, analyses were conducted in the intervention group only. When disaggregating by group, we found positive association between session quality and outcomes in the control group only. This may be due to insufficient variability within the treatment group (over 80% of intervention FAMI-mothers scored above four out of a maximum of five on the rating scales), suggesting that with the initial training and ongoing coaching provided throughout the intervention period, a high and fairly uniform level of implementation quality was achieved. There may also be a threshold which could serve as a benchmark in program monitoring. In this study, training and support was provided by professional tutors hired by the research team. In the future, it will be important to test whether it is possible to maintain implementation quality using the government supervisors of the FAMI program, or whether additional child development supervisors are required.

The finding of positive associations between session quality and outcomes in the control group suggest that the observation tool, although informed by Reach-Up methods, could be a useful measure of quality in general, not only for interventions based on Reach-

Up. Use of the observation tool could be incorporated into ongoing supervisory visits which would guide program supervisors in providing appropriate feedback and support to facilitators during each visit, and would provide timely data on implementation quality and thus inform wider training needs. However, the tool would likely need to be supplemented with some program-specific checklists that record aspects related to the content.

The study's strengths include using observational measures to assess the quality of sessions, good psychometric properties of the observational measure, the use of masked assessors, and the fact that the study was nested within a cluster-randomised trial with a treatment and control group. Due to cost constraints, we could not randomly select FAMI-mothers to participate in this nested study; however, the subsample was reasonably representative of the full sample. We minimized FAMI-mother reactivity to being observed by using a camera on a tripod without a camera operator. We also conducted 3 rounds of observations to maximise the likelihood that the quality score was an accurate indicator of quality across groups. Unfortunately, due to logistical and technical challenges, only half of the sample had all 3 video recordings; however, there were few differences in FAMI-mother's characteristics and quality of implementation between those with 1, 2 or 3 videos. Likewise, session quality for those with 3 videos was similar across rounds. Although the FAMI-program consists of group sessions and home visits, we measured the quality of the group parenting sessions only. In addition, we did not measure aspects related to the content of the sessions. Instead, we focused on the process quality of implementation.

Our results show that the process quality of parenting group sessions was associated with benefits to mother engagement, parenting practices and child development. The observational measure used in this study has potential for monitoring the effectiveness of

261 training and support provided to frontline staff delivering ECD group-based parenting
262 programs in LMIC and improving program delivery.

REFERENCES

1. 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. PLoS Med (2021) 18:e1003602. doi:10.137/journal.pmed.1003602.
2. Britto PR, Lyle SJ, Proulx K, et al. Nurturing care: promoting early childhood development. *Lancet.* 2017;389:91-102. doi:10.1016/S0140-6736(16)31390-3
3. Richter LM, Daelmans B, Lombardi J, et al. Investing in the foundation of sustainable development: pathways to scale up for early childhood development. *Lancet.* 2017:389:103–118. doi:10.1016/S0140-6736(16)31698-1.
4. Durlak JA, DuPre EP. Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American J Community Psychol.* 2008;41☹3–4), 327–350. doi: 10.1007/s10464-008-9165-0
5. Schodt S, Parr J, Araujo MC, Rubio-Codina M. Measuring the quality of home-visiting services. IDB Technical Note Numer IDB-TN-881. Washington DC: Inter-Amercian Development Bank, 2015
6. Leer J, Lopez-Boo F. Assessing the quality of home visit parenting programs in Latin America and the Caribbean. *Early Child Dev Care*. 2019;189(13):2183–2196. doi: 10.1080/03004430.2018.1443922
7. Araujo MC, Dormal M, Rubio-Codina M. (2018). Quality of Parenting Programs and Child Development Outcomes: The Case of Peru's Cuna Mas. IDB Working Paper Series No IDB-WP-951. Washington DC: Inter-American Development Bank.
8. Rubio-Codina M, Dormal M, Araujo MC. Observing home-visiting quality at scale with the home visit rating scales and a supervisor checklist in Peru. *Infant Ment Health J.* 2019;40(3):343–6. doi: 10.1002/imhj.21775
9. Walker S, Chang S, Powell C, Smith J, Baker-Henningham H, and the Reach-Up Team. The Reach Up Early Childhood Parenting Program: origins, content and implementation. *Zero to Three* 2018;38(4):37-43.
10. Attanasio OP, Fernández C, Fitzsimons EOA, Grantham-McGregor SM, Meghir C, Rubio- Codina M. Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in Colombia: Cluster randomized controlled trial. *BMJ.* 2014;349:g5785. doi: 10.1136/bmj.g5785.
11. Attanasio O, Baker-Henningham H, Bernal R, Meghir C, Pineda D, Rubio-Codina M. (2022) Early stimulation and nutrition: The impacts of a scalable intervention. *J Eur Econ Assoc.* doi: 10.1093/jeea/jvac005.
12. Gomez ML, Bernal R, Baker-Henningham H. Qualitative evaluation of an early childhood parenting program integrated into government services in rural Colombia. *Child Care Health Dev.* 2021;48:225-238. doi: 10.1111/cch.12921.
13. Bayley N. Bayley Scales of Infant and Toddler Development: third edition. Cambridge, UK: Pearson Publishing 2005.
14. Kariger P, Frongillo EA, Engle P, Britto PMR, Sywulka SM, Menon P. Indicators of family care for development for use in multicountry surveys. *J Health Popul Nutr.* 2012;30:472-86. doi:10.3329/jhpn.v30i4.13417.

15. Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior Research Methods, 40(3), 879–891. doi: 10.3758/BRM.40.3.879.
16. MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, 7(1), 83–104. doi: 10.1037/1082-989X.7.1.83
17. Aboud FE, Prado EL. Measuring the implementation of early childhood development programs. *Ann NY Acad Sci.* 2018;1419:249-263. doi: 10.1111/nyas.13642
18. Yousafzai AK, Rasheed MA, Siyal S. Integration of parenting and nutrition interventions in a community health program in Pakistan: an implementation evaluation. *Ann NY Acad Sci.* 2018;1419:160-178. doi: 10.1111/nyas.13649.
19. Luoto JE, Lopez-Garcia I, Aboud FE, et al. An Implementation Evaluation of A Group-Based Parenting Intervention to Promote Early Childhood Development. *Front Public Health*. 2021;9:653106. doi: 10.3389/fpubh.2021.653106
20. Smith JA, Baker-Henningham H, Brentani A, Mugweni R, Walker SP. Implementation of Reach Up early childhood parenting program: acceptability, appropriateness, and feasibility in Brazil and Zimbabwe. *Ann N Y Acad Sci*. 2018;1419(1):120–40. doi: 10.1111/nyas.13678
21. Singla DR, Kumbakumba E. The development and implementation of a theory-informed, integrated mother-child intervention in rural Uganda. *Soc Sci Med.* 2015;147: 242–251. doi: 10.1016/j.socscimed.2015.10.069.
22. Tomlinson M, Hunt X, Rotheram-Borus MJ. Diffusing and scaling evidence-based interventions: eight lessons for early child development from the implementation of perinatal home visiting in South Africa. *Ann NY Acad Sci.* 2018;1419(1): 218–229. doi: 10.1111/nyas.13650

TABLE 1. Description of the Observational Instrument

Construct	Questions	Coding	Score
COUNT VARIABLES			
Praise: FAMI-mother praises the mothers and children	- FAMI-mother praises beneficiary mothers - FAMI-mother praises babies - FAMI-mother praises the group - FAMI mother says good things about the children to beneficiary mothers	Event sampling used to code each praise statement	Count variable: sum of all items
Promoting participation: FAMI-mother encourages mothers' contributions to the group session.	- FAMI-mother expands on beneficiary mothers' contributions - Beneficiary mothers participate in the session - FAMI-mother asks open-ended questions		
RATING SCALES			
Demonstrations: FAMI-mother demonstrates how activities with mothers and children	- FAMI-mother demonstrates activities. - Clarity of demonstrations.	5-point rating scale: - 1=inadequate - 3=adequate - 5=excellent	Mean score of 2 items
Practice: FAMI-mother provides opportunities for mothers to practice activities introduced in the session.	- Beneficiary mothers practice activities. - FAMI-mother supports beneficiary mothers while practicing - FAMI-mother gives sufficient time for beneficiary mothers to practice the activities	5-point rating scale: - 1=inadequate - 3=adequate - 5=excellent	Mean score of 3 items
Atmosphere: FAMI-mother creates a welcoming, supportive, engaging and collaborative atmosphere during the session	- Seating arrangement facilitates collaborative approach - FAMI-mother sits at the same level as beneficiary mothers and babies - FAMI-mother gives positive affirmations to beneficiary mothers and babies - FAMI-mother involves beneficiary mothers (using a collaborative and participatory approach) - FAMI-mother uses beneficiary mothers' names - FAMI-mother uses babies' names - Beneficiary mothers answer questions and share information and ideas	5-point rating scale: - 1=inadequate - 3=adequate - 5=excellent	Mean score of all 7 items
Fun and enjoyment: Evidence of enjoyment by beneficiary mothers, babies and FAMI-mother	- Sufficient toys for all the babies - Toys are available for babies for the entire session - Beneficiary mothers enjoy the sesión (show positive affect) - Babies enjoy the session (laugh, smile, play) - FAMI-mother enjoys the sesión (shows positive affect)	5-point rating scale: - 1=inadequate - 3=adequate - 5=excellent	Mean score of all 5 items

See Appendix 4 for the observational tool used in this study.

TABLE 2 Observational Sample Characteristics at Baseline and Follow-Up by Treatment Arm

	Intervention	Control	p-value
Baseline Characteristics			
Panel A: FAMI-mother characteristics	N = 83	N = 67	
FAMI-mother's age, years, mean (SD)	43.66 (9.52)	42.00 (9.63)	0.36
FAMI-mother's years of schooling, mean (SD)	13.24 (1.60)	13.57 (1.98)	0.48
FAMI-mother's years of experience, mean (SD)	12.62 (8.33)	13.64 (8.84)	0.47
Early childhood certification, n (%)	69 (83)	61 (91)	0.26
PPVT (raw score), mean (SD)	30.69 (11.07)	25.37 (11.03)	0.09
Depressive symptoms (CES D-10), n (%)[a]	14 (18)	5 (7)	0.06
Panel B: Characteristics of FAMI-mother's group	N = 83	N = 67	
Number of children between 0 and 12 months, mean (SD)	4.75 (1.87)	4.85 (2.31)	0.87
Number of pregnant mothers, mean (SD)	1.82 (1.36)	2.01 (1.61)	0.53
Number of meetings (last month), mean (SD)	5.72 (5.57)	4.45 (3.13)	0.37
Number of home visits (last month), mean (SD)	11.98 (5.72)	14.79 (8.65)	0.24
Activities' planning time (hours/week), mean (SD)	4.72 (3.20)	5.84 (6.18)	0.29
Panel C: Child characteristics	N = 347	N = 295	
Age in months, mean (SD)	5.88 (3.28)	5.41 (3.31)	0.15
Male, n (%)	173 (50)	153 (52)	0.65
Low birth weight, n (%)	30 (9)	22 (7)	0.63
Stunting, n (%)	27 (8)	42 (16)	0.15
Panel D: Household characteristics	N = 347	N = 295	
Maternal years of schooling, mean (SD)	8.88 (3.41)	9.14 (3.11)	0.50
Maternal age, years, mean (SD)	25.89 (6.76)	26.78 (6.34)	0.13
Maternal PPVT (raw score), mean (SD)[b]	23.22 (8.82)	19.22 (7.56)	0.04
Father present, n (%)	243 (70)	222 (75)	0.25
Household in poverty, n (%)[c]	202 (59)	174 (62)	0.73
Quality of the Home Environment (FCI), mean (SD)	0.12 (0.94)	0.00 (0.90)	0.46
Follow-Up Characteristics			
Panel E: Video Observations Characteristics	N=83	N=67	
Number of observations			
• One video, n (%)	9 (11)	9 (13)	
• Two videos, n (%)	26 (31)	31 (46)	0.10
• Three videos, n (%)	48 (58)	27 (40)	
Number of different child age ranges present (0-5 months, 6-11 months, 12-24 months), median (SD)	1 (0.63)	2 (0.65)	0.12
Number of children present, mean (SD)	3.55 (1.21)	3.93 (1.91)	0.39
Duration of observations (minutes), mean (SD)	36.21 (10.82)	36.96 (10.50)	0.72
Panel F: Quality of Sessions	N=83	N=67	
Sum Count Variables, mean (SD)	64.11 (23.96)	23.83 (13.86)	<0.001
Mean demonstration over all observations, mean (SD)	4.63 (0.73)	3.06 (1.44)	<0.001
Mean practice over all observations, mean (SD)	4.57 (0.70)	3.16 (1.28)	<0.001
Mean atmosphere over all observations, mean (SD)	4.48 (0.62)	3.08 (0.76)	<0.001
Mean fun over all observations, mean (SD)	4.55 (0.60)	3.54 (1.04)	<0.001

[a] Three FAMI-mothers in the intervention arm have missing data in the CES D-10 due to incomplete baseline survey. [b] Spanish version of the Peabody Picture Vocabulary, a proxy for maternal IQ. [c] Indicator variable that equals one if the household's total income is below the poverty line in 2014 ($50 USD person/month). We present 2-sided p-values in column 3. While for continuous and indicator variables we calculated p-values using t-tests, for categorical variables (with more than two categories) we used a Pearson's chi-squared test.

TABLE 3 Treatment Effect on the Quality of Group Sessions and Mediation Analysis

Dependent Variable:	Quality of Group Sessions Factor Score	Bayley-III Factor	Bayley-III Factor	Parental Practices (FCI)	Parental Practices (FCI)	Number of Groups Sessions Attended
	(1)	(2)	(3)	(4)	(5)	(6)
Independent Variables:						
Treatment	1.67	0.27	0.15	0.26	0.14	
95% CI	(1.23 to 2.11)	(0.05 to 0.49)	(-0.10 to 0.41)	(0.08 to 0.45)	(-0.08 to 0.35)	
p-value	[<0.001]	[0.02]	[0.24]	[0.006]	[0.22]	
Quality of Group Sessions Factor Score			0.09		0.10	4.68
95% CI			(-0.01 to 0.19)		(0.01 to 0.19)	(1.37 to 7.98)
p-value			[0.07]		[0.04]	[0.006]
Observations	150	585	585	602	602	347
Indirect Effect			0.12		0.13	
95% CI			(-0.01 to 0.25)		(0.00 to 0.25)	
p-value			[0.08]		[0.04]	

Estimated coefficients in columns 1 to 5 are expressed in SDs of the control group. Estimates in column 1 are at the FAMI-mother level; columns 2 to 6 are at the child level. While in columns 1 to 5 the sample includes all FAMI-mothers with at least one video, in column 6 we restricted the sample to the intervention group, as we do not have information on attendance for the control group. In the treatment group, 101/347 (29.1%) attended zero sessions, the median number of sessions attended was 17; the maximum number of sessions was 42. A family could have attended a maximum of 44 weekly group sessions during the study period. Estimates controlled for baseline FAMI mother's years of experience, years of education, level of depressive symptoms by CESD10, verbal ability using the Spanish version of the Peabody Picture Vocabulary Test, early childhood certificate, district fixed effects, and the total number of videos. Columns 2 to 5 also include interviewer fixed effects. The p-values are 2-tailed conventional p-values. CIs were constructed by using conventional critical values for individual hypotheses. The intracluster correlation coefficient for the primary outcome (quality of group sessions factor score) was 0.24. Missing data in control variables were replaced by sample means. We explored alternative imputation strategies for missing values (i.e., replacement with sample median and regression imputation). Results are robust to these alternative approaches. To test the statistical significance of the indirect effect we follow Preacher and Hayes (2008)[15] and bootstrapped the indirect effect with 2,000 replications to compute the p-value. Results are robust to using the test of the joint significance, as describe by MacKinnon et al (2002)[16].

Figures List:

Legend	File Name
FIGURE 1. Trial Profile	figure1-trial-profile.pdf

Appendix 1: Description of the Program

A. Description of the Existing FAMI-Program

The existing FAMI program is run by the Colombian Family Welfare Agency (ICBF for its acronym in Spanish). The program supports vulnerable families with nutrition, health monitoring, and childrearing and targets pregnant women and parents with children younger than 2 years old. It is delivered through group sessions and home visits by the FAMI-mother, paraprofessional women from the community. The FAMI program provides general operational guidelines and broad learning standards. FAMI mothers are expected to use these guidelines and standards to plan the content to be introduced through group sessions and home visits. Group meetings take place in schools, churches or the FAMI mother's own home. FAMI units vary between 10 and 24 beneficiaries (Mean=13, SD=1.4). Close to 80% are parents of children 0-24 months of age and 20% are pregnant women. FAMI mothers participate in an initial training workshop of approximately 60 hours provided directly by ICBF and also attend 8 hours of additional training every month. FAMI mothers are supervised by ICBF staff at the local (municipality) level. Supervision involves on-site visits to document aspects related to FAMI mothers' record keeping and planning and the physical characteristics of the setting for the group venue. The program also delivers a nutritional supplement corresponding to 22% to 27% of the recommended calorie intake (monthly). The average cost of the FAMI program is $318 US (US dollars or USD) per child per year, and it is publicly funded.

In sections B to L below, we describe the enhanced FAMI program used in this study.

B. Aims of the Enhanced FAMI Program

- Strengthening the child development component of existing curriculum in order to improve the children's development
- To strengthen the nutrition component of the existing FAMI program by providing psychoeducation around feeding and nutrition
- To improve mothers' knowledge, practices and enjoyment of bringing up her child
- To improve mothers' self-esteem and mental health

C. Program Components

The aims above are achieved through the following activities that are implemented during the parenting group sessions and the home visits. The FAMI mothers generally work with the mother but the father, grandparents and other family members are also be encouraged to participate in the activities below if they are available during the group sessions and home visits.

- Making the mother agent of change in promoting her child's development
- Demonstrating the use of age-appropriate play materials and activities
- Providing opportunities to practice age-appropriate activities and provide supportive feedback
- Setting up a toy and child library to use at home.
- Providing opportunities for mothers to make low-cost toys and practice using them in ways that promote specific child development goals.
- Providing opportunities for social support, sharing of experiences, and group problem-solving

- Increasing mothers' motivation to improve her child's development by helping her to understand how her actions can make a difference to her child.
- Improving mother's self-esteem through the use of praise, support and encouragement
- Promoting sensitive and responsive parenting and appropriate behavior management
- Encouraging positive mother-child interactions and preventing child maltreatment

D. Goal of the Home Visit Curriculum

The majority of the program content is delivered through the group visits as they are held on a weekly or fortnightly basis. However, monthly home visits provide the opportunity to reinforce the material covered, to introduce activities that require more individualized instruction, and to personalize to specific needs of the family as necessary. The goals of the monthly home visit are:

- To give and receive feedback about the group sessions and reinforce mothers' participation in all program activities
- To reinforce the key messages learnt in the group meetings
- To ensure the activities that the mother and child are bringing home from the group sessions are developmentally appropriate and to introduce additional activities targeted to the child's age and developmental level.
- To integrate the program activities discussed and practiced in the group sessions into every day family activities
- To identify materials in the home that can be used to promote child development and to help the family to use their home environment to promote child development
- To encourage appropriate mother-child interaction with a strong focus on promoting children's language development in a variety of ways. This is a priority for the home visits as the focus is on individual mother-child dyads.
- To encourage family involvement in program activities
- To engage in problem solving around attendance at group sessions, the ability to do activities at home in addition to any individual concerns or issues that the family may have.

E. Methodology for Home Visits (approx. 1 hour in length)

- Greeting
- Enquire about child and mother and family, and follow up on information shared during previous visit
- Enquire about experiences of group meetings & reinforce participation (problem-solve if necessary)
- Review frequency and type of engagement with toy or book currently in the home and review nutrition message from previous visit (problem-solve if necessary)
- Introduce new play and language activities with a particular focus on play activities that are more difficult to introduce in the group setting (e.g., puzzles, sorting and matching activities, crayon and paper)
- Discuss ways and introduce specific activities to promote child language development
- Discuss how to integrate play and language activities into everyday routines

F. Goals of the Group Curriculum

- Provide opportunities to share parenting experience in a group setting
- Provide opportunities to discuss and practice effective child rearing skills and positive interactions with children

- Demonstrate and practice the use of age-appropriate play materials and demonstrate and practice appropriate language activities and discuss how these help in children's development
- Set up a toy library for home use and show mothers how to make simple toys

G. Methodology and Structure for Group Sessions (approx. 1 hour in length)

Each group session consists of six components:

- Arrival and free play and song
- Feedback from previous session
- Discussion around a parenting theme or activities
- Demonstration and practice of age-appropriate play activity and language activity for the week (with material that will be taken home)
- Review of session – to ensure mothers understand the activities
- Snack

H. Organisation of Group Sessions:

Mothers are asked to attend a group meeting according to the age of their children.

- Pregnant and lactating with children up to 6 months
- Mothers with children from 6 to 11 months
- Mothers with children aged 1-2 years

However, in practice this did not always occur and the curriculum was designed so that it can be delivered to groups with children over the entire age range. The song and parenting message are common to all mother-child dyads and the play and language activities are divided into age bands (birth-5 months, 6-11 months and 1-2 years).

I. Curricula for the Enhanced FAMI Program

Two curricula are used in the enhanced FAMI Program: a group session curriculum and a home-visiting curriculum. The curricula include discussion topics or key parenting messages, a selection of age-appropriate activities to promote child development using simple play materials (e.g., home-made toys, materials in the home, puzzles) and activities to promote children's language development (using games, books, pictures and a using everyday activities to encourage mothers to talk more with their child). Mothers are given one developmentally appropriate book or toy at each session and then the book or toy is swapped for a different book or toy at the next session. Mothers and other family members are also encouraged to make their own toys and books for their child. The curriculum also includes a set of nutrition cards that are discussed the mother during each home visit. Mothers receive a nutrition card relevant to their child's age at these monthly home visits.

J. Training and Coaching of FAMI Mothers

In addition to the set of activities and materials, the enhanced FAMI program includes a coaching component (in-service training) to support and maintain the quality of home visits and group meetings. Shifting away from a supervision model, the new approach consists of a team of tutors who provide the initial pre-service training and then continue to provide in-service training and support during the implementation period. Tutors were required to have professional degrees in psychology or social work. They also had to provide evidence of experience with children younger than 5 -preferably younger than 2- and having worked with communities. Interested professionals applied to an open call. Shortlisted candidates were requested to provide an essay explaining the reasons for their interest in the project and interviewed by members of the research team, Fundacion Exito and an expert in human resources. Tutors train FAMI mothers sequentially by town

and in each town, all FAMI mothers were trained simultaneously. The average training time was 3.5 weeks and 85 hours. However, training time differs depending on the number of FAMI units:

- Towns with $\leq$ 5 FAMI mothers received 75 hours of training in 3 weeks
- Towns with 6-9 FAMI mothers received 100-125 hours for 5-6 weeks
- Towns with 10 or more FAMI mothers received 150-175 hours of training over 6-7 weeks

Training involves demonstration and practice of all play and language activities, toy making sessions, demonstration and practice on how to conduct the entire group session (including feedback, discussions and review) and how to conduct the individual home visits, including discussing the nutrition cards.

Tutors also coach FAMI mothers during one group session and one home visit approximately once every 6 weeks. During these coaching visits, tutors assist the FAMI mother with planning, provide assistance and support during the session and provide supportive feedback and advice to the FAMI mother based on observations recorded on a structured checklist. When feasible, tutors also facilitate a group meeting of FAMI mothers in each town to discuss and share positive experiences and challenges and engage in collaborative problem-solving. The facilitators were supervised by an intervention supervisor who conducted visits with each facilitator every 2 months.

K. Food Supplementation, Nutrition Cards and Messages

Each family received nutritional supplementation every month that correspond to 35% of the daily calorie intake for pregnant women breastfeeding mothers and children young than 2 years of age for a 30-day period. The nutrition package included tuna, sardines, canola oil, iron-fortified whole milk, beans and lentils. The cost of the package is $26 US per month including shipping costs, it is delivered for 11 of the 12 months of the year. Parents were also provided with a recipe book and we prepared 18 nutrition cards as detailed below. These cards are given out and discussed at each monthly home visit. Two of the group sessions also focus on nutrition. A list of the cards and the ages given are shown below:

- Birth: Nutrition for lactating mothers
- Birth: Breastfeeding your baby
- Birth: How to breastfeed
- 1 month: Benefits of breastfeeding
- As necessary: For mothers who are bottle feeding
- 2 months (and 4 months): Chatting while breastfeeding/feeding
- 3 months (and 5 months): How to extract and store breastmilk
- 6 months (and 7 months): Beginning to eat new foods
- 6 months (and 15 and 22 months): What to do when your child has diarrhea
- 7 months (and 14 months): Hygiene
- 8 months (and 10 months): Finger foods
- 9 months (and 11 months): Feeding babies aged 9-12 months
- 12 months (and 15 months): Feeding 1–2-year-old children
- 13 months (and 19 months): Making mealtimes a special time
- 14 months (and 18 and 21 months): Iron
- 16 months (and 23 and 24 months): Menu ideas
- 17 months (and 20 months): Chatting to baby while feeding
- 24 months: Feeding your 2–3-year-old child

L. Key Content for Group Meetings (Discussion Topics)

The key content for the group sessions is shown in the table below. There are 20 group sessions which are suitable for all ages and 4 group sessions specially designed for mothers of babies from birth to 5 months. The sessions repeat every 5-8 months depending on how frequently the FAMI-mother conduct the group sessions. On average, FAMI-mothers conduct 3-4 group sessions a month with mothers of children aged 6-24 months and 1-2 visits a month with children aged from birth to 5 months. In addition, FAMI mothers conduct a home visit with each mother once a month.

Summary of Group Curriculum

Session	Parenting Topic	Play Activity	Language Activity
SESSIONS FOR ALL AGES			
1	Importance of spending time playing with baby	Blocks	Having conversations and learning new words
2	Importance of praising your child	Soft ball	Chatting with baby while bathing
3	Talking with our baby	Picture book	Walk and talk with baby
4	Share in things your child likes to do	Shaker	Responding to your child
5	Things to do at bath time	Doll	Body parts
6	Helping our child learn	Stacking bottle tops	Learning new words2
7	Learning to trust	Blocks and container	Learning names of people
8	Looking at books and pictures with your child	Picture book	Find-it-game
9	Things to do while dressing child	Teething ring on string and container/posting bottle	Playing peek-a-boo
10	Giving sense to your child's world	Picture book	Chatting to baby while feeding him/her
11	Understanding your child's feelings	Teething ring/tin to roll/ putting rings on a bottle	Using baby's name
12	Singing with baby	Drum rattle	Chatting to baby while dressing
13	Making mealtimes a special time	Doll	Making mealtimes a special time
14	Things to do while doing chores	Shaker and container	Learning new words3
15	Finding things in the home to play with	Ring on a string / Push-a-long	Find it game
16	Helping your child learn action words	Picture book	Chatting to baby while doing housework
17	Teaching baby about their environment	Blocks	Responding to your child
18	Helping your child learn to behave 1	Books and pictures	Having fun playing game with baby
19	Helping your child learn to behave 2	Teething ring / sock doll and car	Having fun playing games with baby 2
20	Good nutrition	Doll	Following directions
SESSIONS FOR BABIES AGED FROM BIRTH TO 5 MONTHS			
1	Love and comfort your baby	Ring necklace	Imitate sounds
2	Getting to know your baby	Soft ball	First conversations
3	Babies learn from birth	Looking at pictures	Chatting with baby
4	Breastfeeding your baby	Teething ring	Singing with baby & responding to baby

Appendix 2: Webtables

WEBTABLE 1. Interobserver reliabilities of Observational Assessment of Quality of Group Sessions

ITEM	Intraclass Correlation Coefficient N=54
FAMI-mother praises beneficiary mothers	0.97
FAMI-mother praises babies	0.95
FAMI-mother praises the whole group	0.92
FAMI-mother says good things to mothers about babies	0.95
FAMI-mother expands what mothers say	0.91
Mothers' contributions	0.96
FAMI-mother asks open questions	0.96
FAMI-mother demonstrates activities	0.98
FAMI-mother gives clear demonstrations	0.97
Mothers practice activities with baby	0.98
FAMI-mother supports mothers as they practice	0.98
FAMI-mother gives time for mothers to practice	0.97
Mothers seating arrangements	0.71
FAMI-mother sits at the same level	0.92
FAMI-mother gives supportive feedback to mothers/babies	0.91
FAMI-mother involves all mothers	1.00
FAMI-mother calls mothers by their names	0.86
FAMI-mother calls babies by their names	0.89
Mothers participate in the session	0.98
There are sufficient toys	0.85
Toys are available to children throughout the session	0.98
Mothers have fun	0.96
Babies have fun	0.97
FAMI-mother enjoys session	0.89

WEBTABLE 2. Internal Reliabilities (Cronbach's Alpha) of Observations by Round

Score	Round 1	Round 2	Round 3
Sum of count variables	0.69	0.73	0.70
Demonstration subscale	0.97	0.96	0.97
Practice subscale	0.92	0.93	0.89
Atmosphere subscale	0.84	0.79	0.79
Fun subscale	0.87	0.84	0.83

WEBTABLE 3. Factor analysis of Subscales of Observations of Group Quality

	Factor Loading
Atmosphere	0.93
Practice	0.91
Demonstration	0.90
Fun	0.89
Sum of count variables	0.73
Variance explained	*76.38%*

WEBTABLE 4. Attrition Analysis in the Observational Sample

	Surveyed at follow-up	Lost to follow-up	p-value
Panel A: Child characteristics	N = 602	N = 40	
Treatment, n (%)	322 (53)	25 (63)	0.41
Age in months, mean (SD)	5.61 (3.32)	6.40 (2.97)	0.05
Male, n (%)	311 (52)	15 (38)	0.06
Birth weight, grams, mean (SD)	3,155.38 (487.71)	3,039.34 (368.96)	0.08
Low birth weight, n (%)	50 (8)	2 (5)	0.41
Stunting, n (%)	65 (12)	4 (12)	0.99
Panel B: Household characteristics	N = 602	N = 40	
Maternal years of schooling, mean (SD)	8.98 (3.32)	9.24 (2.63)	0.64
Maternal age, years, mean (SD)	26.48 (6.62)	23.63 (5.38)	<0.001
Maternal PPVT (raw score), mean (SD) [a]	21.48 (8.51)	20.02 (8.24)	0.39
Father present, n (%)	441 (73)	24 (60)	0.08
Household in poverty, n (%) [b]	354 (61)	22 (55)	0.55
Household income > median, n (%)	317 (53)	21 (53)	0.98
Quality of the Home Environment (FCI), mean (SD)	0.06 (0.93)	0.10 (0.76)	0.77

Observational sample refers to children assigned to a FAMI-mother with at least one video recording. The unit of observation in this table is the child. [a] Spanish version of the Peabody Picture Vocabulary Test, a proxy for maternal IQ. [b] Indicator variable that equals one if the household's total income is below the poverty line in 2014 ($50 USD person/month). We present 2-sided p-values using *t*-tests.

WEBTABLE 5. Balance Between the Full Sample and the Observational Sample

	Full sample	Observational Sample	p-value
Panel A: FAMI-mother characteristics	N = 340	N = 150	
FAMI-mother's age, years, mean (SD)	41.60 (10.19)	42.92 (9.57)	0.17
FAM-mother 's years of schooling, mean (SD)	13.13 (1.82)	13.39 (1.78)	0.15
FAMI-mother 's years of experience, mean (SD)	11.79 (8.24)	13.07 (8.55)	0.12
Early childhood certification, n (%)	260 (76)	130 (87)	0.005
Married, n (%)	77 (23)	33 (22)	0.87
Number of FAMI-mothers' own children, mean (SD)	2.63 (1.43)	2.67 (1.33)	0.76
FAMI's household size, mean (SD)	3.94 (1.45)	3.89 (1.45)	0.70
TVIP's total score (Z), mean (SD)	28.49 (10.44)	28.31 (11.33)	0.87
Knowledge's total score, mean (SD)	7.20 (1.57)	7.18 (1.73)	0.90
Depressive symptoms (CES D-10), n (%) [a]	47 (14)	19 (13)	0.76
Panel B: Characteristics of FAMI-mother's group	N = 340	N = 150	
Number of children between 0 and 12 months, mean (SD)	4.99 (2.18)	4.79 (2.07)	0.34
Number of pregnant mothers, mean (SD)	1.91 (1.39)	1.91 (1.48)	0.99
Number of meetings (last month), mean (SD)	5.27 (3.99)	5.15 (4.66)	0.77
Number of home visits (last month), mean (SD)	12.77 (6.92)	13.24 (7.29)	0.51
Activities' planning time, mean (SD)	5.91 (5.43)	5.23 (4.81)	0.17
Panel C: Child characteristics	N = 1456	N = 642	
Age in months, mean (SD)	5.61 (3.32)	5.66 (3.30)	0.76
Male, n (%)	748 (51)	326 (51)	0.80
Birth weight, grams, mean (SD)	3,171.83 (535.64)	3,148.42 (481.98)	0.33
Low birth weight, n (%)	104 (7)	52 (8)	0.46
Stunting, n (%)	158 (12)	69 (12)	0.89
Panel D: Household characteristics	N = 1456	N = 642	
Maternal years of schooling, mean (SD)	8.62 (3.37)	9.00 (3.28)	0.02
Maternal age, years, mean (SD)	26.32 (6.77)	26.30 (6.58)	0.94
Maternal PPVT (raw score), mean (SD) [b]	20.99 (8.39)	21.38 (8.50)	0.33
Father present, n (%)	1056 (73)	465 (72)	0.96
Household in poverty, n (%) [c]	874 (61)	376 (60)	0.58
Household income > median, n (%)	728 (50)	338 (53)	0.26
Quality of the Home Environment (FCI), mean (SD)	0.00 (0.99)	0.06 (0.92)	0.17

Observational sample refers to children assigned to a FAMI-mother with at least one video recording. [a] Three FAMI-mothers in the intervention arm have missing data in the CES D-10 due to incomplete baseline survey. [b] Spanish version of the Peabody Picture Vocabulary Test, a proxy for maternal IQ. [c] Indicator variable that equals one if the household's total income is below the poverty line in 2014 ($50 USD person/month). We present 2-sided p-values using *t*-tests.

WEBTABLE 6. Balance at Baseline in the Observational Sample by Number of Available Video Recordings

	One Video	Two Videos	Three Videos	p-value
Panel A: FAMI-mother characteristics	N = 18	N = 57	N = 75	
FAMI-mother 's age, years, mean (SD)	41.39 (9.59)	42.58 (10.04)	43.55 (9.27)	0.69
FAMI-mother 's years of schooling, mean (SD)	13.61 (1.61)	13.07 (1.86)	13.57 (1.74)	0.23
FAMI-mother 's years of experience, mean (SD)	12.99 (8.63)	11.85 (7.89)	14.03 (8.98)	0.37
Early childhood certification, n (%)	16 (89)	46 (81)	68 (91)	0.32
Married, n (%)	6 (33)	15 (26)	12 (16)	0.27
Number of FAMI-mothers' own children, mean (SD)	2.50 (1.10)	2.81 (1.55)	2.61 (1.21)	0.54
FAMI's household size, mean (SD)	3.94 (1.35)	3.79 (1.57)	3.95 (1.39)	0.80
TVIP's total score (Z), mean (SD)	27.44 (12.80)	27.84 (11.74)	28.88 (10.76)	0.88
Knowledge's total score, mean (SD)	6.50 (2.04)	7.40 (1.58)	7.17 (1.75)	0.09
Depressive symptoms (CES D-10), n (%) [a]	5 (29)	5 (9)	9 (12)	0.33
Panel B: Characteristics of FAMI-mother's group	N = 18	N = 57	N = 75	
Number of children between 0 and 12 months, mean (SD)	4.22 (1.73)	4.84 (2.14)	4.89 (2.10)	0.50
Number of pregnant mothers, mean (SD)	1.39 (1.61)	2.02 (1.61)	1.95 (1.32)	0.21
Number of meetings (last month), mean (SD)	4.33 (3.33)	4.91 (3.54)	5.52 (5.59)	0.53
Number of home visits (last month), mean (SD)	15.22 (8.52)	13.55 (8.00)	12.53 (6.38)	0.43
Activities' planning time, mean (SD)	6.94 (10.00)	5.86 (4.66)	4.36 (2.48)	0.11
Panel C: Child characteristics	N = 71	N = 247	N = 324	
Age in months, mean (SD)	5.14 (3.81)	5.62 (3.22)	5.81 (3.23)	0.31
Male, n (%)	34 (48)	116 (47)	176 (54)	0.38
Birth weight, grams, mean (SD)	3151.32 (370.51)	3125.30 (478.53)	3165.25 (505.65)	0.78
Low birth weight, n (%)	5 (7)	23 (9)	24 (7)	0.74
Stunting, n (%)	12 (20)	29 (12)	28 (10)	0.41
Panel D: Household characteristics	N = 71	N = 247	N = 324	
Maternal years of schooling, mean (SD)	8.90 (3.18)	9.08 (3.19)	8.95 (3.37)	0.91
Maternal age, years, mean (SD)	26.10 (6.42)	26.56 (6.75)	26.14 (6.50)	0.71
Maternal PPVT (raw score), mean (SD) [b]	20.85 (9.59)	20.14 (7.63)	22.45 (8.76)	0.05
Father present, n (%)	52 (73)	182 (74)	231 (71)	0.78
Household in poverty, n (%) [c]	50 (70)	126 (53)	200 (63)	0.10
Quality of the Home Environment (FCI), mean (SD)	-0.23 (0.82)	-0.04 (0.90)	0.21 (0.94)	0.02

Observational sample refers to children assigned to a FAMI-mother with at least one video recording. [a] Three FAMI-mothers in the intervention arm have missing data in the CES D-10 due to incomplete baseline survey. [b] Spanish version of the Peabody Picture Vocabulary Test, a proxy for maternal IQ. [c] Indicator variable that equals one if the household's total income is below the poverty line in 2014 ($50 USD person/month). We present 2-sided p-values using *F*-tests.

WEBTABLE 7. Quality of Group Sessions by Number of Available Video Recordings

	One Video	Two Videos	Three Videos	p-value
	N = 18	N = 57	N = 75	
Quality of Group Sessions Factor Score	0.04 (1.09)	-0.21 (1.03)	0.16 (0.95)	0.15
Demonstration subscale	4.06 (1.55)	3.68 (1.43)	4.10 (1.22)	0.21
Practice subscale	4.00 (1.30)	3.75 (1.23)	4.07 (1.19)	0.46
Atmosphere subscale	3.83 (1.17)	3.66 (1.01)	4.01 (0.88)	0.17
Fun subscale	4.11 (1.05)	3.88 (1.08)	4.26 (0.82)	0.15
Sum of count variables	46.99 (34.35)	41.47 (26.28)	49.43 (28.26)	0.17

The Factor Score is presented in standard deviations. The Demonstration, Practice, Atmosphere, and Fun subscale have a 5-point rating scale (1=inadequate, 5=excellent). The sum of count variables is the sum of all count items per 30 minutes. We present 2-sided p-values using *F*-tests.

WEBTABLE 8. Quality of Group Sessions Across Rounds

	Round 1	Round 2	Round 3	p-value
Subsample of units with 3 videos	N = 75	N = 75	N = 75	
Quality of Group Sessions Factor Score	0.14 (0.92)	0.10 (0.96)	0.10 (0.93)	0.95
Demonstration subscale	50.73 (33.30)	49.42 (31.48)	48.14 (28.61)	0.66
Practice subscale	4.27 (1.36)	4.07 (1.61)	3.96 (1.61)	0.33
Atmosphere subscale	4.21 (1.35)	3.97 (1.54)	4.03 (1.43)	0.59
Fun subscale	4.05 (1.02)	4.07 (0.96)	3.90 (0.98)	0.16
Sum of count variables	4.32 (1.07)	4.29 (1.03)	4.18 (1.00)	0.37

We present 2-sided p-values using *F*-tests.

WEBTABLE 9. Association Between Quality of Group Sessions and Children's Outcomes in the Observational Sample

Dependent Variable:	Bayley-III Factor	Bayley-III Factor	Bayley-III Factor	Parental Practices (FCI)	Parental Practices (FCI)	Parental Practices (FCI)
	(1)	(2)	(3)	(4)	(5)	(6)
Independent Variable:						
Quality of Group Sessions Factor Score	0.140	0.007	0.116	0.139	-0.088	0.145
95% CI	(0.050 to 0.230)	(-0.174 to 0.188)	(-0.016 to 0.248)	(0.037 to 0.240)	(-0.239 to 0.062)	(-0.012 to 0.302)
p-value	[0.003]	[0.94]	[0.08]	[0.008]	[0.25]	[0.07]
Sample	All	Treated	Control	All	Treated	Control
Observations	585	311	274	602	322	280

Estimated coefficients are expressed in SDs of the control group. While columns (1)-(3) show the association between the quality of group sessions and the Bayley-III score, columns (4)-(6) show the association between the quality of group sessions and the FCI score. Results come from a linear regression in which we control for baseline FAMI-mother years of experience, years of education, level of depressive symptoms by CESD10, verbal ability using the Spanish version of the Peabody Picture Vocabulary Test, early childhood certificate, district fixed effects, interviewer fixed effects, and the total number of videos. The coefficient is the expected change in Bayley-III and parental investment for every additional SD increase the quality of group sessios. In columns (2) and (5) we restrict sample to treated towns and in columns (3) and (6) we restrict sample to control towns.

WEBTABLE 10. ITT Impacts and Dosage by Sample

	Full Sample					Observational Sample					Cross-model hypotheses p-value
	Point Estimate	Confidence Interval		p value	n	Point Estimate	Confidence Interval		p value	n	
		Lower Bound	Upper Bound				Lower Bound	Upper Bound			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: Treatment effect											
Bayley-III Factor	0.16	0.03	0.29	0.02	1292	0.28	0.09	0.48	0.01	585	0.18
Parental Investment (FCI)	0.34	0.20	0.47	<0.001	1331	0.21	-0.003	0.43	0.05	602	0.18
Panel B: Dosage effect											
Bayley-III Factor	0.09	0.03	0.14	0.003	599	0.11	0.04	0.18	0.005	311	0.34
Parental Investment (FCI)	0.04	-0.02	0.11	0.21	626	0.04	-0.08	0.16	0.48	322	0.95

Estimates in Panel A replicate the program's effects from Attanasio et al. (2022) on the total sample and the video sub-sample. Thus, we use the same baseline controls: child's gender, an indicator of high household wealth index, maternal PPVT score, teenage mother, an indicator of high municipality population, previous attendance to a child care center, department and interviewer fixed effects, and baseline weight-for-age and height-for-age z-scores. Table 2 differs from estimates in column 6 due to different sets of controls in each regression. While in this table, we use the same set of controls as in Attanasio et al. (2022). Table 2 controls for additional unbalanced characteristics between the total sample and the observational sample and the number of available videos. For the analyses of the dosage effect, in Panel B, the explanatory variable is the number of group sessions attended divided by 10. We control for the same baseline characteristics as above, and interviewers fixed effects. The coefficient is the expected increase in Bayley-III and parental investment for every ten additional sessions attended. The number of group sessions attended is only available for the treatment group.

Appendix 3: Disaggregated Analyses Using Subscales from the Observational Instrument

For completeness, we present our main results disaggregated by each measure of quality (i.e., the sum of count variables and four rating scales). We first present ITT effects between the treatment and control group on each measure of quality separately. All quality measures are standardized with respect to the control group for comparability across estimates. We estimated these analyses at the FAMI mother level, clustered SEs at the town level, calculated 2-sided p-values using t-tests, and controlled for the same covariates as in the main results.

We then conducted a mediation analysis using these five separate measures of quality on the impacts of the intervention on child development (Bayley-III) and parenting practices (FCI). We compared the total ITT effect on the outcome variable with the ITT effect when each mediator was included separately. We estimated these analyses at the child level, clustered SEs at the FAMI mother level, calculated 2-sided p-values using t-tests, controlled for the same covariates as before, and included child's age, sex, and tester fixed effects. We follow Preacher and Hayes's (2008) approach to test the statistical significance of the indirect effect.

We used Poisson regression to estimate the association between participant attendance to group sessions and the quality of sessions in the treatment arm only (as attendance data is only available for the treatment group). We present average marginal effects of each measurement of quality separately. We estimated these analyses at the child level, clustered SEs at the FAMI-mother level, calculated 2-sided p-values using *t*-tests, and controlled for the same covariates as before.

In ITT analysis, we found that the intervention significantly improved all the disaggregated measures of quality of group sessions. The interventions had an effect size of 2.76 SD (95% CI: 2.05 to 3.48) on the sum of count variables, 1.20 SD (95% CI: 0.88 to 1.52) on the demonstration subscale, 1.12 SD (95% CI: 0.75 to 1.48) on the practice subscale, 1.76 SD (95% CI: 1.14 to 2.37) on the atmosphere subscale, and 1.02 SD (95% confidence interval: 0.65 to 1.39) on the fun subscale (Webtable 11).

While only the demonstrations subscale (95% CI: 0.02 to 0.24) partially mediated the effect of the intervention on child development (Bayley-III) (Webtable 12); the demonstration (95% CI: 0.00 to 0.19), practice (95% CI: -0.01 to 0.16), and fun (95% CI: 0.01 to 0.13) subscales mediated the effect the intervention on parental practices (FCI) (Webtable 13). Finally, Webtable 14 shows that all five disaggregated measures of quality of group sessions are positively associated with attendance in the treatment group. One SD increase in the sum of the count variables, demonstration subscale, practice subscale, atmosphere subscale and fun subscale predicted an increase on treatment mothers' attendance of 0.83, 3.70, 3.92, 2.02, and 6.09 sessions respectively.

WEBTABLE 11. Treatment Effect on Disaggregated Measures of Quality of Group Sessions

Dependent Variable	Sum Count Variables	Demonstration Subscale	Practice Subscale	Atmosphere Subscale	Fun Subscale
	(1)	(2)	(3)	(4)	(5)
Treatment	2.76	1.20	1.12	1.76	1.02
95% CI	(2.05 to 3.48)	(0.88 to 1.52)	(0.75 to 1.48)	(1.14 to 2.37)	(0.65 to 1.39)
p-value	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Observations	150	150	150	150	150

This table presents intention-to-treat (ITT) effects between the treatment and control group on the disaggregated quality measures of group sessions. Estimated coefficients are expressed in SDs of the control group. Estimates are at the FAMI mother level and the sample includes all FAMI mothers with at least one video. Estimates controlled for baseline FAMI mother's years of experience, years of education, level of depressive symptoms by CESD10, verbal ability using the Spanish version of the Peabody Picture Vocabulary Test, early childhood certificate, district fixed effects, and the total number of videos. The p-values are 2-tailed conventional p-values. CIs were constructed by using conventional critical values for individual hypotheses. The intracluster correlation coefficient for the primary outcome (quality of group sessions factor score) was 0.24.

WEBTABLE 12. Mediation Analysis with Disaggregated Measures of Quality of Group Sessions on the Bayley--III

Dependent Variable	Bayley-III	Bayley-III	Bayley-III	Bayley-III	Bayley-III	Bayley-III
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.27	0.17	0.14	0.22	0.22	0.22
95% CI	(0.05 to 0.49)	(-0.08 to 0.42)	(-0.10 to 0.39)	(-0.03 to 0.47)	(-0.03 to 0.47)	(-0.01 to 0.45)
p-value	[0.02]	[0.17]	[0.25]	[0.09]	[0.09]	[0.07]
Sum Count Variables		0.03				
95% CI		(-0.03 to 0.10)				
p-value		[0.32]				
Demonstration Subscale			0.12			
95% CI			(0.03 to 0.21)			
p-value			[0.01]			
Practice Subscale				0.06		
95% CI				(-0.05 to 0.16)		
p-value				[0.27]		
Atmosphere Subscale					0.04	
95% CI					(-0.05 to 0.12)	
p-value					[0.41]	
Fun Subscale						0.08
95% CI						(-0.02 to 0.17)
p-value						[0.12]
Observations	585	585	585	585	585	585
Indirect Effect p-value		0.31	0.02	0.28	0.42	0.13

Column 1 presents intention-to-treat (ITT) effects between the treatment and control group on the Bayley-III composite score. Columns 2 to 6 show ITT effects after including each disaggregated measurement of quality separately. Both estimated coefficients and independent variables are expressed in SDs of the control group. Estimates are at the child level and the sample includes all FAMI mothers with at least one video. Estimates controlled for baseline FAMI mother's years of experience, years of education, level of depressive symptoms by CESD10, verbal ability using the Spanish version of the Peabody Picture Vocabulary Test, early childhood certificate, district fixed effects, the total number of videos, and include interviewer fixed effects. The p-values are 2-tailed conventional p-values. CIs were constructed by using conventional critical values for individual hypotheses. Missing data in control variables were replaced by sample means. To test the statistical significance of the indirect effect we follow Preacher and Hayes (2008)[16] and bootstrapped the indirect effect with 2,000 replications to compute the p-value.

WEBTABLE 13. Mediation Analysis with Disaggregated Measures of Quality of Group Sessions on Parental Investment

Dependent Variable	FCI	FCI	FCI	FCI	FCI	FCI
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.26	0.27	0.17	0.19	0.16	0.19
95% CI	(0.08 to 0.45)	(0.02 to 0.51)	(-0.03 to 0.37)	(-0.01 to 0.39)	(-0.05 to 0.38)	(0.01 to 0.38)
p-value	[0.006]	[0.03]	[0.10]	[0.06]	[0.13]	[0.04]
Sum Count Variables		-0.00				
95% CI		(-0.05 to 0.05)				
p-value		[0.97]				
Demonstration Subscale			0.08			
95% CI			(-0.01 to 0.17)			
p-value			[0.06]			
Practice Subscale				0.08		
95% CI				(-0.01 to 0.17)		
p-value				[0.09]		
Atmosphere Subscale					0.07	
95% CI					(-0.01 to 0.14)	
p-value					[0.09]	
Fun Subscale						0.10
95% CI						(0.02 to 0.19)
p-value						[0.02]
Observations	602	602	602	602	602	602
Indirect Effect p-value		0.97	0.06	0.09	0.11	0.02

Column 1 presents intention-to-treat (ITT) effects between the treatment and control group on the FCI. Columns 2 to 6 show ITT effects after including each disaggregated measurement of quality separately. Both estimated coefficients and independent variables are expressed in SDs of the control group. Estimates are at the child level and the sample includes all FAMI mothers with at least one video. Estimates controlled for baseline FAMI mother's years of experience, years of education, level of depressive symptoms by CESD10, verbal ability using the Spanish version of the Peabody Picture Vocabulary Test, early childhood certificate, district fixed effects, the total number of videos, and include interviewer fixed effects. The p-values are 2-tailed conventional p-values. CIs were constructed by using conventional critical values for individual hypotheses. Missing data in control variables were replaced by sample means. To test the statistical significance of the indirect effect we follow Preacher and Hayes (2008)[16] and bootstrapped the indirect effect with 2,000 replications to compute the p-value.

WEBTABLE 14. Association Between Disaggregated Measures of Quality of Group Sessions and Attendance

Dependent Variable	Number of Groups Sessions Attended	Number of Groups Sessions Attended	Number of Groups Sessions Attended	Number of Groups Sessions Attended	Number of Groups Sessions Attended
	(1)	(2)	(3)	(4)	(5)
Sum Count Variables	0.83				
95% CI	(-0.07 to 1.73)				
p-value	[0.07]				
Demonstration Subscale		3.70			
95% CI		(-0.01 to 7.42)			
p-value		[0.05]			
Practice Subscale			3.95		
95% CI			(0.11 to 7.79)		
p-value			[0.05]		
Atmosphere Subscale				2.03	
95% CI				(-0.41 to 4.47)	
p-value				[0.10]	
Fun Subscale					6.11
95% CI					(2.86 to 9.37)
p-value					[<0.001]
Observations	347	347	347	347	347

This table presents the average marginal effects of a Poisson regression to estimate the association between participant attendance to group sessions and the disaggregated measures of quality of sessions. Estimated coefficients are expressed in number of sessions for a SD increase of each measure of quality. Estimates are at the child level and the sample includes all FAMI mothers with at least one video in the intervention group, as we do not have information on attendance for the control group. In the treatment group, 101/347 (29.1%) attended zero sessions, the median number of sessions attended was 17; the maximum number of sessions was 42. Estimates controlled for baseline FAMI mother's years of experience, years of education, level of depressive symptoms by CESD10, verbal ability using the Spanish version of the Peabody Picture Vocabulary Test, early childhood certificate, district fixed effects, and the total number of videos. The p-values are 2-tailed conventional p-values. CIs were constructed by using conventional critical values for individual hypotheses.

Appendix 4: Observation Instrument to Measure the Quality of a Group Early Childhood Development Parenting Session

Facilitator_____ **Town:** _____ **Date:**_____

Number of participants in the session:

Women_____ Children: Infants_____
Men_____ Sitting/crawling____
Others _____ Walking_____

Time:

a. Start: _____
b. End: _____
c. Total duration of session: _____

Count Variables

	Tally # of times	TOTAL
Facilitator praises mothers		
Facilitator praises babies		
Facilitator praises the group		
Facilitator says positive things about the children to mothers		
Facilitator asks an open question		
Mothers make verbal contribution		
Facilitator expands on what a mother says		

RATING SCALES

Each item is scored as inadequate (score=1), adequate (score=3) or excellent (score=5).
The subscale score is the mean score for all items in the subscale

Demonstration Subscale

1. Inadequate	2.	3. Adequate	4.	5. Excellent
☐ 1.1. Facilitator rarely *demonstrates* an activity or play. ☐ 2.1. Facilitator gives unclear		☐ 1.3 Facilitator occasionally demonstrates an activity or play. ☐ 2.3. Facilitator gives partially clear *demonstrations*		☐ 1.5. Facilitator frequently demonstrates an activity or play.

Practice Subscale

1. Inadequate	2.	3. Adequate	4.	5. Excellent
☐ 1.1. Few mothers *practice* the activity or play. ☐ 2.1. Facilitator rarely supports mothers while practicing (also score 1 if mothers do not practice the activities) ☐ 3.1. Facilitator gives very little or no time to practice the activities or games		☐ 1.3. Some mothers *practice* the activities activity or play. ☐ 2.2. Facilitator occasionally supports mothers while practicing ☐ 3.3. Facilitator gives not enough time to practice the activities or games		☐ 1.5. Most of the mothers *practice* the activities or plays. ☐ 2.5. Facilitator frequently supports mothers while practicing ☐ 3.5. Facilitator gives enough time to practice the activities or games.
PRACTICE SUBSCALE SCORE = (mean of 3 items)				

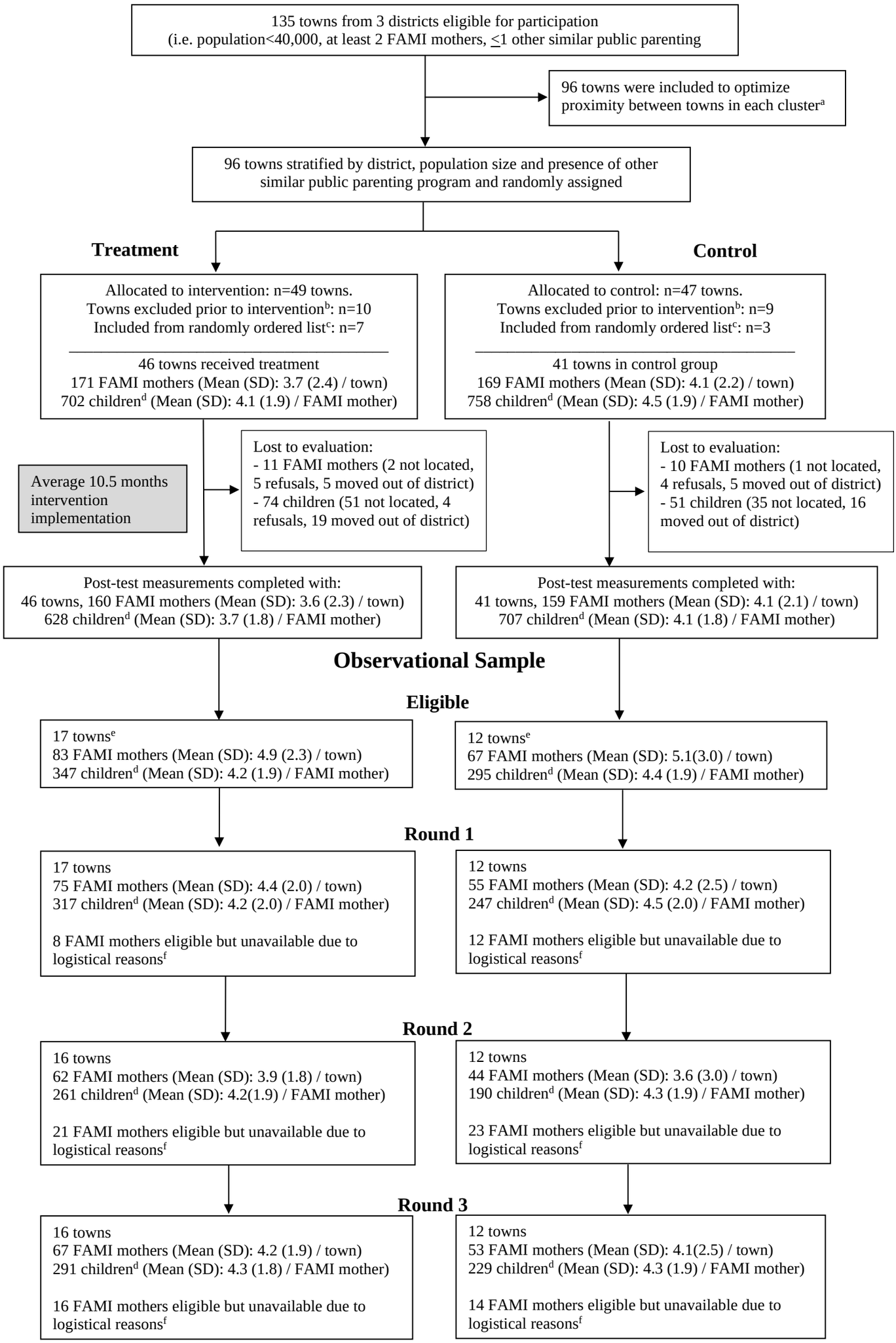
Atmosphere Subscale

(Seating, proximity, position; comfort of the mothers, names)

1. Inadequate	2.	3. Adequate	4.	5. Excellent
☐ 1.1. Most of the moms are sitting in rows or in a way where they cannot see each other during the session ☐ 2.1. Facilitator stands in front of the mothers most of the time while mothers are sitting		☐ 1.3. Some of the moms are sitting behind others most of the time of the session. ☐ 2.3. Facilitator sometimes stands in front of the mothers and sometimes sits or crouches at the same height of the babies and mothers		☐ 1.5. All of the moms are sit in a way where all of them can see each other ☐ 2.5. Facilitator spends most of the time of the session at the same height of the mothers and babies
☐ 3.1. Facilitator rarely gives positive affirmations to mothers and babies ☐ 4.1. Facilitatorinvolves few of the mothers ☐ 5.1. Facilitator rarely uses mothers' names ☐ 6.1. Facilitator rarely uses babies' names		☐ 3.3. Facilitator occasionally gives positive affirmations to mothers and babies ☐ 4.3. Facilitator involves some of the mothers. ☐ 5.3. Facilitator occasionally uses mothers' names ☐ 6.3. Facilitator occasionally uses babies' names		☐ 3.5. Facilitator frequently gives positive affirmations to mothers and babies. ☐ 4.5. Facilitator makes sure to involve mos Facilitator t of the mothers ☐ 5.5. Facilitator frequently uses mothers' names ☐ 6.5. Facilitator frequently uses babies' names

Fun Subscale

1. Inadequate	2.	3. Adequate	4.	5. Excellent
☐ 1.1. There are few toys for all the babies ☐ 2.1. Toys are available a little of the time		☐ 1.3. There are some toys for some babies. ☐ 2.3. Toys are available for some of the time.		☐ 1.5.There are enough toys for all of the babies ☐ 2.5. Toys are available for most of the time
☐ 2.1. Few of the mothers look like they are having fun and rarely laugh or play ☐ 3.1. Few of the babies look like they are having fun and rarely laugh or play ☐ N.A if babies are sleeping most of the time, score item 3 as N/A.		☐ 2.3. Some of the mothers look like they are having fun and occasionally laugh or play ☐ 3.3. Some of the babies look like they are having fun and occasionally laugh or play		☐ 2.3. Most of the mothers look like they are having fun and frequently laugh or play ☐ 3.5. Most of the babies look like they are having fun and frequently laugh or play
☐ 5.1. Facilitator rarely laughs and enjoys the activities and discussions with mothers and babies		☐ 5.3. Facilitator occasionally laughs and enjoys the activities and discussions with mothers and babies		☐ 5.5. Facilitator frequently laughs and enjoys the activities and discussions with mothers and babies
FUN SUBSCALE SCORE = (mean of 5 items)				



[a] According to power calculations, only 96 towns were needed for the study. We excluded 39 municipalities because the remaining 96 allowed shorter routes for training and supervision of FAMI-mothers. [b] Once in the field for data collection, we realized some towns did not have any FAMI mothers as they had made the transition to other public parenting programs (Modalidad Familiar or MF). [c] Towns in the list of 39 towns excluded initially from the sample, were randomly ranked, and used as replacements. However, we did not have enough replacement towns in all randomization strata. [d] Children in the evaluation sample only. [e] Towns included in the video sample, towns were selected for logistical reasons to maximize observations given time and resource constraints. [f] Logistical reasons include (i) being unable to go to all FAMI mothers in a town due to time constraints; (ii) FAMI mothers were running concurrent group sessions being unable to videotape both; (iii) technical issues with the video's audio recording. In the video sample, 25 children in treatment arm (20 not located, 5 moved out of district) and 15 in the control arm (13 not located, 2 moved out of district) were lost to follow-up.