

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

abstract

OBJECTIVES: We conducted a cluster-randomized trial of an enhancement to an existing parenting program in rural Colombia (called the Family, Women, and Infancy Program [FAMI]), 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 aged <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 intention-to-treat analyses, we found significant benefits of intervention for the observed quality of group sessions (1.67 SD [95% confidence interval, 1.23–2.11]). An SD increase in session quality predicted an increase in treatment mothers' attendance of 4.68 sessions (95% confidence interval, 1.37–7.98). 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.



^aUniversidad de los Andes and CEDE, Bogota, Colombia, and ^bSchool of Human and Behavioral Sciences, Bangor University, Bangor, Wales, United Kingdom

Dr 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 Gomez 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 Perez-Cardona conducted the data analyses and contributed to the initial draft of the manuscript; Dr 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; and all authors approved the final manuscript as submitted and agreed to be accountable for all aspects of the work.

The trial is registered at ISCTRN, #ISRCTN93757590 (doi.org/10.1186/ISRCTN93757590). Deidentified individual participant data (including data dictionaries) will be made available through the World Bank Open Data Depository within a year after publication.

DOI: <https://doi.org/10.1542/peds.2023-060221J>

Accepted for publication Feb 17, 2023

Address correspondence to Raquel Bernal, PhD, Universidad de los Andes, Calle 19A, No. 1-37 Este. Bogotá 111711, Colombia. E-mail: rbernal@uniandes.edu.co

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

Copyright © 2023 by the American Academy of Pediatrics

FUNDING: Funded by Grand Challenges Canada Grant #0462-03-10 and Fundacion Exitó. The funders had no role in the design or conduct of this study.

CONFLICT OF INTEREST DISCLOSURES: The authors have indicated they have no conflicts of interest relevant to this article to disclose.

There is a strong evidence base showing early childhood parenting programs benefit child development in low- and middle-income countries (LMICs).¹ The challenge is to extend the reach of these programs.^{2,3} An essential component of scaling evidence-based interventions is to 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. Although 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 semiurban and rural areas of Colombia, the Family, Women, and Infancy Program (FAMI) provides training and support for economically disadvantaged pregnant women and parents of children up to 2 years of age. The FAMI 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 Up⁹ 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-randomized 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 (eg, demonstration, practice, positive feedback) and the interactive 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-randomized 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 (Fig 1). The mean (SD) beneficiaries per FAMI mother were 11.6 (2.8), comprising 9.5 (2.9) children aged <2 years and 2.1 (1.7) pregnant women. Within each unit, we enrolled all children aged <12 months at baseline in the evaluation sample to give a total of 1456 children (Fig 1). We selected children aged <12 months to maximize the potential time of exposure to our intervention before children outgrew the FAMI at age 2. At posttest, 319 (93.8%) FAMI mothers (160 intervention, 159 control) and 1335 children (91.4%) (628 intervention, 707 control) were evaluated (Fig 1).

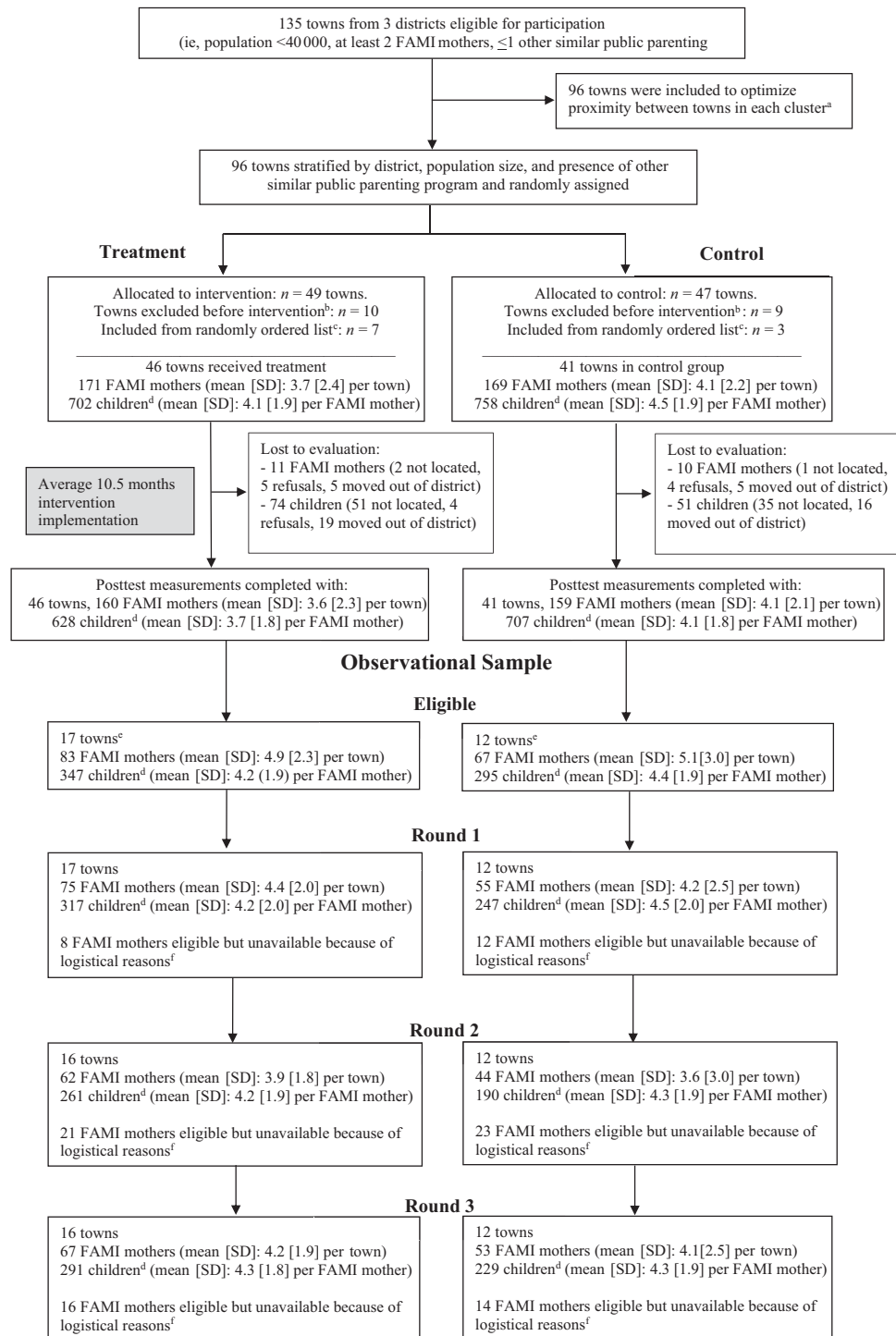
For this study, we selected a subsample of towns to participate in the assessment of the quality of the group sessions through observation. The subsample was not randomly selected; rather, it was selected for logistical reasons and includes towns with more FAMI mothers and those that permitted a shorter 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, 295 control) were included in this subsample.

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. Posttest measurements were collected between April and July 2016. Written informed consent to participate in the study was obtained from FAMI mothers and beneficiary mothers by survey staff at baseline, before the observational assessments, and at endline. No participants refused to participate. The study protocol was approved by the Universidad de los Andes ethics committee (No. 287/2014) and the University College London ethics committee (No. 2168/011). The trial registration number is ISRCTN93757590.

Intervention

The enhancement to the FAMI in the intervention group consisted of 4 main components:

1. 2 structured curricula: 1 for home visiting and 1 for group sessions;



^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 because they had made the transition to other public parenting programs (Modalidad Familiar). ^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 because of 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.

FIGURE 1
Trial profile.

2. developmentally appropriate and low-cost play materials (eg, picture books, puzzles, homemade 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 in Supplemental Information. 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, IQartil, 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 mothers' use of praise and efforts to promote mothers' participation (7 items) with 4 rating

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 infants. 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 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 infants. FAMI mother gives positive affirmations to beneficiary mothers and infants. FAMI mother involves beneficiary mothers (using a collaborative and participatory approach). FAMI mother uses beneficiary mothers' names. FAMI mother uses infants' 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, infants, and FAMI mother	Sufficient toys for all the infants Toys are available for infants for the entire session. Beneficiary mothers enjoy the session (show positive affect). Infants enjoy the session (laugh, smile, play). FAMI mother enjoys the session (shows positive affect).	5-point rating scale: 1 = inadequate 3 = adequate 5 = excellent	Mean score of all 5 items

See Appendix 3 in Supplemental Information for the observational tool used in this study.

scales: (1) demonstration (2 items), (2) practice (3 items), (3) atmosphere (7 items), and (4) fun and enjoyment (5 items) (Table 1). Further details of the observational instruments used are given in Appendix 3 in Supplemental Information. The categories were designed to include the core delivery components of Reach Up, adapted for the group setting and suitable for use with video recordings. All videos were coded over a 3-month period after posttest 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 interobserver reliabilities. We randomly selected 15% of videos from each round of filming and conducted ongoing interobserver reliabilities every week. Interobserver reliabilities (intraclass correlation coefficients) were mean (SD) = 0.93 (0.06), with a range of 0.86–1.0 (Supplemental Table 4). All subscales had good internal consistency (Cronbach's α mean [SD] = 0.85 [0.09], with a range of 0.69–0.97) (Supplemental Table 5).

We also examined whether quality of the group sessions was associated with child development and parenting practices, 2 outcomes that showed significant benefits in our impact evaluation.¹¹ Child development was measured at posttest only using the Bayley Scales of Infant and Toddler Development, Third 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 posttest 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 3 days.

Randomization and Blinding

Towns were randomized 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 because 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 intraclass 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 0.70 SD 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 (ie, sum of the count variables and 4 rating scales), prorated to 30 minutes of observation, across the number of observations available for each FAMI mother. Exploratory factor analysis gave 1 factor (Supplemental Table 6); 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, the baseline FAMI mother's 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 tested the statistical significance of the indirect effect (IE) 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 presented 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 parenting practices in 29 towns (Fig 1). Losses were balanced across groups (Supplemental Table 7). The

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

	Intervention	Control	<i>P</i>
Baseline characteristics			
Panel A: FAMI mother characteristics			
	<i>N</i> = 83	<i>N</i> = 67	
FAMI mother's age, <i>y</i> , mean (SD)	43.66 (9.52)	42.00 (9.63)	.36
FAMI mother's <i>y</i> of schooling, mean (SD)	13.24 (1.60)	13.57 (1.98)	.48
FAMI mother's <i>y</i> of experience, mean (SD)	12.62 (8.33)	13.64 (8.84)	.47
Early childhood certification, <i>n</i> (%)	69 (83)	61 (91)	.26
PPVT (raw score), mean (SD)	30.69 (11.07)	25.37 (11.03)	.09
Depressive symptoms (CES D-10), <i>n</i> (%) ^a	14 (18)	5 (7)	.06
Panel B: characteristics of FAMI mother's group			
	<i>N</i> = 83	<i>N</i> = 67	
Number of children aged between 0 and 12 mo, mean (SD)	4.75 (1.87)	4.85 (2.31)	.87
Number of pregnant mothers, mean (SD)	1.82 (1.36)	2.01 (1.61)	.53
Number of meetings (last mo), mean (SD)	5.72 (5.57)	4.45 (3.13)	.37
Number of home visits (last mo), mean (SD)	11.98 (5.72)	14.79 (8.65)	.24
Activities' planning time (h per wk), mean (SD)	4.72 (3.20)	5.84 (6.18)	.29
Panel C: child characteristics			
	<i>N</i> = 347	<i>N</i> = 295	
Age in mo, mean (SD)	5.88 (3.28)	5.41 (3.31)	.15
Male, <i>n</i> (%)	173 (50)	153 (52)	.65
Low birth weight, <i>n</i> (%)	30 (9)	22 (7)	.63
Stunting, <i>n</i> (%)	27 (8)	42 (16)	.15
Panel D: household characteristics			
	<i>N</i> = 347	<i>N</i> = 295	
Maternal <i>y</i> of schooling, mean (SD)	8.88 (3.41)	9.14 (3.11)	.50
Maternal age, <i>y</i> , mean (SD)	25.89 (6.76)	26.78 (6.34)	.13
Maternal PPVT (raw score), mean (SD) ^b	23.22 (8.82)	19.22 (7.56)	.04
Father present, <i>n</i> (%)	243 (70)	222 (75)	.25
Household in poverty, <i>n</i> (%) ^c	202 (59)	174 (62)	.73
Quality of the home environment (FCI), mean (SD)	0.12 (0.94)	0.00 (0.90)	.46
Follow-up characteristics			
Panel E: video observations characteristics			
	<i>N</i> = 83	<i>N</i> = 67	
Number of observations			.10
1 video, <i>n</i> (%)	9 (11)	9 (13)	
2 videos, <i>n</i> (%)	26 (31)	31 (46)	
3 videos, <i>n</i> (%)	48 (58)	27 (40)	
Number of different child age ranges present (0–5 mo, 6–11 mo, 12–24 mo), median (SD)	1 (0.63)	2 (0.65)	.12
Number of children present, mean (SD)	3.55 (1.21)	3.93 (1.91)	.39
Duration of observations (min), mean (SD)	36.21 (10.82)	36.96 (10.50)	.72
Panel F: quality of sessions			
	<i>N</i> = 83	<i>N</i> = 67	
Sum count variables, mean (SD)	64.11 (23.96)	23.83 (13.86)	<.001
Mean demonstration over all observations, mean (SD)	4.63 (0.73)	3.06 (1.44)	<.001
Mean practice over all observations, mean (SD)	4.57 (0.70)	3.16 (1.28)	<.001
Mean atmosphere over all observations, mean (SD)	4.48 (0.62)	3.08 (0.76)	<.001
Mean fun over all observations, mean (SD)	4.55 (0.60)	3.54 (1.04)	<.001

CES D-10, Center for Epidemiologic Studies Depression Scale; PPVT, Peabody Picture Vocabulary Test.

^a Three FAMI mothers in the intervention arm have missing data in the Center for Epidemiologic Studies Depression Scale because of incomplete baseline survey.

^b Spanish version of the Peabody Picture Vocabulary Test, a proxy for maternal IQ.

^c Indicator variable that equals 1 if the household's total income is below the poverty line in 2014 (US\$50 person per month). We present 2-sided *P* values in column 3. Although for continuous and indicator variables we calculated *P* values using *t* tests, for categorical variables (with >2 categories), we used a Pearson's χ^2 test.

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* = .01) and higher maternal education (9.00 years versus 8.62, *P* = .03) in the observational sample (Supplemental Table 8). We control for these differences in the analyses. Eighteen FAMI mothers were video recorded once,

57 were recorded twice, and 75 were recorded 3 times, with a similar number of video recordings available per FAMI mother across study groups (Table 2), and few differences in the numbers of video observations conducted (Supplemental Table 9). 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 (Supplemental Tables 10 and 11).

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

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 Group Sessions Attended
	(1)	(2)	(3)	(4)	(5)	(6)
Independent variables						
Treatment	1.67	0.27	0.15	0.26	0.14	—
95% CI	(1.23–2.11)	(0.05–0.49)	(–0.10 to 0.41)	(0.08–0.45)	(–0.08 to 0.35)	—
<i>P</i>	[<.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)	—	(0.01–0.19)	(1.37–7.98)
<i>P</i>	—	—	[.07]	—	[.04]	[.006]
Observations	150	585	585	602	602	347
IE	—	—	0.12	—	0.13	—
95% CI	—	—	(–0.01 to 0.25)	—	(0.00–0.25)	—
<i>P</i>	—	—	[.08]	—	[.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. Although in columns 1 to 5, the sample includes all FAMI mothers with at least 1 video, in column 6, we restricted the sample to the intervention group, because we do not have information on attendance for the control group. In the treatment group, 101 of 347 (29.1%) attended 0 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 the baseline FAMI mother's years of experience, years of education, level of depressive symptoms by the Center for Epidemiologic Studies Depression Scale, 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 (ie, replacement with sample median and regression imputation). Results are robust to these alternative approaches. To test the statistical significance of the IE, we follow Preacher and Hayes (2008)¹⁵ and bootstrapped the IE with 2000 replications to compute the *P* value. Results are robust to using the test of the joint significance, as described by MacKinnon et al (2002).¹⁶ —, variable or statistic not included in the regression.

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 ES of 1.67 SD (95% confidence interval [CI], 1.23–2.11) (Table 3). The intervention had an effect of 0.27 SD (95% CI, 0.05–0.49) on child development (Bayley-III composite score) and 0.26 SD (95% CI, 0.08–0.45) 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 (IE, 0.12; 95% CI, –0.01 to 0.25) and parental practices (IE, 0.13; 95% CI, 0.00–0.25) (Table 3). When analyzing treatment and control groups separately, associations between session quality and child and parent outcomes were evident in the

control group only (Supplemental Tables 12).

Finally, an SD increase in the quality of the group parenting sessions predicted an increase on treatment mothers' attendance of 4.68 sessions (95% CI, 1.37–7.98) (Table 3). Mothers' attendance predicted child and maternal outcomes: for every 10 group sessions attended, child Bayley test scores increased by 0.10 SD and parental practices increased by 0.04 SD (Supplemental Table 13).

In Appendix 2 in Supplemental Information, 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 programs¹⁸ and group parenting sessions.¹⁹ In this study, video recordings of group sessions were conducted after ~6 months of implementation, and even within

this relatively short time frame, 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 because of insufficient variability within the treatment group (>80% of intervention FAMI mothers scored >4 out of a maximum of 5 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 were provided by 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-randomized trial with a treatment and control group. 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 maximize the likelihood that the quality score was an accurate indicator of quality across groups. The study also had several limitations. Because of cost constraints, we were unable to randomly sample FAMI mothers to participate in this nested study; however, the subsample was

reasonably representative of the full sample. As a result of logistical and technical challenges, only half of the sample had all 3 video recordings; however, there were few differences in FAMI mothers' characteristics and quality of implementation among those with 1, 2, or 3 videos. Likewise, session quality for those with 3 videos was similar across rounds. Although the FAMI 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 training and support provided to frontline staff delivering ECD group-based parenting programs in LMICs and improving program delivery.

ABBREVIATIONS

Bayley-III: Bayley Scales of Infant and Toddler Development, Third Edition
CI: confidence interval
ECD: early childhood development
ES: effect size
FAMI: Family, Women, and Infancy Program
FCI: Family Care Indicators
IE: indirect effect
ITT: intention-to-treat
LMIC: low- and middle-income countries

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(5):e1003602
2. Britto PR, Lye SJ, Proulx K, et al. Early Childhood Development Interventions Review Group, for the Lancet Early Childhood Development Series Steering Committee. Nurturing care: promoting early childhood development. *Lancet.* 2017;389(10064):91–102
 3. Richter LM, Daelmans B, Lombardi J, et al. Paper 3 Working Group and the Lancet Early Childhood Development Series Steering Committee. Investing in the foundation of sustainable development: pathways to scale up for early childhood development. *Lancet.* 2017;389(10064): 103–118
 4. Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am J Community Psychol.* 2008;41(3-4):327–350
 5. Schodt S, Parr J, Araujo MC, Rubio-Codina M. *Measuring the Quality of Home-Visiting Services.* Washington, DC: Inter-American 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
 7. Araujo MC, Dormal M, Rubio-Codina M. 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; 2018
 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):343–362
 9. Walker S, Chang S, Powell C, Smith J, Baker-Henningham H. Reach Up Team. The Reach Up Early Childhood Parenting Program: origins, content and implementation. *Zero 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
 11. Attanasio O, Baker-Henningham H, Bernal R, Meghir C, Pineda D, Rubio-Codina M. Early stimulation and nutrition: the impacts of a scalable intervention. *J Eur Econ Assoc.* 2022;20(4):1395–1432
 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(2):225–238
 13. Bayley N. *Bayley Scales of Infant and Toddler Development*, 3rd ed. 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(4):472–486
 15. Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav Res Methods.* 2008;40(3):879–891
 16. MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. *Psychol Methods.* 2002;7(1):83–104
 17. Aboud FE, Prado EL. Measuring the implementation of early childhood development programs. *Ann N Y Acad Sci.* 2018;1419(1):249–263
 18. Yousafzai AK, Rasheed MA, Siyal S. Integration of parenting and nutrition interventions in a community health program in Pakistan: an implementation evaluation. *Ann N Y Acad Sci.* 2018; 1419(1):160–178
 19. Luoto JE, Lopez Garcia I, Aboud FE, et al. An implementation evaluation of a group-based parenting intervention to promote early childhood development in rural Kenya. *Front Public Health.* 2021;9:653106
 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):120–140
 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
 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 N Y Acad Sci.* 2018;1419(1):218–229