

Rubio-Codina, Marta; Dormal, Marta; Araujo, M. Caridad

Working Paper

Measuring quality and characterizing Cuna Mas home visits: Validation of the HOVRS-A+2 in Peru and of a short checklist for use at-scale

IDB Working Paper Series, No. IDB-WP-930

Provided in Cooperation with:

Inter-American Development Bank (IDB), Washington, DC

Suggested Citation: Rubio-Codina, Marta; Dormal, Marta; Araujo, M. Caridad (2018) : Measuring quality and characterizing Cuna Mas home visits: Validation of the HOVRS-A+2 in Peru and of a short checklist for use at-scale, IDB Working Paper Series, No. IDB-WP-930, Inter-American Development Bank (IDB), Washington, DC, <https://doi.org/10.18235/0001269>

This Version is available at:

<https://hdl.handle.net/10419/208134>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>

IDB WORKING PAPER SERIES N° IDB-WP-930

Measuring Quality and Characterizing Cuna Mas Home Visits:

Validation of the HOVRS-A+2 in Peru and of a Short
Checklist for Use At-Scale

Marta Rubio-Codina
Marta Dormal
M. Caridad Araujo

Measuring Quality and Characterizing Cuna Mas Home Visits:

Validation of the HOVRS-A+2 in Peru and of a Short
Checklist for Use At-Scale

Marta Rubio-Codina
Marta Dormal
M. Caridad Araujo

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library
Rubio-Codina, Marta.

Measuring quality and characterizing Cuna Mas home visits: validation of the HOVRSA +2 in Peru and of a short checklist for use at-scale / Marta Rubio-Codina, Marta Dormal, Maria Caridad Araujo.

p. cm. — (IDB Working Paper ; 930)

Includes bibliographic references.

1. Child care services-Government policy-Peru. 2. Child development-Government policy-Peru. 3. Child welfare-Government policy-Peru. I. Dormal, Marta. II. Araujo, Maria Caridad. III. Inter-American Development Bank. Social Protection and Health Division. IV. Title. V. Series.

IDB-WP-930

<http://www.iadb.org>

Copyright © 2018 Inter-American Development Bank. This work is licensed under a Creative Commons IGO 3.0 Attribution-NonCommercial-NoDerivatives (CC-IGO BY-NC-ND 3.0 IGO) license (<http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) and may be reproduced with attribution to the IDB and for any non-commercial purpose, as provided below. No derivative work is allowed.

Any dispute related to the use of the works of the IDB that cannot be settled amicably shall be submitted to arbitration pursuant to the UNCITRAL rules. The use of the IDB's name for any purpose other than for attribution, and the use of IDB's logo shall be subject to a separate written license agreement between the IDB and the user and is not authorized as part of this CC-IGO license.

Following a peer review process, and with previous written consent by the Inter-American Development Bank (IDB), a revised version of this work may also be reproduced in any academic journal, including those indexed by the American Economic Association's EconLit, provided that the IDB is credited and that the author(s) receive no income from the publication. Therefore, the restriction to receive income from such publication shall only extend to the publication's author(s). With regard to such restriction, in case of any inconsistency between the Creative Commons IGO 3.0 Attribution-NonCommercial-NoDerivatives license and these statements, the latter shall prevail.

Note that link provided above includes additional terms and conditions of the license.

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.



scl-sph@iadb.org

www.iadb.org/SocialProtection

Measuring Quality and Characterizing *Cuna Mas* Home Visits: Validation of the HOVRS-A+2 in Peru and of a Short Checklist for Use At-scale

Marta Rubio-Codinaⁱ, Marta Dormalⁱ, and M. Caridad Araujoⁱ

Abstract

As home visiting programs continue to expand, it is crucial to develop cost-effective methods to monitor their quality that are feasible at-scale. This paper compares two instruments widely used among home visiting services in the US to a simpler checklist in the context of Peru's *Cuna Mas* Program. The paper aims to document the structure, content and level of process quality of the *Cuna Mas* home visits. Its main contribution is to empirically identify a subset of twenty-eight items that focus on critical aspects of quality and that are feasible to collect on a routine-basis as part of program monitoring efforts.

JEL codes: I12, J24, O15

Keywords: home visiting, quality, child development, monitoring, HOVRS, HVCCF, scalability.

ⁱ Social Protection and Health Division, Inter-American Development Bank, Washington DC, USA.

We would like to acknowledge the valuable contributions of Jenny Encina and Sara Schodt, at different stages of this project. We are also grateful to the firms Instituto Cuanto in Lima and Habitus Investigaciones in Quito, for their work in data collection and coding of the videos, and to Caridian Niama for leading the training of the video recording. Special thanks to Lori Roggman who generously shared her knowledge and experience in the field and provided valuable suggestions in this project's design, implementation and analysis phases. We would also like to thank Kim Boller, Sally Grantham-McGregor, Norbert Schady, and all participants to the workshop that initiated this work in January 2015 for valuable exchanges and inputs. Lastly, this study would have not been possible without the support and commitment of the program staff of *Cuna Mas*, and the technical teams at the Ministries of Social Development and Inclusion and of Economics and Finance in Peru.

1. Introduction

Home visiting has grown in popularity in Latin America and the Caribbean (LAC) and in other regions of the world. Evidence coming mostly from efficacy trials and small-scale rigorously evaluated pilots has shown that home visiting programs can have high and sustained impacts on child development (Britto et al. 2016). An important policy question is whether it is feasible to scale-up this type of interventions whilst maintaining fidelity so that at-scale programs lead to similarly large and sustained benefits as those observed in smaller experiences.

A key to ensuring fidelity, in particular when thinking about delivering home visits at-scale, is to understand what attributes of service delivery are most critical to generate and improve over time. By critical, we refer to aspects that are necessary to achieve those behavioral changes in caregivers that result in more and better-quality adult-child interactions and in higher levels of psychosocial stimulation in the home. There is a relatively small, but growing literature on measuring the quality of home visits and understanding how it relates to child development outcomes. Schodt et al. (2015) offer a recent review.

Paulsell et al. (2010) propose three dimensions of home visiting quality: dosage, content, and relationships. Dosage and content can be understood as *structural* elements of quality and can therefore be easily measured through a variety of checklist-type tools. These tools contain information on whether the visit was completed, its length and location, the participants, and the activities carried out or material covered, amongst others. They are most often filled out by the visitor or a trained observer at the end of a visit (Barrett, Zaveri & Strong 2010; Paulsell et al. 2010). A number of such checklist-type tools have been used in the study of several large-scale programs in the United States, such as the Early Head Start Family and Child Experiences (Baby FACES) and the Partnering with Families for Early Learning (PFEL).

Relationships, on the other hand, and the quantity and nature of the interactions that occur during the home visit are a *process* element of quality. Process variables focus on more dynamic aspects of quality such as *how* the intervention content is implemented, as well as on the relationships between the home visitor, the caregiver, and the child (Thomasen and La Paro 2009).¹ These interactive processes are at the heart of a programs' successful service delivery (Paulsell et al. 2010): the home visitor's role involves the complex task of transmitting parenting and developmental skills to caregivers with the aim of changing her day-to-day interactions with her child. Process quality is more intricate and lengthy to record and quantify in a standardized manner than structural quality, and requires well-trained observers.

¹ There is some debate as to whether content (and fidelity to content in the delivery) is an element of structural or of process quality. In this work we will refer to content as structural quality.

A number of instruments have been developed to measure process quality. By and large, these instruments place emphasis on measuring the overall quality of the home visit and of the home visitor's relationship with the caregiver. Nonetheless, they differ on a number of dimensions such as the specific areas they assess; the scoring protocols; on whether they can—or should preferably—be administered by live observation, video or both; and on whether they can also be used as tools for supervision and mentoring. A common element to their administration and scoring is the requirement of rigorously trained observers, even if their background may differ.

Schodt et al. (2015) provides a detailed description of the instruments commonly used to measure home visit quality, alongside past experiences of their administration. The following are those for which there exists a larger body of published evidence of their use:

- The *Home Visit Rating Scale series* (HOVRS, HOVRS-A, HOVRS-A+ and HOVRS-A+2; Roggman et al. 2006, Roggman et al. 2010) is the most used observation instrument in home visiting to date and focuses on the strategies used by the home visitor, on the nature of the engagement of parent and child, as well as on the interactions between them.
- The *Home Visit Content and Characteristics Form* (HVCCF; Boller et al. 2009) was designed to complement the HOVRS, but can also be used on its own to document a range of more structural characteristics of the home visit, such as dosage and content.
- The *Home Visit Observation Form* (HVOF; McBride et al. 1993) focuses primarily on content delivery and on the ways and degree to which the home visitor is successful in engaging the family.
- The *COACH* (Dishion et al. 2010) is an observational system designed to quantify fidelity of program content delivery.

In this paper, we compare two of these instruments in the context of the home visiting service of the Peruvian *Programa Nacional Cuna Mas*, known as the *Servicio de Acompañamiento a Familias* (Cuna Mas, hereafter). More specifically, we consider the *Home Visit Rating Scale, version A+2* (HOVRS-A+2), the *Home Visit Content and Characteristics Form* (HVCCF) and compare them to a *checklist* specifically designed to measure structural and process quality aspects of the visit in a routinely manner. The HOVRS-A+2 and the HVCCF were both scored on recorded visits (videos) by trained coders, whereas the checklist was scored live by Cuna Mas supervisors on the same visits. We carry out the analysis of these instruments at the item and construct levels with a two-fold objective. Firstly, to document the structure, content and level of process quality of the Cuna Mas home visits amongst various dimensions, as assessed by these instruments. And secondly, to identify a subset of cost-efficient and reliable measures (i.e. set of items) of critical quality aspects that can be used by programs for monitoring purposes on a regular basis.

Whilst instruments such as the HOVRS-A+2 capture critical process quality aspects and can be used to structure constructive feedback to home visitors, they are complex, time-consuming and costly to administer. As home visiting programs continue to expand and grow in popularity in LAC and elsewhere, it is crucial to develop cost-effective quality assessments that can be implemented at-scale. This is an area of work still in its very early days and to which this study aims to contribute to.

The remainder of the paper is organized as follows. Section 2 briefly describes the Cuna Mas Program and Section 3 describes the study design, the sample and the data. Section 4 presents descriptive statistics of the analysis sample and characterizes the average Cuna Mas visit in terms of structural and process aspects of quality, as assessed by the HVCCF, the HOVRS-A+2 and the checklist. In Section 5 we analyze the performance and psychometric properties of the HOVRS-A+2, our reference for the measurement of interactions and engagement of participants during the visit, in the context of our study sample. In Section 6 we use the items in the checklist to propose a simplified, cost-efficient instrument that can be used to monitor quality at-scale. Finally, Section 7 concludes.

1. The Cuna Mas home visiting Program

Cuna Mas is Peru's largest national early childhood development service provider, reaching over 83 thousand children 0-36 months of age in rural areas in 2015 with its home visiting services. It delivers one-hour weekly home visits, carried out by a community member who has been trained by the Program. Although the Program requires home visitors to meet certain minimum requirements—being literate, at least 21 years old, speaking the local language, having past experience working with children, and having a certain level of recognition within the community—these are often difficult to find in the context of the poor, disperse rural communities where the Program operates. Community Council members² nominate a list of candidates to the Program and local supervisors select home visitors from that pool of applicants. Every home visitor is in charge of ten families. In turn, for every ten home visitors, there is one supervisor who works closely helping them plan their work and providing unstructured on-the-job training and mentoring. Supervisors, who are required to have higher education, play a key role in the continuous mentoring and in-service training of home visitors. They are expected to accompany each of the home visitors in their team for a whole day of work twice a month and observe their visits. Based on these observations, they help home visitors plan the activities for the coming weeks and they provide suggestions on how to improve the delivery of the visit. They are mentored by more senior staff from the Program's regional offices.

² *Comités de Gestión*, as they are known in the areas.

During the weekly home visit, the home visitor follows a structured curriculum that promotes more and better-quality adult-child interactions and learning through play. Each visit is organized around three ‘moments’. The first moment, ‘Family Life’ (*Vida en Familia*), seeks to help caregivers identify opportunities for learning and play in their daily routine activities. Often, more general messages related to health, nutrition, and sanitation are also delivered. The second moment, ‘Learning through Play’ (*Jugando Aprendo*), uses intentional play activities and materials—i.e. toys, games, and form boards, either provided by the Program or homemade with the caregiver—to promote the development of language, fine motor and cognitive skills. The third moment, ‘Tell me a Story’ (*Cuéntame un Cuento*) encourages vocabulary acquisition and language development through the use of picture books provided by the Program and helps caregivers learn how to use these materials with infants and toddlers. In addition to these moments, at the beginning of each visit, the home visitor and the caregiver spend time revising the activities introduced in the previous visit. Similarly, at the end of each visit, they recap and identify the specific activities that caregiver and child will practice during the coming week. To this end, all materials used in the visit are left in the home and collected in the next visit. The home visitor rotates these materials amongst her families as a ‘toy library’.

Cuna Mas targets the poorest in the country and, in particular, districts³: (i) with a poverty rate of at least 50 percent; (ii) with a chronic malnutrition rate of at least 30 percent; (iii) targeted by the conditional cash transfer program Juntos; and (iv) where over 50 percent of the population lives in rural communities (*Centro Poblados* or CP). Within these districts, Cuna Mas intervenes exclusively in rural CPs, which are defined as those with less than 2,000 families or 400 dwellings. All children from birth to 24 months in these communities are eligible to enroll in the Program and receive visits until they are 36 months of age, when they graduate.

The government was committed to a rigorous evaluation of the Program. Therefore, an experimental evaluation design was built taking advantage of its gradual roll out to assess Cuna Mas’ impact on child rearing practices and child development outcomes. The impact evaluation sample included 5,859 children ages 1-24 months at baseline in 360 communities in 180 Cuna Mas eligible districts. One hundred and twenty of these districts were randomly assigned to the treatment group and received visits, and 60 districts were assigned to the control group and did not receive the intervention. These districts were spread across 12 departments out of the 24 departments in the country, mostly located in the Andean and Amazon regions. Baseline data was collected in mid-to-late 2013 amongst children in both the treatment and control groups and their families. Similarly, endline data was collected approximately 24 months later, in mid-to-late

³ Districts are the equivalent of municipalities, or counties, in Peru.

2015. DGSE-MIDIS et al. (2015) offers more details on the evaluation sample and design and DGPP-MEF et al. (2016) reports the evaluation findings.

2. Study design, sample and data

2.1. Quality analysis sample

Data on home visit quality was collected between August and October 2015 on a sample of 554 children enrolled in Cuna Mas and receiving home visits at the time of data collection, and on their 176 home visitors. These children lived in 103 rural communities, in 69 districts and in nine departments. Communities for the quality sample were selected from the complete set of treatment communities in the Program's impact evaluation sample taking into consideration budgetary and practical constraints, as described in Appendix I. Children in the quality sample range between 9 and 38 months of age. In the descriptive analysis of the quality measures, we explore whether there is an association between child age and the quality of the home visit.

However, not all children in the quality sample were part of the impact evaluation sample. Within a community, we first identified children in the impact evaluation sample who were still receiving visits and their home visitor. For each home visitor, we filmed her visits to three or four different children, including the visit to *any* child in the impact evaluation sample.⁴ This resulted in 229 children in the quality sample who were also part of the impact evaluation sample and 325 children visited by the *same* home visitors but not included in the evaluation sample.

2.2. The data: measures, adaptation and training

The data for the analysis of home visit quality comes from various sources. For each child in the quality sample, one home visit was filmed and a checklist administered. The videos of the visits were later coded using instruments designed to assess home visit quality, as detailed next. In addition, child and family background characteristics were collected as part of the Cuna Mas impact evaluation, both at baseline and at endline. A home visitor and a community questionnaire were also collected at endline only.

2.2.1. Home visit quality measures scored from videos: HOVRS-A+2 and HVCCF

Quality of filmed visits was rated using suitably translated and adapted versions of the *Home Visit Rating Scale, version A+2* (HOVRS-A+2; Roggman et al. 2010) and the *Home Visit Content and Characteristics Form* (HVCCF; Boller et al. 2009).

⁴ For 67% of the home visitors, three visits to three different children were filmed, and for 22% of the home visitors, four visits were filmed. The distribution of the number of filmed visits for the remaining home visitors is as follows: for 3%, only one visit was filmed; two visits, for 6%; five visits for 1%; and six visits for another 1% of home visitors.

Developed from the original HOVRS (Roggman et al. 2006), the HOVRS-A+2 is the latest of a series of instruments specifically designed to assess home visit process quality. It can be scored by observation of a visit or of a video recording of a visit and consists of seven scales. These are organized in two domains: the first, *Home Visitor Practices*, includes four scales that focus on the strategies used by the home visitor to effectively carry out the visit—this is to say, the extent to which the home visitor is prepared for the visit, observes and responds to caregiver and child, and shows a warm and positive attitude towards them, amongst others. The second domain, *Family Engagement*, includes three scales aimed at measuring the nature of the interactions between caregiver and child, and their interest and level of participation throughout the visit.

The HOVRS family of scales has been widely used in the United States, both in the context of research studies and to support programmatic efforts. Indeed one of its strengths is that it has proven adaptable to diverse program goals, visit formats, and different cultures within the US, such as Spanish-only Latino families, rural Caucasian families, and urban African-American families.⁵

Originally developed as a complement to the HOVRS-A+2, the HVCCF takes account of more structural aspects of the visit such as the number of participants (adults and children), number of interruptions, language used, as well as a detailed record of the content topics covered and time allocated to each, including conversations unrelated to the visit. Similar to the HOVRS-A+2, the HVCCF can be scored by observation of a live visit or of a video recording.

Adaptation of the HOVRS-A+2 and HVCCF and training of the trainer. The HOVRS-A+2 was translated to Spanish and independently back-translated into English. The back-translation was reviewed by the developers, who suggested minor modifications. The developers also trained the trainer and the research team in the use of the instrument for two days, following which the trainer and two of the researchers continued to score videos of home visits until satisfactory inter-observer reliabilities (according to the criteria discussed below) were obtained.

Contrary to the HOVRS-A+2, many items in the HVCCF were substantially modified to better accommodate the specific structure of a Cuna Mas visit. The modified version was translated to Spanish and independently back-translated.

Filming of videos. The home visits were filmed between August and October 2015 by a team of 14 cameramen, who had been trained by a professional for a week including practice filming. A total of 24 people participated in the training, half of which had prior experience filming. The field team was selected based on an

⁵ For more information, see:
<http://healthandwelfare.idaho.gov/Portals/0/Children/HomeVisiting/MIECHV%20Assessment%20Guide%204-9-2013.pdf>

assessment of their performance during the training and practices. The content of the training was finalized after a month-long pilot exercise and included the specifics of how to use a video camera, how to select the optimal angle for filming and other strategies to obtain high quality videos given the limitations of the setting (filming indoors, background noise, static camera, etc.). It also included a session introducing the Program, the structure of a home visit, and revising a filming protocol that would minimize the disturbance created to the family and to the visit itself. This session was given by the trainer of the HOVRS who oversaw the entire training process.

Training of coders. The coding of the videos was carried out between November 2015 and May 2016 by a data collection firm based in Quito, Ecuador with experience in the coding of videos from child care, preschool and school settings using observational instruments to measure quality. Nine coders and a supervisor were trained during six days on the HOVRS-A+2 and the HVCCF. The training included between six and eight practice coding of entire videos, until acceptable reliability was consistently obtained. The five coders with the highest reliability with the trainer—all with higher education studies in psychology, education or related fields—were selected to code the study videos. The supervisor, a psychologist, acted as a back-up coder.

Two scores on the HOVRS-A+2 were considered non-reliable if: (i) one or more scales differed by more than one point (scale-level reliability criterion); and/or (ii) within a given scale, two or more items differed by two or more points (item-level reliability criterion).⁶ No inter-observer reliability criterion was defined for the HVCCF during the training. Reliability between the selected coders and the trainer was acceptable: throughout the training, item-level reliability ranged from 80 to 100 percent on average per coder and was above 90 percent for three of the coders; and scale-level reliability ranged from 61 to 86 percent on average per coder, with values above 80 percent for three coders.

Coding process. Each video was coded twice, independently, by two coders randomly selected from our pool of coders. If scores from both coders did not satisfy the item-level and/or scale-level reliability criteria, the video was assigned to a third coder—different from the first two and also selected at random—for a third coding.

To sustain reliability throughout the coding period, we put in place a calibration exercise whereby coders coded a 20-minute fragment daily. Scores were compared to the trainers' or the supervisors' and disagreements discussed. Half way through the training, however, we modified the calibration exercise to the coding of an entire visit twice a week since it was felt that coding 20-minute

⁶ Items in the HOVRS-A+2 are scored on a 4-point scale, with possible values 1, 3, 5 or 7. As such, a 1-point difference at the item-level corresponds to, for example, the difference between 1 and 3, 3 and 5, or 5 and 7.

fragments generated fewer disagreements than coding 1-hour visits.⁷ This strategy significantly reduced the number of videos that had to be coded a third time. The percentage agreement of the coders with the trainer during this exercise was very high, ranging from 93 to 98 percent.⁸

Of the 554 videos in the quality sample, a total of 214 videos (39 percent) were coded a third time: 142 videos (66 percent) did not comply with the scale-level reliability criterion, 17 (8 percent) did not comply with the item-level reliability criterion, and 55 videos (26 percent) did not comply with either criterion. After a third coding, average reliability increased from 65 to 82 percent at the scale level and from 87 to 94 percent at the item level.⁹

The average video was about one hour long. Coders coded no more than three videos a day in four-hour sessions, which included the calibration exercise. Any interactions that took place in Quechua were subtitled in Spanish by a team of Peruvian translators prior to coding. This occurred in 13 percent (74) of the videos.

Scoring. For each item, we keep the scores of the two coders with the highest scale- and item-level reliability and compute the average.

2.2.2. Home visit quality measure scored live: the checklist

During the course of the home visit being filmed, the Program supervisor administered a *checklist*, specifically designed to measure content and process quality aspects of the visit live. Based on the record forms used by Cuna Mas and by other home visiting programs in low- and middle-income countries (Jamaica, Bangladesh, Colombia and India), the checklist also incorporated items suggested by a team of experts that included the developers of the HOVRS-A+2 and the HVCCF. The aim was to design a relatively easy to train and administer tool that could be collected routinely by program staff while observing a visit and that could be used both to monitor quality in service delivery and to support home visitors in the performance of their role—this is to say, a tool that could be used simultaneously for supervision and for mentoring. On the last page of the checklist, we added a ‘complementary’ form to collect a few socio-demographic characteristics of the home visitor and the supervisor, time the family had been on the Program, and the total number of visitors the family had had to date, amongst other variables.

⁷ Interactions of different nature are more likely to occur during the full hour than in 20 minutes.

⁸ For the calibration exercise, agreement at the item level was defined as the number of items that did not differ by more than 2 points. The item-level reliability criterion, which considers the number of differences in items *within* a scale, was not used since coders were not always coding entire videos.

⁹ For this exercise, average reliability at the scale level is constructed by first computing the percentage of scales that meet the scale-level reliability criterion for each video and then taking the sample average. An analogous procedure was used for the average reliability at the item level.

The supervisors were trained for two days in the administration of the checklist by Program staff that, in turn, had been trained by the research team.

2.2.3. Other measures

As part of the impact evaluation of Cuna Mas, a household survey was collected both at baseline and at endline. It included basic socio-economic information (demographic composition of the household, education and employment of its members, dwelling characteristics and assets, etc.); the child's health history (birth weight, gestational age, vaccinations, and the like); maternal knowledge on child development (only at endline); and measures of the quality of the home environment using the responsiveness and acceptance scales in the *Home Observation for Measurement of the Environment* (HOME; Bradley et al., 2003) and the play materials and play activities scales in the *Family Care Indicators* (FCI, Frongillo et al. 2003).

At endline only, a home visitor and a community surveys were also collected. The home visitor questionnaire comprised home visitor characteristics (age, education, prior employment, offspring) and other information relevant to her role as home visitor, such as prior experience with children and/or as a social worker, time on the job, and knowledge on child development. These variables complement the information on the home visitor collected in the 'complementary' form of the checklist and described above.

The community survey collected data on the availability of services in the community (health center, pharmacy, bank, police station, etc.), the level of penetration of other social protection, education, and nutritional services, as well as distances to larger towns with more services (including the district capital), amongst others. It was administered to a community leader, either a member of the Community Council or someone else knowledgeable of daily life in the community.

3. Descriptive statistics of the sample and quality of visits

In this section we describe the sample, contextualize it, and characterize the average home visit in terms of structural and process quality, drawing on the information we collected with the HOVRS-A+2, the HVCCF and the checklist. As mentioned earlier, data on home visiting quality spanned a sample of 554 home visits in an identical number of distinct families. These visits were carried out by 176 home visitors who regularly work with 80 supervisors. Families in the sample

reside in 103 different rural communities, implying more than one home visitor per community.¹⁰

3.1. Characteristics of the children, home visitors, supervisors and communities in the sample

Table 1 combines data from different sources and presents descriptive statistics of the sample. Panel I shows information on supervisors, home visitors and children collected as part of the supervisor-administered checklist, and Panel II reports home visitor and community characteristics obtained from the home visitor and community surveys, respectively. These surveys included data for 154 of the 176 home visitors and 97 of the 103 communities in the quality sample.¹¹

Panel I documents that, on average, supervisors and home visitors in our sample had worked in the Program for close to 1.5 years (19 and 17 months, respectively). The ratio of home visitors-per-supervisor is 9.09, which is slightly lower than the 10 home visitors-per-supervisor ratio from the Program guidelines. Similarly, the ratio of children-per-home visitor observed (9.47) is also slightly lower than that allowed (up to 10). It is noteworthy that the standard deviation is smaller in the latter ratio than in the former one. Also, virtually all supervisors (99 percent) have completed their tertiary education, as required by the Program. Home visitors report having worked with the family whose visit was assessed for an average of 14 months.

Families report children have been enrolled in the Program for 21 months. Fifty-four percent of them also report having had more than one home visitor since their enrollment. Those who had more than one visitor worked with an average of two home visitors during their time in the Program. These children changed home visitors, on average, every 9 months (SD=3.58). The average child in the sample is 28.51 months old and. The children filmed are therefore on the older side of the age range the Program covers.

As described in Panel II, the home visitors are, on average, 31 years old and have 10 years of education, with a wide range of variation: whilst the majority (84 percent) has at least some secondary schooling, 14 percent have at most completed primary education and 14 percent have studied past secondary education. Twenty-eight percent of home visitors report having some prior experience working with children in the 0-3 age range and a similar percentage report having worked with families before their current role. Home visitors were also administered an adapted version of the Knowledge of Infant Development Inventory (KIDI; MacPhee, 2002), an instrument aimed at measuring an adult's knowledge of child development and behavior in the 0-5 years of age range. A

¹⁰ As explained in Appendix I, we had originally selected 98 communities for inclusion in the quality sample. The extra 5 communities are new communities, not included in the evaluation sample, where other children visited by the home visitors in the sample live.

¹¹ We were unable to find all 176 home visitors in the home visitor questionnaire because this questionnaire was only administered to one home visitor per community, selected at random.

total of 33 questions were administered: 22 of them adapted from the original version of the instrument, and 11 related to the training received by the Program. Home visitors answered, on average, 74 percent of the questions correctly.

The communities where these visits take place are rural, small and isolated. All of them are more than one hour away from the district capital, either by car or on foot. In almost half of them walking is the most common means of transportation to get there. Access to public and private services varies greatly: 88 percent of the communities in the sample have a health center, 96 percent a public preschool, 43 percent a pharmacy, 11 percent a bank, and 21 percent a police station. In turn, coverage of flagship social programs—including school feeding, conditional cash transfers, college grants, and non-contributory pensions—is large in these communities.

3.2. Home visit structure and content: the HVCCF data

Table 2 reports information about the all 554 visits that were filmed. These data come from the HVCCF and are very helpful to describe the structure and setting of a typical Cuna Mas home visit. Panel I reports ‘factual’ information on the visit length, participants and language in which it took place. Panels II and III focus on visit content and report the time allocated to topics covered and the activities carried out, respectively. Panel IV shows variables related to the overall tone of the visit.

We report the average score from the two coders with the highest reliability in the HOVRS-A+2 for all items except the dichotomous variables (Panels I and III), for which we report the responses of only one of the two coders. Reliability—defined as percent agreement between the two coders—was highest (97 percent on average, range from 92 to 99 percent) for the 0/1 items, particularly those in Panel I or those items in Panel III that enquired whether an activity had taken place. However, reliability on other items in Panel III was lower, especially for the item *Home visitor explained the objectives* (agreement between 61 and 75 percent, depending on the moment), which reduced the average reliability to around 75 to 80 percent. Percentage agreement was also lower for items in Panels II and IV, with average values of 76 percent (range from 40 to 99 percent) and of 58 percent (range from 45 to 72 percent), respectively. Nonetheless, the average difference between any two scores for items in these panels was between 1 and 2 points.

As shown in Panel I, and in line with Program guidelines, each visit was almost one hour long on average. However, there is substantial variation behind this mean: 25 percent of home visits lasted between 21 and 50 minutes, 60 percent between 51 and 70 minutes and 15 percent lasted between 71 and 120 minutes. In the large majority of visits (88 percent), mother and child were the only two people present, in addition to the home visitor. In 8 percent of the visits, between 1 and 3 children other than the focus child also participated, and in 5 percent of the cases, 1 or 2 adults, in addition to the main caregiver and the home visitor,

participated. Also, the child was awake during the entire visit in practically all the cases (99 percent). Mother and home visitor spoke exclusively in Spanish in 88 percent of the visits, exclusively in Quechua in 2 percent of the visits, and in a mix of both languages in 10 percent of the visits. As noted earlier, any interactions in Quechua were translated and subtitles were added to the videos, prior to coding.

The data shows that the conversations between mother and home visitor during the visit were largely focused on contents related to child development and well-being (Panel II in Table 2). Coders were asked to rate in a scale of 1-4 how long a set-list of 12 themes were covered for during the visit. A score of 1 indicates 'topic not treated', 2 'topic touched upon briefly', 3 'topic discussed at least for 10-15 minutes' and a score of 4 indicates 'topic is the main focus of the visit'. As shown, the topics discussed more at length are, in this order: child development (cognitive, language, motor), nutrition and health, and the importance of adult-child interactions¹². In turn, the topics that were addressed for shorter amounts of time were: other community services relevant for the family, topics unrelated to the home visit, and family dynamics or problems.

The Program structures every visit around three main moments: 'Family Life', 'Learning through Play' and 'Tell me a Story' (described in Section 2). Within each moment, four steps are expected to be implemented. First, the home visitor explains the purpose of the activity planned to be covered during the visit. Second, the child is allowed to freely explore the materials. This can happen while the home visitor explains the purpose of the activity to the caregiver. Third, the home visitor demonstrates the activity with the child. Fourth and last, caregiver and child carry out the activity on their own. In addition, home visitors are always expected to devote some time at the beginning of the visit to greet the family and review the activities from last week and, at the end, to wrap up and recap the activities caregiver and child will practice during the week. There is also some time devoted to singing songs, and at the very end, reviewing messages delivered and commitments acquired by the family, and filling forms. These activities were scored with a dummy variable taking the value of one if they took place, and zero otherwise.

As shown in Panel III of Table 2, the large majority of visitors followed the expected visit structure. The part of the visit that was observed with the lowest frequency—although it was carried out in 81 percent of cases—was the closure and wrap-up. Whilst in 94 percent of the visits, the review of the activities from the previous weeks took place, it was only in 78 percent of the visits that caregiver and child actually demonstrated these activities and hence could show that they had been practicing and the extent to which the child mastered the activity and was therefore ready to move to a higher difficulty level activity or not. Review of

¹² We note that this item was the one with lowest inter-coder reliability. One possibility is that coders were not in full agreement on what constituted 'child development' talk.

activities and recap are thought to be extremely important steps to ensure the Program is effective at changing caregiver behaviors and practices.

Within each structured moment, the steps that were most frequently carried out were letting the child explore the material, which occurs in 81 percent of the cases on average, and letting the child and her caregiver carry out the activity on their own, with a prevalence of 87 percent. At the same time, the steps that were most often missed were explaining the activity's objectives and demonstrating it, which only occurred in 62 to 65 percent of the occasions. It is also worth noting that the moments of the visit where the home visitors more frequently complied with the expected four steps were 'Learning through Play' and 'Tell me a Story', where 87 and 83 percent of the expected steps were followed, respectively. At the other extreme were the review from last week's activities and 'Family Life', where only 54 and 71 percent of the expected steps were carried out.

The last panel of the table describes some observations from the supervisor on the overall climate during the visit and on the presence of distractors. When asked to rank the extent to which environmental distractions (loud noises, people coming in, etc.) interfered with the visit, on a scale from 1 to 5 where 1 'no interruption or non-interfering' and 5 'very interfering interruption', the supervisors reported an average of 1.23 (SD=0.39). This implies that there were few interruptions during the visit or that existing interruptions were little disturbing. In addition, their perception was that the visits were largely entertaining (score of 4.07 out of 5) and had very little conflict or tension (score of 1.47 out of 5).

3.3. Interactions and engagement during the home visit: the HOVRS-A+2 data

As noted in Section 3.2.1, the HOVRS-A+2 has two distinct domains: *Home Visitor Practices* and *Family Engagement*. The first domain includes 24 items distributed in four scales: (i) *Home visitor responsiveness to family* (6 items); (ii) *Home visitor relationship with family* (8 items); (iii) *Home visitor facilitation of parent-child interaction* (6 items); and (iv) *Home visitor non-intrusiveness and collaboration* (5 items). The second domain includes 17 items in three scales: (i) *Parent-child interaction during the home visit* (7 items); (ii) *Parent engagement during the home visit* (6 items); and (iii) *Child engagement during the home visit* (4 items).

Each item describes a specific practice and is scored on a 7-point scale, with anchor points of 1 'needs training or support', 3 'adequate', 5 'good', and 7 'excellent'. The coder then assigns a score to the scale from 1 to 7 based on the general pattern of scores observed for the items.¹³ For easiness in the

¹³ This results in very similar scores to those obtained by computing the mathematical average of all items within the scale (correlations range from 0.92 for Responsiveness to family to 0.97 for Child engagement during home visit).

interpretation, we categorize this continuous scale following the categories defined in the scoring manual: 1 'needs training or support' (if 1 ≤ average scale score <3), 2 'adequate' (if 3 ≤ average scale score <5), 3 'good' (if 5 ≤ average scale score <7) and 4 'excellent' (if average score = 7). Domain scores are constructed averaging scale scores within the domain and a total score is constructed as the average of the two domain scores.

Table 3 reports descriptive statistics for each of the items, scales and domains in the HOVRS-A+2 for our sample of Cuna Mas home visits, and Table 4 shows how scale scores are distributed along the instrument's quality categories. Figures A1 and A2 in Appendix II show the distribution of scores at the item- and at the scale-level, respectively.

Overall, the *Home Visitor Practices* domain scored one point lower than *Family Engagement*, with a mean of 3.45 (SD=0.52) compared to 4.49 (SD=0.73) (Table 3). This resulted from relatively different quality distributions between both domains, as shown in Table 4. For *Home Visitor Practices*, 16 percent of the visits scored in the lowest quality category, 'needs training or support', and the large majority (84 percent) scored in the second lowest level of quality, 'adequate quality'. Less than 1 percent of the visits were classified as 'good quality'. On the other hand, for *Family Engagement*, only 1 percent of visits were rated as 'needs training or support', 72 percent of visits were considered of 'adequate quality' and 27 percent were classified as 'good quality'.

We next describe the item and scale scores in detail by domain.

Domain 1: Home visitor practices. Three of the four scales in the first domain scored, on average, within the Adequate quality range. These were: *Relationship with family* (mean=3.75, SD=0.69), *Facilitation of parent-child interaction* (mean=3.37, SD=0.67), and *Non-intrusiveness and collaboration* (mean=3.79, SD=0.57).

The scale that scored the lowest—within this domain and across all scales in the instrument—and the one most likely to appear in the Needs training and support category, was *Responsiveness to family* (mean=2.90, SD=0.56). The item that drives this low score is the extent to which the home visitor *Responds to family inputs for the agenda and activities of the home visit* (mean=1.49, SD=0.67). Given the structured nature of the Cuna Mas Program, this does not necessarily reflect poor home visitor performance but rather a proof of fidelity of implementation from the home visitors, who may not be deviating from the structured curriculum in order to consider families' inputs. The other two items with scores below 3 points (Adequate quality) were *To prepare for the home visit using parent-selected activities* (mean=2.95, SD=0.23), which, similar to the item just discussed, could actually reflect fidelity in implementation and conscientiously following the plan the home visitor and supervisor had agreed on, and *To get information about the family's strengths and child's development* (mean=2.71,

SD=1.13), which relates to engaging caregivers in conversations about their children and their development.

Within the other three scales, two items stood out because of lower scores: *To reflect on family's life and activities in relation to child's development* (mean=2.96, SD=0.74) in the scale *Relationship to Family*, and *To encourage the parent's ideas and interests for interactions with child* (mean=2.91, SD=0.44) in the *Non-intrusiveness and collaboration* scale. These items also relate to the ability of the home visitor to engage the caregiver in a discussion on the development of her child and reflect on what she does (or could do) to promote it.

On the other hand, a number of items showed higher scores, with five items attaining a median score of 5 (Good). More specifically, the items were: *To interact sociably with parents focusing on child development* (mean=4.45, SD=0.91), *To set the tone for positive interactions* (mean=4.37, SD=0.98) and *To express positive emotions about the home visit* (mean=4.44, SD=1.05) in the *Relationship with family* scale; and items *To keep parents in the "teacher" role* (mean=4.47, SD=0.80) and *To allow parent-child interactions to continue uninterrupted* (mean=4.50, SD=0.80) in the *Non-intrusiveness and collaboration* scale.

Domain 2: Family engagement. All three scales in the second domain scored within the Adequate quality range: *Parent-child interaction during the home visit* (mean=4.52, SD=0.79), *Parent engagement during the home visit* (mean=4.13, SD=0.83) and *Child engagement during the home visit* (mean=4.81, SD=0.97).

Nonetheless, there was some variation at the item level. Specifically, we observed three items with scores below 4 points: *To adapt activities to child's interests and needs and encourage child engagement* (mean=3.21, SD=0.89) in the *Parent-child interaction during the home visit*, and *To initiate activities and conversations* (mean=3.21, SD=0.98) and *To discuss questions and topics relevant to child and family* (mean=3.29, SD=0.89) in the *Parent engagement during the home visit* scale. As observed in the first domain, some of these lower-scoring items seem to require more subtle abilities such as redirecting interactions and conversations towards the child's development and caregiving practices in the home environment and responding or building on them as part of the activities carried out in the visit. They also require the home visitor to be able to scaffold activities (or their difficulty level) to the child's performance and interest. These are more complex competencies that probably require more and better training (specifically geared towards such goals) to support the home visitor in mastering them, as well as careful observation of progress since the last visit on the activities from such visit.

Five items showed very high scores, with 90 percent of the visits scoring 4 or above (between Adequate and Good). Two such items were in the *Parent-child interaction during the home visit* scale and were *To be available to engage in*

interactions (mean=5.15, SD=0.77) and *To observe and be ready to respond to the child's behavior* (mean=4.92, SD=0.60). Another one, in the *Parent engagement during the home visit* scale, was *To be ready to interact with both child and home visitor* (mean=4.93, SD=0.80). The last two were in the *Child engagement* scale: *To participate in home visit activities* (mean=5.04, SD=0.90) and *To sustain interactions with parent or home visitor* (mean=5.02, SD=1.04). It is worth noting that, on the home visitor's and parent's side, these items relate to readiness and good disposition whilst, on the child's side, they have to do with sustained interest, participation, and interactions during the visit. While all of these are important ingredients for the visits to be successful, items referring to the caregiver and home visitor seem essential considering that the engagement of the child often depends on the ability of adults to interest and engage her in the activity. On the other hand, the engagement of adults can depend on their abilities, readiness and good disposition to engage with each other and to collaborate in supporting the child's development.

HOVRS and Cuna Mas: Overall, we identified three items in the instrument that do not align with the Program's spirit. All of these are in the *Responsiveness to Family* scale: *To plan activities and topics of the home visit with the parent*, *To prepare for the home visit using parent-selected activities*, and *To respond to family input for the agenda and activities of the home visit*. As mentioned above, Cuna Mas' home visitors are required to follow a set curriculum that does not give them much room for caregiver input. We recalculated the HOVRS-A+2 scale scores dropping these items. As expected, scores improved for both the scale and the overall domain score (mean=3.28, SD=0.78 for Responsiveness to Family; and mean=4.02, SD=0.59 for domain Home Visitor Practices).

3.4. The checklist data

The checklist is comprised of 54 items organized into three sections: content and planning (38 items), interactions (10 items), and overall assessment of the child and caregiver's participation (6 items).¹⁴ Items in the first section are organized in different subsections that map one-on-one to the moments defined in the visit guidelines: *Welcome and review of previous visit*, *Family Life*, *Learning through Play*, *Tell me a Story*, and *Closure and recap*. Some of the items in these subsections collect similar information to items in the HVCCF. The remaining two sections provide an overall assessment of the quality of the interactions and the degree of engagement of visit participants, and therefore are more related to items covered in the HOVRS-A+2.

¹⁴ This excludes sections 1 and 2 which collect general information about the visit (i.e. identifying information of the home visitor, supervisor and child, the date and duration of the visit, etc.) similar to that reported in Tables 1 and 2, and items 20 and 21 which record the codes of the activities carried out during the visit for administrative purposes. Also, item 46 *Overall the home visitor's attitude towards the main caregiver* was turned into five dummy variables (i.e. one for each scoring option) for this analysis.

Summary statistics for items in these sections are reported in Panels I, II and III of Table 5 and described next. We structure the discussion around the moments of each visit.¹⁵

Introduction and closure. It is worth noting that, overall, the supervisor assessment of the visit tended to be more favorable than that of the coder. These types of discrepancies were found in items that coincide closely with those of the HVCCF, such as items 19, 51 and 52. For example, the percentage of supervisors that reported that the *Caregiver and child demonstrated the last visit's activities* was 87 percent for the checklist, compared to 78 percent in the HVCCF (Table 2). Also, when asked whether the *visit was fun* on a scale from 1 'almost never' to 5 'almost always', 16 percent of visits had a score from 1 to 3 in the checklist compared to 26 percent in the HVCCF. Similarly, when asked whether the *visit was conflictive/tense*, 86 percent of supervisors in the checklist recorded 'almost never' compared to 59 percent in the HVCCF.

Consistent with what was observed in the analysis of the HVCCF data, the moment of the visit where home visitors seemed to forget more of the expected steps and, on more occasions, was the closure. Home visitors *asked the caregiver to name the activities to carry out during the week* in only 68 percent of the visits, as reported by the supervisor. Consistent with what was observed with the HOVRS-A+2, *asking the caregiver to think of how to integrate these activities into the family routine or provide other ideas of games to play* was even less common (30 and 18 percent, respectively). We consider these to be low frequencies given the pivotal role these questions play in motivating behavioral changes amongst caregivers.

In the specific case of the Cuna Mas program, there are two important factors which render those practices nuanced and complex for home visitors to put in practice. The first one relates to the poor pre-service and in-service training home visitors receive, which is currently not designed in a way that supports the development of those skills. The second one relates to the home visitors' cultural background. Praising and providing constructive, detailed positive feedback is not at all common in the Andean culture and can in some cases be perceived as uncomfortable and unnatural to home visitors.

Family Life, Learning through Play and Tell me a Story. Supervisors recorded that the vast majority of home visitors brought the materials needed for the activities planned for each of the three moments (range between 93 to 100 percent). Supervisors were also asked to rank the quality of the explanations and demonstrations provided by the home visitor, independently, as well as her skill engaging caregiver and child in the activity, on a scale from 1 to 5, where: 1 'did

¹⁵ The distribution of scores for these items is shown in Figure A3 in Appendix II. Note that the item numbers in Figure A3 match those in Table 5.

not explain/demonstrate/engage', 2 'poor explanation/demonstration/engagement', 3 'average', 4 'good', and 5 'very good explanation/demonstration/engagement'. Results in Panel I show that the supervisors rated the *explanation of the activities' objectives* very similarly across moments, with ranges from 3.63 to 3.82. A similar pattern of coincidence of scores across moments is observed for the quality of the *demonstration of the activity* to the caregiver and for the quality of the *engagement of caregiver and child in the activity*, with scores between 3.81 and 3.83 and 4.01 and 4.18, respectively. Supervisors also ranked from 1 to 5 *how often the home visitor commented, labeled or responded to what the main caregiver and child were doing when they were trying the activity on their own* (conditional on the home visitor allowing them to try the activity on their own) as follows: 1 'almost never commented, labeled or responded', 2 'rarely', 3 'some of the time', 4 'most of the time' and 5 'almost always commented, labeled or responded'. The average rating across the various moments ranged from 3.66 to 3.86.

There was some more variation across moments in the reports of the frequency of occurrence of some actions. For example, supervisors reported that *The home visitors gave children time to explore the materials before starting the activity* in 80 and 82 percent of cases during 'Family Life' and 'Tell me a Story', compared to 93 percent in 'Learning through Play'. Interestingly, there are two other items where the frequencies between the first and last moment of the visit were very close to one another, while 'Learning through Play' seemed to follow a somewhat different pattern. These were whether *The Home visitor carried out all the steps for the activity* (84 and 85 percent vs. 92 percent for the first, third and second moments) and whether *The activity was of interest to the child* (83 and 83 percent vs. 95 percent). This is suggestive that, in the eyes of the supervisor, those activities that were most interesting to the child were also those where the home visitors followed the protocols more closely. We observe even a larger range of variation on whether *The home visitor gave the caregiver and child the opportunity to explore the activity on their own*. This was a lot more common in 'Learning through Play' (92 percent) than in 'Family Life' and 'Tell me a Story' (77 and 87 percent, respectively). In addition, the majority of supervisors reported that the transitions between moments were smooth and adequate.

In contrast, a 0/1 item with very little variation across moments was *The activity was adequate to the developmental level of the child*, which ranged from 95 to 100 percent. Given the large heterogeneity in the level of mastering a skill—and in developmental levels more broadly—across children of a given age, it is very unlikely that such a large proportion of activities were well-adjusted to the child's ability level. The large figures reported are more likely to reflect poor judgment of the supervisor to assess the adequacy of an activity to the child's developmental level and are possibly indicative of the pertinence of further covering this aspect during the trainings.

Interactions and overall assessment of the visit. The last set of items focused on rating the quality of the interactions and participation of caregiver and child during the visit. Each item is rated on a scale from 1 to 5, where: 1 'almost never', 2 'rarely', 3 'some of the time', 4 'most of the time' and 5 'almost always'. These items, reported in Panel II, are those in the checklist more closely related to those in the HOVRS-A+2. As shown, most scores ranged between 3 and 4. However, an action that occurred less frequently was for *The home visitors to ask the caregivers about other activities and games that the child usually plays with* (mean=2.19, SD=1.27). It is also worth noting that the home visitor generally felt more comfortable *praising the child* than *praising the caregiver*, as reported by the supervisor. With regard to the attitude of the home visitor, the supervisors reported that she *was encouraging* in 89 percent of the visits, whereas in 3-4 percent of the visits she was reported as *being intrusive/bossy, shy or indifferent*.

A positive aspect to highlight from the supervisor's assessment of the visit is the low scores given to unwanted practices. For example, *Topics of conversation unrelated to the home visit* were reported to occur 'sometimes' or more frequently in less than 5 percent of the visits (mean=1.26, SD=0.64). *The home visitor taking care of the child alone during the entire visit* showed similar frequencies (mean=1.18, SD=0.70). Moreover, only 1 percent of the home visits were reported to be *tense or conflictive* 'almost always' and less than 3 percent were reported to be so 'sometimes' (mean=1.20, SD=0.53); whereas 84 percent of the visits were considered fun 'most of the time' or 'almost always' (mean=4.20, SD=0.80).

4. Performance of the HOVRS-A+2 in the study sample

As explained in Section 2.2.1, the HOVRS-A+2 was translated to Spanish and the phrasing of some items was marginally modified to facilitate comprehension in our study context. In this section we study the performance—namely, internal consistency reliability and some measures of validity—of the modified instrument in rural Peru and compare it to the performance of the instrument in the US, whenever such comparison is available.

Internal consistency reliability explores the extent to which items in a scale measure the same underlying construct. For each scale, we computed it using Cronbach's alpha (α) on all items in the scale and compared this α to that reported in Roggman (2015). We complemented this analysis computing Pearson correlation coefficients (r) between each item with the scale score, each scale score with the domain score, and each domain score with the full score. We also conducted confirmatory factor analysis to check whether all items in a scale load into one factor (scale construct) and the relative contribution (factor load) of each item to the scale.

We report results in Table 6.¹⁶ The first two columns show that the scale α s in the sample are relatively similar to the scale α s in the original HOVRS-A+2 and generally good as denoted by $\alpha \geq 0.7$ in all but one instance. This corresponds to the *Non-intrusiveness and collaboration scale* ($\alpha=0.65$), which is also the scale exhibiting the largest difference in internal consistency with the manual. The scale with the second next poorest internal consistency is *Responsiveness to Family* ($\alpha=0.70$). Generally speaking, as in the original instrument, all scales in the *Family Engagement* domain show larger Cronbach's alpha than those in the *Home Visitor Practices* domain. Consistently, the *Family Engagement* scales also tend to have larger eigenvalues (Table 6, fourth column). As expected, all eigenvalues are larger than 1, confirming that items in a scale contribute to one common underlying construct.

The α s reported next to each item in the first column of Table 6 indicate what the α for the scale would be if such item was not included in the computation. As shown, for over 90 percent of the items, the α would stay the same or would decrease. This tends to coincide with those items with the lowest Pearson correlations with the scale score and those with the lowest contribution to the scale construct, and is yet another indication of the high levels of internal consistency amongst items in a scale, for all scales.

It is worth noting that those items with the lowest performance for scales *Responsiveness to Family* are precisely those that were identified as not being a good fit to the Cuna Mas' spirit and structure in Section 3.4. Replicating the analysis in Table 6 without these items showed slightly higher item correlations with the scale score and a marginal improvement in the scale α .¹⁷ Items in other scales related to harder practices for the home visitor to implement and/or those with low variability showed lower performance (e.g. *To help parents use available resources to support child development* in scale *Facilitation of parent-child interaction*, where 95 percent of home visitors scored a 3).

Further to this analysis, Table A1 in Appendix III shows that the various scales are highly and significantly correlated with each other, as are the various domains, thus indicating high congruence between scales and domains. This speaks to the interrelatedness between home visitor practices, positive interactions, and engagement of all participants as critical elements to achieve a high-quality visit.

Next, we investigate validity—this is, how well the instrument measures what it is supposed to measure—by studying the extent to which the scales and domains in

¹⁶ The sample size in this analysis is 554 observations (videos) for all scales except two: *Relationship with Family* (scale 2) and *Child Engagement during Home Visit* (scale 7). Following the manual instructions, item 5 in scale 2 was only coded for the 92 visits for which other family members were present; and item 2 in scale 7 was only coded for the 547 visits with children 12 months or older. Results are robust to excluding these items from the analysis and recomputing the statistics reported using all 554 observations available for the remaining items.

¹⁷ Results available from the authors upon request.

the HOVRS-A+2 are correlated with other variables as it would be theoretically expected. Specifically, we correlated the scale and domain scores with characteristics of the child, the home visitor (age, years of education, years of experience working with families, time in the Program, KIDI score etc.), structural characteristics of the visit (duration and the level of distractions), and the visit overall tone. These variables come from the checklist, the home visitor and community surveys, or the HVCCF, as indicated in the various panels of Table 7, where we report results.¹⁸ Note that, for this analysis we have no comparison benchmark. Similarly, we cannot investigate criterion validity with other instruments measuring similar constructs since such instruments are not available.

As shown in Panel II, education (in years) and experience working with families in prior jobs—most likely, in the context of other social programs—are the only home visitor characteristics that are significantly correlated with the *Home Visitor Practices* domain and the scales within: $r=[0.22, 0.33]$ with years of education and $r=[0.18, 0.24]$ with experience with families, all p-values <0.01 . These characteristics are also associated with *Family Engagement* and some of its scales, the exception being *Child Engagement during the Home Visit* for years of education, and *Parent-child interaction* for experience with families. Correlations are however lower. It is intriguing that the *Time the home visitor has been working with the family* is not correlated with any of the HOVRS-A+2 scores.

Age of the child is only significantly associated with *Child Engagement during the Home Visit* ($r=0.17$, p-value <0.01) (Panel I, Table 7). The only other child characteristic associated with some of the HOVRS-A+2 scales is whether the child has had at least another home visitor since enrolment. This might be indicative of the child being more used to ‘strangers’ (i.e. more used to more/new home visitors) and therefore more willing to engage in the visit and collaborate with the visitor. On the other hand, consistency in the home visitor might be expected to be positively associated with quality of the visits, which would not explain this finding. In any event, the magnitude of these correlations is very low ($r=[0.07,0.09]$, all p-values <0.01).

A few variables in the HVCCF exhibit relatively large associations with the HOVRS-A+2 scales. These are duration of the visits ($r=[0.14,0.26]$, p-value <0.01), the extent to which there were distractions ($r=[-0.13,-0.26]$, p-value <0.01), and the extent to which the visit was considered to be fun ($r=[0.39,0.71]$, p-value <0.01) or tense and conflictive ($r=[-0.24,-0.56]$, p-value <0.01). The high correlations for the latter two variables, alongside their content, seem to suggest

¹⁸ For the correlations with the home visitor characteristics in Panels I and II, we averaged the HOVRS-A+2 scores of each home visitor for the various scales, then estimated the correlations at the home visitor level (rather than at the visit-level as for the other variables in Table 7).

that these may be good ‘summary’ measures of home visit quality of interactions and engagement of participants.

5. Towards a cost-efficient instrument to assess home visit quality

In this section we analyze the checklist in detail to select a subset of items that could constitute a simplified and cost-efficient instrument (i.e. a shorter checklist) suitable for use in the field on a routinely basis for the monitoring of service quality. Ideally, this instrument should: (i) focus on meaningful aspects of quality; and (ii) be useful for supervisors to structure their mentoring, feedback, and planning sessions with home visitors. While the HOVRS-A+2 captures aspects of process quality that would be useful on both accounts, it is also complex, time-consuming and costly to train and administer in an at-scale Program, especially in a low-resource, low-skill context like the one where Cuna Mas operates. Therefore, substantial gains would arise from providing services with a simplified checklist that could be administered by observation by the Program supervisors during their routine observation of home visits.

We started by reorganizing and grouping the 54 items in the checklist according to the type(s) of information (‘constructs’) they aim to capture. We considered aspects related to the (i) visit structure; (ii) visit content; and (iii) other process quality aspects more related to the nature of the interactions, the nature of the facilitation practices by the visitor, the level of engagement of participants, and the overall tone of the visit.

For the later set of quality practices—interaction, engagement, facilitation, overall climate—we used the HOVRS-A+2 as reference and aimed to map items in the checklist to scales in the HOVRS-A+2. We found items that seemed closely related to each of the three scales in the first HOVRS-A+2 domain, *Home Visitor Practices*, and which we grouped under three ‘constructs’ with the names of those scales. A few other items seemed to capture aspects measured across the three scales in the *Family Engagement* domain and were collected under ‘Child and caregiver engagement during visit’, constituting a fourth ‘construct’ in the checklist. We also considered another group of items directly enquiring how well the visitor performed certain critical aspects in the visit during the visit moments—explanation of the activity’s objective, demonstration, and involvement of the caregiver and child during the activity. It also included the supervisor’s assessment on whether the visit was fun since this might reflect the easiness with which the home visitor conducted the visit. This fifth ‘construct’ was labeled ‘Key practices during visit moments’. For all these items, we carried out confirmatory factor analysis to investigate the extent to which they constituted a ‘construct’.

Finally, we identified other items in the checklist that were not covered in the HOVRS-A+2 and which we analyzed on their own. We refer to these as stand-alone items.

Table 8 presents the resulting set of items, organized in the five ‘constructs’. The final list of items met the following criteria:

1. The item could not have more than 15 percent of missing observations. This criterion was imposed since a few of the items did not apply in all occasions. For instance, item 28 *While the caregiver and the child were carrying out the activity, the home visitor commented/named/responded to what they doing* was only administered if the caregiver and the child had been given the opportunity to carry out the activity on their own.
2. The item had to have enough variability and in particular, at least 10 percent of the responses had to be different from 0 or 1. For those non-dichotomous items, often scored on a 1 to 5 scale, we investigated variability by combining adjacent categories (e.g. we looked at the variability of the item when the ‘always’ and ‘almost always’ categories were combined) and applied the same criterion—namely, more than 10 percent of answers in either category.
3. Whenever we identified a factor with eigenvalue larger than 1, we kept items that had a factor load equal or larger to 0.40 (a cut-off value often used in the literature).

Items that fulfilled criteria 1 and 2 but not criterion 3 and were thought to be meaningful, were left in the list as stand-alone items. Table A2 in Appendix III documents items that were not included alongside the criterion why they were discarded.¹⁹ Note that it is possible for a checklist item to appear in more than one of the ‘constructs’ identified.

The structure of Table 8 is very similar to that of Table 6. For each identified ‘construct’, and next to the ‘construct’ name, the first column shows the Cronbach’s α on all items classified under such ‘construct’.²⁰ The α s reported next to each item indicate what the α for the ‘construct’ would be if such item was not included in the computation. The second column reports the Pearson correlation coefficients (r) between each item with the ‘construct’ score (i.e. the first factor constructed by confirmatory factor analysis). The third and fourth columns show, respectively, the eigenvalue and factor loads of each item to the ‘construct’.

We observe that the internal consistency at the checklist ‘construct’ level is generally very good as denoted by $\alpha \geq 0.81$ in all but one instance, which corresponds to the ‘Child and caregiver engagement during the visit’ ($\alpha = 0.59$). It is also the ‘construct’ with the smallest number of items and that with the smallest eigenvalue ($\alpha = 1.06$). As expected, after the item selection process, all eigenvalues were larger than 1 and all factor loads were reasonably large. These

¹⁹ As shown in Table A2, in addition to the five ‘constructs’ mentioned above, two additional ‘constructs’ were considered: ‘Knowledge of protocols’ and ‘Recap of activities’. However, most items that had been mapped to them did not comply with the criteria and hence were dropped.

²⁰ The stand-alone items are presented in Table 9.

findings suggest that the selected items do load into the proposed set of 'constructs'.

In Table 9 we investigate the association of the checklist 'constructs' and the scale, domain and total scores in the HOVRS-A+2. We obtain moderate significant correlations ($r=[0.31,0.38]$, all p -values <0.01) between the 'constructs' in the checklist and the total HOVRS-A+2 score. As expected, these associations are larger between the first three checklist 'constructs' and the first HOVRS-A+2 domain ($r=[0.35,0.36]$, all p -values <0.01), and between the fourth checklist 'construct' and the second HOVRS-A+2 domain ($r=0.365$, p -value <0.01). Within the first three 'constructs', the size of the correlation is not always larger with the matching scale than with the other scales—in fact, for all checklist 'constructs', the largest correlations are with the *Relationship with Family* and *Facilitation of Parent-Child Interaction*. This is not surprising since Cuna Mas' home visitors are trained to follow a highly-structured curriculum without much adaption to individual families, while the other two subscales of this domain, *Responsiveness* and *Non-Intrusiveness*, require the home visitor to individualize the visit to the family using caregiver input and responding to such input. These are more complex skills. Interestingly, 'Key practices during visit moments' always displays a correlation with any of the HOVRS-A+2 scales, domains or total score at least as large as any other 'construct' or individual item in the checklist. We interpret this as indicative of the relevance of these 'how to' questions as indicators of good home visitor practices, positive interactions amongst all participants and adequate family engagement during the visit. The correlations between the HOVRS-A+2 scores and the stand-alone items vary substantially and rarely attend values higher than $r=0.22$.²¹ Nonetheless, as shown in Table A3 in Appendix III, items that are part of a checklist 'construct' do attain correlations over this value with the HOVRS-A+2 scores.²²

Finally, we assess the correlations between the 'constructs' in the checklist and the variables reported in Table 7 related to the child, home visitor, and visit characteristics. Results are reported in Table 10. Years of education is significantly correlated with all checklist 'constructs' ($r=[0.24, 0.32]$, all p -values <0.01), except for *Child and caregiver engagement*. Another home visitor characteristic associated with two of the 'constructs' is her knowledge of child development ($r=[0.14, 0.15]$; all p -values <0.01). It is worth noting that the size of these correlations is larger than that of years of education with the HOVRS-A+2 scores. The other two variables that exhibit significant correlations with the checklist 'constructs' are the extent to which the visit was considered to be fun ($r=[0.20,0.34]$, p -value <0.01) or tense and conflictive ($r=[-0.13,-0.29]$, p -value

²¹ Note that given the insignificant correlations with the HOVRS-A+2 scales, checklist item *The home visitor gave the child the opportunity to explore the material before beginning the activity (moment 3)* was removed at this stage from the final set of items.

²² Results in Table 9 are robust to constructing the HOVRS-A+2 scale, domain and total scores using factor analysis instead of using the standard score (results available upon request).

<0.01). This in line with what observed with the HVCCF, and further supports the notion of these variables possibly being good ‘summary’ measures of home visit quality of interactions and engagement of participants.

Overall, we consider the set of items in Table 8 good candidates for a shorter checklist, and those in the fifth checklist ‘construct’ in particular (see Table A4 in Appendix III for the proposed checklist). A few considerations are however in order. First, the ideal shorter checklist should also include additional information related to the visit structure, the child and the home visitor, as proposed in Table A3. Some of these variables would only need to be recorded during the first visit to the family (e.g. date of entry into the Program).

Second, it is unclear the extent of the contribution of the stand-alone items due to the generally lower correlations they exhibited with both the HOVRS-A+2 scales and the socio-demographic variables. These items also tend to have lower variability. It is important to highlight, however, that these lower correlations might simply reflect the fact that the stand-alone items include information that is not picked up by the HOVRS-A+2. In fact, most refer to the home visitors’ knowledge of the protocols—an aspect the HOVRS-A+2 does not aim to capture.

Third, it is also worth mentioning the possibility that some items showed very little variability—and were therefore not considered good candidates for the shorter checklist—because they were not sufficiently well trained and/or administered in the field, and not because they were items that did not capture valuable information *per se*. One example is item *The activity was adequate to the developmental of the child* which, intuitively, should be important for the visit’s success and is also more difficult to score as it requires knowledge and experience. In some cases, these more difficult items may be substituted by easier-to-train/administer alternatives. For example, in the case of item *The activity was adequate to the developmental of the child*, research on master motivation and task persistence (see McCall, 1995) suggests that children will be more interested in activities that are neither below nor above their developmental level (i.e., neither too easy nor too difficult). As such, supervisors may be better able to determine whether the child remained interested in the activity—an item that was included in construct ‘Child and caregiver engagement during the visit’—rather than whether it was developmentally appropriate. One obvious limitation, however, is that children may not be interested in the activity for reasons other than the challenge of the activity itself (e.g., if there are distractions in the environment, if the child is hungry, etc.). Another potential substitute for this more difficult item includes asking *Could the child do the activity*, with answers “yes, right away”, “yes, after some trying”, and “no, not even after trying.”

Fourth, a limitation that should be kept in mind when having Program staff administer themselves the quality assessment tool is the potential tendency to inflate scores. While reports from the HVCCF and the checklist were largely

consistent with each other, some items (e.g., caregiver and child demonstration of last visit's activities; whether the visit was fun or conflictive) showed evidence of supervisors 'overstating' quality. This could be due to several reasons, including inadequate training, supervisors wanting to 'protect' the home visitor they supervise, or supervisors believing the home visiting quality is a reflection of the quality of their own supervision, amongst others.

Finally, some of the items in this final set are quite specific to the Cuna Mas Program—particularly those that refer to different moments of the visit—and as such should be revised if applied to another context.

6. Conclusion

In this paper, we have documented the quality of the Cuna Mas home visits along a number of both structural and process aspects of quality, as assessed by two well-known instruments: the HOVRS-A+2 and the HVCCF. Both instruments have been used to evaluate quality of home visiting services in the US and were adapted and contextualized to the areas where Cuna Mas operates—namely, rural disperse communities in Peru. Both instruments were scored by trained coders on video recordings of a sample of 554 home visits, in an identical number of homes. During the filming of these visits, the Cuna Mas supervisor filled in a checklist, specifically designed by the research team to collect content and process quality aspects of the visit live and in a more cost-efficient manner. This checklist aimed to be an easy to train and administer tool that could be collected routinely by Program staff whilst offering reliable information of home visit quality. Ultimately, the checklist—or a shortened and simplified version of it—would become a basic tool not only to assess quality in service delivery but also to provide feedback, mentoring and support to home visitors in the performance of their role. To this end, in the last section of the paper and after documenting home visit quality levels, we analyzed the items in the checklist in detail to select those that would be good candidates to constitute such simplified and cost-efficient instrument.

The filmed home visits were carried out by a total of 176 home visitors. As mandated by Cuna Mas' guidelines, the home visitors in the sample worked with between nine and ten children. They had been in the Program for almost two years, a majority had some secondary education—some completed, others not—and about a quarter had worked with families in previous employment. The children in the sample were between two and two and a half years old on average and a bit more than half of them had had at least one other home visitor since their enrollment in the Program. This is indicative of high home visitor turnover and might have implications for quality in service delivery.

Overall, the home visits were of the expected duration (one hour) and largely covered the content they were supposed to cover. Whilst the home visitors tended to follow the established steps, some of these were more likely to be missed than

others, and in particular, explaining the objectives and demonstrating the activity. Other important steps not always carried out were the review and demonstration (by caregiver and child) of the activities from the previous week and the recap with caregiver of the activities to be practiced during the following week. These steps are key to ensure the visit is modifying behaviors and practices in the home, beyond providing an hour of play. Reports from the HVCCF and the checklist were largely coinciding on all these aspects, although supervisors showed a general tendency to overstate 'quality'.

Regarding process quality aspects, the majority of visits scored 'adequate' or 'good' quality levels in the second domain of the HOVRS-A+2, which is related to caregiver and child participation and engagement in the visit. However, scores on the first domain, more specifically related to the role of the home visitor during the visit, were lower with almost all visits scoring in the 'adequate' or 'needs training or support' quality levels. The items with the lowest scores within domains and scales were often those related to the ability of the home visitor to effectively engage the caregiver, obtain her opinion, views or already existing practices, and incorporating these in the activities, praising the caregiver, and scaffolding activities to the child's performance and interest. These are more complex and often more subtle competencies that probably require more specific training.

The detailed analysis of the checklist items and comparison with the HOVRS-A+2 scores so as to identify good candidate items that could constitute a simplified and cost-efficient instrument proposes a set of 27 items. Within these, those directly enquiring *how well* did the home visitor carry out specific actions (involved caregiver and child, praised, commented or labeled, etc.) seem to be more highly correlated with the quality aspects identified in the HOVRS-A+2. Usually, the highest correlations were obtained with the second HOVRS-A+2 domain and with the *Relationship with Family* and *Facilitation of Parent-Child* scales for the first domain.

It is also interesting to note that both the supervisor and the coder coincided in rating the majority of visits as fun and non-conflictive or tense. Moreover, these scores showed the highest correlations with the HOVRS-A+2 domains and total scores, hence suggesting that these might be good 'summary' measures of home visit quality of interactions and engagement of participants. In addition to all these items, the ideal shorter checklist for monitoring purposes would also need to include a number of structural quality elements in the visit—this is to say, total length, number of participants, fidelity to content, and interference, as well as some home visitor's characteristics (as shown in Table A4 in Appendix III).

All items should be very specific and use simple, direct language so that they can be quickly filled in as the visit develops and without taking too much of the supervisor's attention from it. This is also important in order to ensure that there is no room for ambiguity in their interpretation and scoring. Training of the

supervisors in its administration is critical to attain a reliable measure. A clear limitation of the current study is the short training that the supervisors received on the observation and scoring of the checklist variables, as reflected in the little variation observed on many items and on the fact that supervisors seemed to overrate scores compared to coders.

Another important limitation of the study is that the HOVRS-A+2 is a tool that was developed for a very different context than rural Peru. Similarly, not all aspects in the Cuna Mas visits—given both the structure, specificities of the curriculum, and the profile of the visitors—are fully aligned with some of the practices considered in the instrument as important quality elements. Nonetheless, the instrument showed good internal consistency, compared to that reported in the manual.

With these limitations in mind, programs aiming to use the suggested shorter checklist might consider revisiting some of the items we did not analyze as well as reconsider and reformulate those that were selected but are specific to the Cuna Mas' structure.

Future next steps for investigation include correlating the HOVRS-A+2 scores and the final checklist score—based on the selected subset of highest performing items, including items measuring structural quality aspects—with child development outcomes as a further test of validity. Additionally, the information content of these scores could be improved using Item Response Theory techniques.

References

- Ángeles, G., Gadsden, P., Galiani, S., Gertler, P., Herrera, A., Kariger, P. & Seira, E. (2011). Evaluación de Impacto del Programa Estancias Infantiles para Apoyar a Madres Trabajadoras. Informe Final de la Evaluación de Impacto. Instituto Nacional de Salud Pública. México.
- Britto, P.R., Stephen, J.L., Proulx, K., Yousafzai, A.K., Matthews, S.G., Vaivada, T., Perez-Escamilla, R., Rao, N., Ip, P., Fernald, L.C.H, MacMillan, H., Hanson, M., Wachs, T.D., Yao, H., Yoshikawa, H., Cerezo, A., Leckman, J.F., Bhutta, Z.A. & the Early Childhood Development Interventions Review Group, for the Lancet Early Childhood Development Series Steering Committee. (2016). Nurturing care: promoting early childhood development. *The Lancet*. 389: 91-102
- Barrett, K., Zaveri, H., & Strong, D. A. (2010). Fidelity Data Collection Manual for the Evidence-Based Home Visiting to Prevent Child Maltreatment Cross-Site Evaluation. Princeton, NJ: Mathematica Policy Research.
- Bayley, N. (1969). Bayley Scales of Infant Development. Manual. New York: Psychological Corp.
- Bayley, N. (2006). Bayley Scales of Infant and Toddler Development—Third Edition: Technical manual. San Antonio, TX: Harcourt Assessment.
- Boller, K., Vogel, C., Cohen, R., Aikens, N., and Hallgren, K. (2009). “Home Visit Characteristics and Content Form.” Princeton, NJ: Mathematica Policy Research
- Bradley, R., Caldwell, B., & Corwyn, R. (2003). The Child Care HOME Inventories: Assessing the quality of family child care homes. *Early Childhood Research Quarterly*. 18: 294-309.
- Dishion, T. J., Knutson, N., Brauer, L., Gill, A., & Risso, J. (2010). Family Check Up: COACH ratings manual. Unpublished coding manual. Available from the Child and Family Center, 195 West 12th Avenue, Eugene, OR 97401.
- DGSE-MIDIS, DGPP-MEF, & BID. (2015). Documento de Línea de Base para la Evaluación de Impacto del Programa Nacional Cuna Más – Servicio de Acompañamiento a Familias., Documento interno del MIDIS.
- DGPP-MEF & BID (2016). Reporte de la evaluación de impacto del Servicio de Acompañamiento a Familias del Programa Nacional Cuna Más. Unpublished manuscript, Inter-American Development Bank.
- Fernald, L., Kariger, P., Engle P., & Raikes, A. (2009). Examining child development in low-income countries: a toolkit for the assessment of children in the first five years of life. Washington, DC: The World Bank.
- Fernandes, M., Stein, A., Newton, C.R., Cheikh-Ismail, L., Kihara, M., Wulff, K., de Leon Quintana, E., Aranzeta, L., Soria-Frisch, A., Acedo, J., Ibanez, D.,

- Abubakar, A., Giuliani, F., Lewis, T., Kennedy, S. & Villar, J. (2014) The INTERGROWTH-21st Project Neurodevelopment Package: A Novel Method for the Multi-Dimensional Assessment of Neurodevelopment in Pre-School Age Children. *PLoS ONE*, 9(11), e113360.
- Frongillo, E. A., Tofail, F., Hamadani, J.D., Warren, A.M., & Mehrin, S.F. (2014) Measures and indicators for assessing impact of interventions integrating nutrition, health, and early childhood development. *Annals of the New York Academy of Sciences*, 1308, 68–88.
- Frongillo, E.A., Sywulka, S.M. & Kariger, P. (2003). UNICEF psychosocial care indicators project. Final report to UNICEF, mimeo Cornell University
- MacPhee, D. (2002). Knowledge of Infant Development Inventory Survey of Child Care Experiences Manual. Mimeo.
- McBride, S.L., & Peterson, C.A. (1993). Home visit observation form. Unpublished manuscript, Iowa State University.
- McCall, R. B. (1995). On definitions and measures of mastery motivation. In R. H. MacTurk & G. A. Morgan (Eds.), *Mastery motivation: Origins, conceptualizations, and applications* (pp. 273–292). Norwood, NJ: Ablex.
- Paulsell, D., Avellar, S., Sama Martin, E., & Del Grosso, P. (2010). Home Visiting Evidence of Effectiveness Review: Executive Summary. Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. Washington, DC.
- Roggman, L. A., Cook, G. A., Innocenti, M. S., Jump Norman, V. K., Christiansen, K., Boyce, L. K., Aikens, N., Boller, K., Paulsell, D., & Hallgren, K. (2010). Home Visit Rating Scales—Adapted and Extended (HOVRS-A+) version 2. Adapted from Roggman, L. A., et al (2008). Home Visit Rating Scales. In L. Roggman, L. Boyce, and M. Innocenti, *Developmental Parenting* (pp. 209-217). Baltimore: Paul H. Brookes Publishing.
- Roggman, L. A., Cook, G. A., Jump, V., Boyce, L. K. & Innocenti, M. S. (2006). Home Visit Rating Scales (HOVRS). Unpublished manuscript. Utah State University, Logan, UT.
- Roggman, L. (2015). Observing home visiting process: Practices and engagement in relation to outcomes. Home Visiting Research Network Day, Washington, DC. May 6.
- Schodt, S., Parr, J., Araujo, M.C., & Rubio-Codina, M. (2015). Measuring the Quality of Home-Visiting Services. A review of the literature. Inter-American Development Bank. Technical Note IDB-TN-881
- Squires, J., Bricker, D., Twombly, E., Nickel, R., Clifford, J., Murphy, K., Hoselton, R., Potter, L., Mounts, L. & Farrell, J. (2009). *Ages & Stages English*

Questionnaires (Third Edition) (ASQ-3): A Parent-Completed Child-Monitoring System. Baltimore: Brookes Publishing Co., Inc.

Thomason, A.C. & La Paro, K.M. (2009). Measuring the quality of teacher-child interactions in toddler child care. *Early Education and Development*. 20(2): 285-304.

Tables and figures

Table 1: Characteristics of the supervisors, home visitors, children and communities in the sample, by data source

	N	Mean/ Proportion	SD
Panel I: Supervisor-administered checklist			
Supervisor			
Time in Program (months)	80	18.58	9.56
Ratio of home visitors-per-supervisor	80	9.09	3.48
Proportion with tertiary technical complete	80	0.40	
Proportion with university incomplete	80	0.01	
Proportion with university complete	80	0.59	
Home Visitor			
Time in Program (months)	176	17.10	9.50
Ratio of children-per-home visitor	175	9.47	1.24
Time working with family (months)	176	14.40	9.31
Child			
Time in Program (months)	553	21.40	6.94
Child has had at least one other home visitor since enrolment	554	0.54	
Number of home visitors since enrolment (if more than 1)	295	1.72	0.87
Ratio of time in Program to number of home visitors (if more than 1)	295	9.25	3.58
Age (months)	554	28.51	5.88
Panel II: Home visitor and community surveys			
Home Visitor			
Age (years)	154	31.18	7.92
Years of education	154	10.17	2.48
Proportion with primary incomplete	154	0.02	
Proportion with primary complete	154	0.14	
Proportion with secondary incomplete	154	0.19	
Proportion with secondary complete	154	0.51	
Proportion with more than secondary complete	154	0.14	
Proportion with experience with children in the 0-3 age range	154	0.28	
Proportion with experience working with families	154	0.29	
Knowledge of Infant Development Inventory (KIDI) score	154	0.74	0.09
Community			
Time to the district capital by foot (hours)	48	1.72	3.51
Time to the district capital by car (hours)	46	1.05	0.87
Has a health center	97	0.88	
Has a preschool center of the Ministry of Education	97	0.96	
Has a pharmacy	97	0.43	
Has a bank or <i>caja municipal/rural</i>	97	0.11	
Has a police station	97	0.21	
Beneficiary of the school breakfast program "Qali Warma"	97	0.97	
Beneficiary of the school lunch program "Qali Warma"	97	0.90	
Beneficiary of "Beca 18" program	97	0.58	
Beneficiary of "Juntos" program	97	1.00	
Beneficiary of "Pensión 65" program	97	1.00	

Notes: Data from three sources: the supervisor-administered checklist during the filming of the home visit (Panel I) and the home visitor and community surveys collected at endline (Panel II).

Table 2: Structure and Content of the Visit from the Home Visit Content and Characteristics Form (HVCCF, N=554)

	Mean/ Proportion	SD	Possible range of scores
Panel I: General information about the visit			
Duration of the visit	58.62	12.33	
Duration of the visit, by range			
21 y 50 minutes	0.25		
51 y 70 minutes	0.60		
71 y 120 minutes	0.15		
Other people present during majority of visit, in addition to home visitor and child:			
only the mother	0.88		[0-1]
mother and father	0.03		[0-1]
mother, father and another person	0.01		[0-1]
mother and another person	0.08		[0-1]
Proportion of visits in which children (other than focus child) participated	0.08		[0-1]
Proportion of visits in which adults (other than main caregiver and home visitor) participated	0.05		[0-1]
Proportion of visits in which the child was awake during the entire visit	0.99		[0-1]
Languages the home visit was conducted in (mother and home visitor)			
Spanish only	0.88		[0-1]
Quechua only	0.02		[0-1]
Spanish and Quechua	0.10		[0-1]
Panel II: Extent of time allocated to topics¹			
Hand washing	1.32	0.46	[1-4]
Personal or household hygiene	1.39	0.41	[1-4]
Nutrition and health	1.91	0.26	[1-4]
Personal care routines (including security)	1.14	0.24	[1-4]
Child development (cognitive, language, motor)	2.14	0.57	[1-4]
Temperament of the child and socio-emotional aspects	1.17	0.26	[1-4]
Importance of the interactions between the child and caregiver (answering and talking to the child, hugging and praising)	1.79	0.34	[1-4]
Father's involvement in caregiving	1.35	0.39	[1-4]
Household dynamics, family problems, etc.	1.11	0.21	[1-4]
Relevant community services for the child	1.22	0.30	[1-4]
Relevant community services for adults and/or the household	1.00	0.04	[1-4]
Topics not related to the home visit	1.01	0.07	[1-4]

Panel III: Activities carried out during the visit**Greetings and arrival at the home**

Activity was carried out 0.91 [0-1]

Reviewing the activities from the previous visit

Activity was carried out 0.94 [0-1]

Home visitor explained the objectives 0.34 [0-1]

Child explored the material 0.63 [0-1]

Home visitor demonstrated the activity with the child 0.45 [0-1]

Child and caregiver carried out the activity on their own 0.78 [0-1]

Family life (moment 1)

Activity was carried out 0.99 [0-1]

Home visitor explained the objectives 0.70 [0-1]

Child explored the material 0.69 [0-1]

Home visitor demonstrated the activity with the child 0.64 [0-1]

Child and caregiver carried out the activity on their own 0.81 [0-1]

Learning through play (moment 2)

Activity was carried out 0.99 [0-1]

Home visitor explained the objectives 0.77 [0-1]

Child explored the material 0.99 [0-1]

Home visitor demonstrated the activity with the child 0.75 [0-1]

Child and caregiver carried out the activity on their own 0.97 [0-1]

Tell me a story (moment 3)

Activity was carried out 0.99 [0-1]

Home visitor explained the objectives 0.67 [0-1]

Child explored the material 0.92 [0-1]

Home visitor demonstrated the activity with the child 0.76 [0-1]

Child and caregiver carried out the activity on their own 0.92 [0-1]

Singing

Activity was carried out 0.92 [0-1]

Closure and wrap-up

Activity was carried out 0.81 [0-1]

Reviewing the commitments and messages

Activity was carried out 0.87 [0-1]

Filling in the forms

Activity was carried out 0.85 [0-1]

Panel IV: Overall impression of the visit

Extent to which environmental distractions (television, phone calls, visitors, pets, other children, noise, etc.) interfered with the home visit.² 1.23 0.39 [1-5]

Was the visit fun?³ 4.07 0.68 [1-5]

Was the visit conflictive/tense?³ 1.47 0.53 [1-5]

Notes: Data in Panels I (except the duration of the visit) and III correspond to the scoring from one of the two coders of each visit selected at random. Panels II and IV (and duration of the visit in Panel I) correspond to the average of the scores of the 2 coders with the highest reliability.

¹ Variables coded as 1 "topic not treated" to 4 "main topic of the visit."

² Extent to which environmental distractions interfered with the home visit is scored on a scale from 1 to 5 where 1 refers to "no interruption or non-interfering" and 5 to "very interfering interruption."

³ Was the visit fun and Was the visit conflictive/tense are scored on a scale from 1 to 5 where 1 refers to "almost never" and 5 to "almost always."

Table 3: Process quality levels of the Cuna Mas visit from the Home Visit Rating Scale (HOVRS-A+2, N=554)

	Mean	SD	Min	Max	P10	P50	P90
Domain 1. Home Visitor Practices	3.45	0.52	2.00	5.25	2.75	3.38	4.13
1. Responsiveness to Family	2.90	0.56	1.50	4.50	2.00	3.00	3.50
1. To plan activities and topics of the home visit with the parent	3.15	0.46	1.00	5.00	3.00	3.00	4.00
2. To prepare for the home visit using parent-selected activities	2.95	0.23	1.00	4.00	3.00	3.00	3.00
3. To get information about the family's strengths and child's development	2.71	1.13	1.00	6.00	1.00	3.00	4.00
4. To provide feedback on family strengths for supporting child development	3.81	0.94	1.00	6.00	3.00	4.00	5.00
5. To adapt activities to the family's interests and needs	3.31	0.85	1.00	6.00	2.50	3.00	4.00
6. To respond to family input for the agenda and activities of the home visit	1.49	0.67	1.00	4.00	1.00	1.00	2.00
2. Relationship with Family	3.75	0.69	1.50	6.00	3.00	4.00	4.50
1. To interact sociably with parent(s), focusing on child development	4.45	0.91	2.00	7.00	3.00	5.00	5.00
2. To set the tone for positive interactions	4.37	0.98	1.00	7.00	3.00	5.00	5.00
3. To express positive emotions about the home visit	4.44	1.05	1.00	7.00	3.00	5.00	6.00
5. To engage other family members if present during the home visit ¹	3.07	1.75	1.00	7.00	1.00	3.00	5.00
7. To reflect on family's life and activities in relation to child's development	2.96	0.74	1.00	6.00	2.00	3.00	4.00
8. To show respect and acceptance of the family, home, culture, and lifestyle	3.16	0.44	2.00	6.00	3.00	3.00	4.00
9. To discuss sensitive issues respectfully and reflectively	3.34	1.13	1.00	6.00	2.00	3.00	5.00
3. Facilitation of Parent-Child Interaction	3.37	0.67	1.50	6.00	2.50	3.50	4.00
1. To elicit ongoing parent-child interactions during the home visit	4.14	0.98	1.00	7.00	3.00	4.00	5.00
2. To promote developmentally supportive interactions	3.12	0.96	1.00	7.00	2.00	3.00	4.00
3. To engage parent and child together	3.33	0.81	1.00	7.00	3.00	3.00	4.00
4. To support parent responsiveness to child cues	3.06	0.94	1.00	6.00	2.00	3.00	4.00
5. To directly encourage or reinforce positive parent-child interactions	3.47	1.06	1.00	7.00	2.00	3.00	5.00
6. To help parents use available resources to support child development	3.04	0.22	2.00	5.00	3.00	3.00	3.00
4. Non-Intrusiveness and Collaboration	3.79	0.57	1.50	5.50	3.00	4.00	4.50
1. To encourage the parent's ideas and interests for interactions with child	2.91	0.44	1.00	5.00	3.00	3.00	3.00
2. To avoid intruding on or ignoring parent-child interactions	3.78	0.93	1.00	6.00	3.00	4.00	5.00
3. To keep parent in the "teacher" role	4.47	0.80	1.00	6.00	3.00	5.00	5.00
4. To follow the lead of parent and child in pace and activities	3.35	0.81	1.00	6.00	3.00	3.00	4.00
5. To allow parent-child interactions to continue uninterrupted	4.50	0.80	1.00	7.00	3.00	5.00	5.00
Domain 2. Family Engagement	4.49	0.73	2.50	6.50	3.50	4.50	5.50
5. Parent-Child Interaction during Home Visit	4.52	0.79	1.50	6.50	3.50	4.50	5.50
1. To engage in interactions	4.73	1.26	1.00	7.00	3.00	5.00	6.00
2. To make contact with each other	4.47	0.95	1.00	7.00	3.00	5.00	5.00
3. To be available to engage in interactions	5.15	0.77	2.00	7.00	4.00	5.00	6.00
4. To observe and be ready to respond to the child's behavior	4.92	0.60	2.00	7.00	4.00	5.00	5.00
5. To respond to child and support child development	4.46	0.86	1.00	7.00	3.00	5.00	5.00
6. To adapt activities to child's interests and needs and encourage child engagement	3.21	0.89	1.00	7.00	2.00	3.00	4.00
7. To sustain positive interactions	4.62	1.02	1.00	7.00	3.00	5.00	6.00
6. Parent Engagement during Home Visit	4.13	0.83	1.50	6.50	3.00	4.00	5.00
1. To show interest in materials and activities	4.59	0.81	1.00	7.00	3.00	5.00	5.00
2. To participate and focus on home visit topics and activities	4.43	1.07	1.00	7.00	3.00	5.00	6.00
3. To engage in play and activities with child	4.48	1.14	3.00	7.00	3.00	5.00	6.00
4. To initiate activities and conversations	3.21	0.98	1.00	7.00	2.00	3.00	5.00
5. To discuss questions and topics relevant to child and family	3.29	0.89	1.00	7.00	2.00	3.00	5.00
6. To be ready to interact with both child and home visitor	4.93	0.80	3.00	7.00	4.00	5.00	6.00
7. Child Engagement during Home Visit	4.81	0.97	2.50	7.00	3.50	5.00	6.00
1. To participate in home visit activities	5.04	0.90	2.00	7.00	4.00	5.00	6.00
2. To initiate successful activities or interactions ¹	4.54	1.16	1.00	7.00	3.00	5.00	6.00
3. To sustain interactions with parent or home visitor	5.02	1.04	3.00	7.00	4.00	5.00	6.00
4. To show interest and enthusiasm about home visit activities	4.53	1.18	2.00	7.00	3.00	5.00	6.00
Total Score	3.90	0.56	2.21	5.71	3.21	3.93	4.57

Notes: ¹ Following the manual instructions, item 5 in scale 2 was only coded for the 92 visits for which other family members were present; and item 2 in scale 7 was only coded for the 547 visits with children 12 months or older. P10, P50 and P90 refer to the 10th, 50th and 90th percentile of the sample, respectively.

Table 4: Quality assessment of Cuna Mas visits according to the HOVRS-A+2 quality categories

	1. Needs training or support (1≤score<3)	2. Adequate (3≤score<5)	3. Good (5≤score<7)	4. Excellent (score = 7)
Domain 1. Home Visitor Practices	0.16	0.84	0.01	
1. Responsiveness to Family	0.39	0.61		
2. Relationship with Family	0.07	0.86	0.07	
3. Facilitation of Parent–Child Interaction	0.16	0.82	0.03	
4. Non-Intrusiveness and Collaboration	0.05	0.92	0.03	
Domain 2. Family Engagement	0.01	0.72	0.27	
5. Parent–Child Interaction during Home Visit	0.02	0.57	0.41	
6. Parent Engagement during Home Visit	0.04	0.74	0.22	
7. Child Engagement during Home Visit	0.01	0.43	0.54	0.02
Total score	0.04	0.93	0.03	

Notes: proportion of visits in each quality category.

Table 5: Structure and Content of the Visit from the supervisor-administered checklist

	N	Mean/ Proportion	SD	Possible range
Panel I: Content and Planning				
Welcome and review of previous visit				
18. Did the home visitor ask the main caregiver if any of the activities demonstrated in the previous week were done	554	0.97		[0-1]
19. Did the caregiver and the child demonstrate the activities from the previous week	552	0.87		[0-1]
Family life (moment 1)				
22. Did the home visitor bring the toys and materials planned for the visit	553	0.93		[0-1]
23. How well did the home visitor explain the objective of the activities shown in the current week	553	3.82	1.04	[1-5]
24. Did the home visitor give the child the opportunity to explore the material before beginning the activity	244	0.80		[0-1]
25. How well did the home visitor demonstrate the activities to the main caregiver	451	3.84	0.96	[1-5]
26. How well did the home visitor engage the main caregiver and child in the activities	554	4.01	0.94	[1-5]
27. Did the home visitor allow the main caregiver and child to try the activity on their own	554	0.77		[0-1]
28. While main caregiver and child were doing the activity, how often did the home visitor comment/label/respond to what the main caregiver and child were doing	428	3.74	1.09	[1-5]
29. Did the home visitor carry out all the steps of the activity	419	0.85		[0-1]
30. Was the activity adequate to the developmental level of the child	553	0.95		[0-1]
31. Was the activity of interest to the child	551	0.83		[0-1]
Learning through play (moment 2)				
22. Did the home visitor bring the toys and materials planned for the visit	552	1.00		[0-1]
23. How well did the home visitor explain the objective of the activities shown in the current week	550	3.89	1.01	[1-5]
24. Did the home visitor give the child the opportunity to explore the material before beginning the activity	554	0.93		[0-1]
25. How well did the home visitor demonstrate the activities to the main caregiver	553	3.86	1.00	[1-5]
26. How well did the home visitor engage the main caregiver and child in the activities	554	4.18	0.76	[1-5]
27. Did the home visitor allow the main caregiver and child to try the activity on their own	554	0.92		[0-1]
28. While main caregiver and child were doing the activity, how often did the home visitor comment/label/respond to what the main caregiver and child were doing	509	3.86	1.01	[1-5]
29. Did the home visitor carry out all the steps of the activity	553	0.92		[0-1]
30. Was the activity adequate to the developmental level of the child	554	1.00		[0-1]
31. Was the activity of interest to the child	551	0.95		[0-1]
Tell me a story (moment 3)				
22. Did the home visitor bring the toys and materials planned for the visit	546	0.99		[0-1]
23. How well did the home visitor explain the objective of the activities shown in the current week	539	3.63	1.12	[1-5]
24. Did the home visitor give the child the opportunity to explore the material before beginning the activity	541	0.82		[0-1]
25. How well did the home visitor demonstrate the activities to the main caregiver	547	3.81	1.01	[1-5]
26. How well did the home visitor engage the main caregiver and child in the activities	549	4.01	0.89	[1-5]
27. Did the home visitor allow the main caregiver and child to try the activity on their own	553	0.87		[0-1]
28. While main caregiver and child were doing the activity, how often did the home visitor comment/label/respond to what the main caregiver and child were doing	479	3.66	1.08	[1-5]
29. Did the home visitor carry out all the steps of the activity	545	0.84		[0-1]
30. Was the activity adequate to the developmental level of the child	548	0.98		[0-1]
31. Was the activity of interest to the child	547	0.83		[0-1]
Closure and recap				
32. Did the home visitor transition easily from one activity to the other:				
A. From Family life to Learn through play	552	0.94		[0-1]
B. From Learn through play to Tell me a story	547	0.91		[0-1]
33. Did the home visitor ask the main caregiver to recall the activities she would be carrying out during the week with the child	554	0.68		[0-1]
34. Did the home visitor ask the caregiver to provide some examples of how she could use the games and activities carried out during the visit in her daily life	353	0.30		[0-1]
35. Did the home visitor ask the main caregiver what might get in the way of doing the activities	354	0.10		[0-1]
36. Did the home visitor ask the main caregiver what else they could do during the week	523	0.18		[0-1]

Panel II: Interactions

37. How often did the home visitor ask the main caregiver's opinion about things like: familiarity with the games and activities proposed, games the child usually plays, what the child can do, what the mother likes to do with child or the whole family enjoy doing together, etc.	550	2.19	1.27	[1-5]
38. How often did the home visitor show interest in what the main caregiver answered	553	3.68	1.24	[1-5]
39. How often did the home visitor build on the information provided by the main caregiver during the visit	551	3.03	1.39	[1-5]
40. How often did the home visitor address the child by her name	553	4.59	0.70	[1-5]
41. How often did the home visitor get down to the child's level to interact with her	554	3.75	0.52	[1-5]
42. How often did the home visitor respond to the child's vocalizations and gestures	554	4.34	0.82	[1-5]
43. How often did the home visitor praise/encourage the child for her attempts (to do things) during the activity	554	4.11	0.98	[1-5]
44. How often did the home visitor prompt the caregiver to respond to and/or praise the child	554	3.40	1.29	[1-5]
45. How often did the home visitor praise/encourage the main caregiver during the activity for his/her efforts	554	3.46	1.24	[1-5]
46. Overall, was the home visitor's attitude towards the main caregiver:				
Encouraging	552	0.89		[0-1]
Intrusive/bossy	552	0.03		[0-1]
Indifferent	552	0.04		[0-1]
Shy	552	0.03		[0-1]
Other	552	0.01		[0-1]

Panel III: Overall assessment of the child and caregiver's participation

47. How often did the child actively participate during the visit	552	3.34	0.76	[1-5]
48. How often did the caregiver actively participate during the visit	553	4.31	0.79	[1-5]
49. How often did the home visitor and caregiver discuss topics unrelated to visit objectives	553	1.26	0.64	[1-5]
50. How often was the home visitor taking care of the child alone during the visit	550	1.18	0.70	[1-5]
51. How often was the visit fun	553	4.20	0.80	[1-5]
52. How often was the visit conflictive/tense	553	1.20	0.53	[1-5]

Notes: Items scored on a scale from 1 to 5 in Panel I correspond to 1 "did not explain/demonstrate/engage" to 5 "very good explanation/demonstration/engagement" except item 28 (only administered when the home visitor gave the child and the caregiver the opportunity to carry out the activity on their own) scored as 1 'almost never commented, labelled or responded' to 5 'almost always commented, labelled or responded'. Items 34 and 35 were only administered when the home visitor asked the main caregiver to recall the activities she would be carrying out during the week with the child. Items in Panels II and III (except item 46) are rated on a scale from 1 to 5 as 1 'almost never' to 5 'almost always'.

Table 6: Internal consistency and validity of the HOVRS-A+2

	Alpha (sample)	Apha (manual)	r	Eigenvalue	Factor Load
Domain 1. Home Visitor Practices					
1. Responsiveness to Family	0.70	0.76	0.69***	1.75	
1. To plan activities and topics of the home visit with the parent	0.70		0.37***		0.35
2. To prepare for the home visit using parent-selected activities	0.71		0.25***		0.26
3. To get information about the family's strengths and child's development	0.62		0.77***		0.65
4. To provide feedback on family strengths for supporting child development	0.59		0.73***		0.70
5. To adapt activities to the family's interests and needs	0.63		0.64***		0.60
6. To respond to family input for the agenda and activities of the home visit	0.65		0.55***		0.53
2. Relationship with Family	0.83	0.70	0.81***	3.09	
1. To interact sociably with parent(s), focusing on child development	0.77		0.79***		0.81
2. To set the tone for positive interactions	0.78		0.80***		0.88
3. To express positive emotions about the home visit	0.78		0.77***		0.83
5. To engage other family members if present during the home visit ¹	0.82		0.60***		0.36
7. To reflect on family's life and activities in relation to child's development	0.81		0.66***		0.55
8. To show respect and acceptance of the family, home, culture, and lifestyle	0.84		0.46***		0.46
9. To discuss sensitive issues respectfully and reflectively	0.82		0.64***		0.56
3. Facilitation of Parent-Child Interaction	0.81	0.85	0.81***	2.55	
1. To elicit ongoing parent-child interactions during the home visit	0.76		0.73***		0.72
2. To promote developmentally supportive interactions	0.76		0.74***		0.72
3. To engage parent and child together	0.78		0.71***		0.65
4. To support parent responsiveness to child cues	0.76		0.74***		0.73
5. To directly encourage or reinforce positive parent-child interactions	0.77		0.77***		0.71
6. To help parents use available resources to support child development	0.84		0.22***		0.18
4. Non-Intrusiveness and Collaboration	0.65	0.87	0.71***	1.42	
1. To encourage the parent's ideas and interests for interactions with child	0.61		0.50***		0.52
2. To avoid intruding on or ignoring parent-child interactions	0.55		0.66***		0.63
3. To keep parent in the "teacher" role	0.59		0.63***		0.48
4. To follow the lead of parent and child in pace and activities	0.66		0.57***		0.37
5. To allow parent-child interactions to continue uninterrupted	0.56		0.66***		0.61
Domain 2. Family Engagement					
5. Parent-Child Interaction during Home Visit	0.88	0.95	0.84***	3.72	
1. To engage in interactions	0.85		0.84***		0.85
2. To make contact with each other	0.86		0.77***		0.79
3. To be available to engage in interactions	0.87		0.72***		0.71
4. To observe and be ready to respond to the child's behavior	0.87		0.59***		0.59
5. To respond to child and support child development	0.88		0.75***		0.71
6. To adapt activities to child's interests and needs and encourage child engagement	0.87		0.69***		0.65
7. To sustain positive interactions	0.86		0.79***		0.78
6. Parent Engagement during Home Visit	0.91	0.88	0.80***	3.74	
1. To show interest in materials and activities	0.90		0.76***		0.73
2. To participate and focus on home visit topics and activities	0.87		0.86***		0.90
3. To engage in play and activities with child	0.88		0.85***		0.88
4. To initiate activities and conversations	0.88		0.84***		0.82
5. To discuss questions and topics relevant to child and family	0.90		0.77***		0.72
6. To be ready to interact with both child and home visitor	0.91		0.70***		0.66
7. Child Engagement during Home Visit	0.88	0.88	0.71***	2.52	
1. To participate in home visit activities	0.84		0.83***		0.81
2. To initiate successful activities or interactions ¹	0.86		0.81***		0.74
3. To sustain interactions with parent or home visitor	0.84		0.85***		0.79
4. To show interest and enthusiasm about home visit activities	0.83		0.85***		0.83

Notes: ¹Following the manual instructions, item 5 in scale 2 was only coded for the 92 visits for which other family members were present; and item 2 in scale 7 was only coded for the 547 visits with children 12 months or older. As such, the eigenvalues and factor loads for those two scales are computed on that set of observations. Results are robust to excluding these items from the analysis and recomputing the statistics in the table using all 554 observations available for the remaining items. "Alpha" refers to the Cronbach's alphas. For each scale, the alpha is the internal consistency of the scale including all items that compose it. The alphas reported next to each item are what the alpha for the scale would be if such item was not included. "r" refers to the Pearson correlation coefficients between each item with the scale score and each scale with the domain score, respectively. Coefficients significant at * p<0.10,** p<0.05, and *** p<0.01.

Table 7: Correlations between the HOVRS-A+2 and other variables

	HOVRS-A+2									
	Domain 1. Home Visitors Practices	1. Responsiveness to Family	2. Relationship with Family	3. Facilitation of Parent-Child Interaction	4. Non-Intrusiveness and Collaboration	Domain 2. Family Engagement	5. Parent-Child Interaction during Home Visit	6. Parent Engagement during Home Visit	7. Child Engagement during Home Visit	Total score
Panel I: Supervisor-administered checklist										
Home Visitor										
Time in Program (months)	0.08	0.07	0.06	0.11	0.01	-0.03	-0.02	0.00	-0.04	0.03
Time working with family (months)	0.03	0.03	0.02	0.02	0.02	-0.02	0.00	-0.01	-0.04	0.00
Child										
Time in Program (months)	0.08*	0.06	0.04	0.02	0.14***	0.09**	0.07	0.09**	0.07	0.09**
Child has had at least one other home visitor since enrolment	0.03	0.02	0.01	-0.02	0.08**	0.08*	0.07*	0.09**	0.03	0.06
Age (months)	0.05	0.03	0.02	0.05	0.06	0.05	-0.06	-0.02	0.17***	0.05
Panel II: Home visitor and community surveys										
Home Visitor										
Age (years)	-0.06	-0.04	-0.09	-0.01	-0.07	-0.06	-0.03	0.01	-0.12	-0.06
Years of education	0.33***	0.29***	0.32***	0.32***	0.22***	0.31***	0.34***	0.38***	0.11	0.34***
Has experience with children in the 0-3 age range	0.04	0.07	0.02	0.00	0.07	-0.04	-0.03	-0.01	-0.05	0.00
Has experience working with families	0.24***	0.21***	0.18**	0.24***	0.21***	0.16*	0.16**	0.09	0.16*	0.21***
Knowledge of Infant Development Inventory (KIDI) score	0.11	0.10	0.12	0.13	0.01	-0.02	-0.04	0.06	-0.05	0.05
Panel III: HVCCF										
Duration of the visit (minutes)	0.24***	0.20***	0.26***	0.21***	0.14***	0.22***	0.19***	0.20***	0.17***	0.25***
Extent to which environmental distractions (television, phone calls, visitors, pets, other children, noise, etc.) interfered with the home visit.	-0.21***	-0.14***	-0.16***	-0.21***	-0.20***	-0.24***	-0.23***	-0.26***	-0.13***	-0.25***
Was the visit fun?	0.59***	0.39***	0.56***	0.56***	0.44***	0.66***	0.53***	0.42***	0.71***	0.69***
Was the visit conflictive/tense?	-0.44***	-0.24***	-0.41***	-0.41***	-0.39***	-0.52***	-0.39***	-0.34***	-0.56***	-0.52***

Notes: Pearson correlation coefficients between scales, domains and the total scores of the HOVRS-A+2, with other variables from three different sources: Panel I includes variables from the checklist, Panel II from the visitor and community surveys, and Panel III from the Home Visit Content and Characteristics Form. Coefficients significant at * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 8: Internal consistency and validity of the “constructs” in the checklist

	Alpha	r	Eigenvalue	Loading
I. Responsiveness to Family (N=464)	0.81		2.37	
28. While main caregiver and child were doing the activity, how often did the home visitor comment/label/respond to what the main caregiver and child were doing (moment 2)	0.78	0.77***		0.70
28. While main caregiver and child were doing the activity, how often did the home visitor comment/label/respond to what the main caregiver and child were doing (moment 3)	0.77	0.82***		0.74
37. How often did the home visitor ask the main caregiver’s opinion about things like: familiarity with the games and activities proposed, games the child usually plays, what the child can do, what the mother likes to do with child or the whole family enjoy doing together, etc.	0.82	0.53***		0.48
38. How often did the home visitor show interest in what the main caregiver answered	0.74	0.81***		0.74
39. How often did the home visitor build on the information provided by the main caregiver during the visit	0.75	0.82***		0.75
II. Relationship with Family (N=548)	0.86		3.59	
42. How often did the home visitor respond to the child’s vocalizations and gestures	0.86	0.54***		0.51
43. How often did the home visitor praise/encourage the child for her attempts (to do things) during the activity	0.84	0.75***		0.71
26. How well did the home visitor engage the main caregiver and child in the activities (moment 1)	0.85	0.63***		0.59
26. How well did the home visitor engage the main caregiver and child in the activities (moment 2)	0.85	0.71***		0.67
26. How well did the home visitor engage the main caregiver and child in the activities (moment 3)	0.85	0.69***		0.64
38. How often did the home visitor show interest in what the main caregiver answered	0.84	0.74***		0.69
44. How often did the home visitor prompt the caregiver to respond to and/or praise the child	0.84	0.80***		0.75
45. How often did the home visitor praise/encourage the main caregiver during the activity for his/her efforts	0.83	0.83***		0.78
III. Facilitation of Parent–Child Interaction (N=470)	0.86		3.47	
26. How well did the home visitor involve the caregiver and the child during the activity (moment 1)	0.85	0.60***		0.56
26. How well did the home visitor involve the caregiver and the child during the activity (moment 2)	0.84	0.76***		0.72
26. How well did the home visitor involve the caregiver and the child during the activity (moment 3)	0.84	0.78***		0.73
28. While main caregiver and child were doing the activity, how often did the home visitor comment/label/respond to what the main caregiver and child were doing (moment 2)	0.84	0.80***		0.75
28. While main caregiver and child were doing the activity, how often did the home visitor comment/label/respond to what the main caregiver and child were doing (moment 3)	0.84	0.79***		0.74
44. How often did the home visitor prompt the caregiver to respond to and/or praise the child	0.84	0.74***		0.70
45. How often did the home visitor praise/encourage the main caregiver during the activity for his/her efforts	0.83	0.76***		0.72
IV. Child and caregiver engagement during visit (N=542)	0.59		1.06	
31. Was the activity of interest to the child (moment 1)	0.57	0.53***		0.41
31. Was the activity of interest to the child (moment 3)	0.59	0.52***		0.40
47. How often did the child actively participate during the visit	0.32	0.89***		0.68
48. How often did the caregiver actively participate during the visit	0.50	0.68***		0.52
V. Key practices during visit moments (N=533)	0.86		3.81	
23. How well did the home visitor explain the objective of the activities shown in the current week (moment 1)	0.84	0.72***		0.68
23. How well did the home visitor explain the objective of the activities shown in the current week (moment 2)	0.84	0.77***		0.72
23. How well did the home visitor explain the objective of the activities shown in the current week (moment 3)	0.84	0.76***		0.71
25. How well did the home visitor demonstrate the activities to the main caregiver (moment 2)	0.85	0.69***		0.65
25. How well did the home visitor demonstrate the activities to the main caregiver (moment 3)	0.85	0.65***		0.61
26. How well did the home visitor engage the main caregiver and child in the activities (moment 1)	0.85	0.67***		0.63
26. How well did the home visitor engage the main caregiver and child in the activities (moment 2)	0.84	0.74***		0.69
26. How well did the home visitor engage the main caregiver and child in the activities (moment 3)	0.85	0.68***		0.64
51. How often was the visit fun	0.86	0.54***		0.50

Notes: The scores for each “construct” of the checklist were computed by factor analysis. Alpha” refers to the Cronbach’s alphas. For each “construct”, the alpha is the internal consistency of such construct including all items that compose it. The alphas reported next to each item are what the alpha for the “construct” would be if such item was not included. “r” refers to the Pearson correlation coefficients between each item with the scale score and each scale with the domain score, respectively. Coefficients significant at * p<0.10, ** p<0.05, and *** p<0.01.

Table 9: Correlations between the ‘constructs’ in the checklist and the HOVRS-A+2

'Constructs' in the checklist	HOVRS-A+2							Total score		
	Domain 1. Home Visitors Practices	1. Responsiveness to Family	2. Relationship with Family	3. Facilitation of Parent–Child Interaction	4. Non-Intrusiveness & Collaboration	Domain 2. Family Engagement	5. Parent–Child Interaction during Home Visit		6. Parent Engagement during Home Visit	7. Child Engagement during Home Visit
I. Responsiveness to Family	0.36***	0.30***	0.38***	0.34***	0.16***	0.23***	0.21***	0.23***	0.15***	0.32***
II. Relationship with Family	0.35***	0.26***	0.35***	0.35***	0.19***	0.32***	0.29***	0.26***	0.26***	0.36***
III. Facilitation of Parent–Child Interaction	0.36***	0.28***	0.36***	0.36***	0.18***	0.27***	0.23***	0.23***	0.21***	0.34***
IV. Child and caregiver engagement during visit	0.22***	0.14***	0.21***	0.20***	0.18***	0.36***	0.27***	0.22***	0.39***	0.31***
V. Key practices during visit moments	0.38***	0.31***	0.38***	0.35***	0.21***	0.32***	0.30***	0.26***	0.25***	0.38***
VI. Stand-alone items										
19. Demonstration from previous week	0.19***	0.12***	0.18***	0.20***	0.14***	0.14***	0.13***	0.13***	0.11**	0.18***
29. All the steps of activity carried out (moment 3)	0.16***	0.10**	0.15***	0.15***	0.11***	0.15***	0.14***	0.13***	0.12***	0.17***
33. Recall of activities during the week	0.22***	0.22***	0.19***	0.17***	0.14***	0.16***	0.14***	0.16***	0.11***	0.21***
36. Other activities for the week	0.22***	0.26***	0.19***	0.21***	0.06	0.14***	0.13***	0.15***	0.08*	0.19***
46. Home visitor's attitude is encouraging	0.23***	0.19***	0.23***	0.20***	0.12***	0.18***	0.18***	0.19***	0.10**	0.22***

Notes: The scores for each ‘construct’ of the checklist were computed using factor analysis. Stand-alone items are items that cannot be mapped to any of the HOVRS-A+2 constructs. Coefficients significant at * p<0.10, ** p<0.05, and *** p<0.01.

Table 10: Correlations between the ‘constructs’ and stand-alone items in the checklist and other variables

	'Constructs' in the checklist					Stand-alone items				
	I. Responsiveness to Family	II. Relationship with Family	III. Facilitation of Parent-Child Interaction	IV. Child and caregiver engagement during visit	V. Key practices during visit moments	19. Demonstration from previous week	29. All the steps of activity carried out (moment 3)	33. Recall of activities during the week	36. Other activities for the week	46. Home visitor's attitude is encouraging
Panel I: Supervisor-administered checklist										
Home Visitor										
Time in Program (months)	0.03	-0.02	0.02	-0.18**	-0.01	0.11	0.03	0.17**	0.14*	-0.02
Time working with family (months)	0.04	-0.06	0.02	-0.084**	-0.02	0.05	0.03	0.11**	0.00	-0.03
Child										
Time in Program (months)	-0.03	-0.06	-0.05	0.02	-0.03	0.02	-0.03	0.01	-0.04	0.04
Child has had at least one other home visitor since enrolment	0.01	0.08*	0.02	0.16***	0.06	0.00	-0.01	-0.05	-0.06	0.10**
Age (months)	0.05	0.02	0.03	0.11**	-0.02	0.06	0.05	0.04	0.00	0.01
Panel II: Home visitor and community surveys										
Home Visitor										
Age (years)	-0.07	-0.01	-0.02	-0.08	-0.04	-0.04	0.00	0.00	0.06	-0.01
Years of education	0.32***	0.26***	0.26***	0.10	0.24***	0.10	0.12	0.10	0.18**	0.18**
Has experience with children in the 0-3 age range	0.11	0.07	0.06	0.04	0.02	0.09	0.02	0.02	-0.09	0.03
Has experience working with families	0.00	0.06	0.02	-0.03	0.05	-0.01	-0.03	-0.02	-0.03	-0.07
Knowledge of Infant Development Inventory (KIDI) score	0.15*	0.07	0.13	-0.02	0.14*	0.05	0.09	0.12	0.08	-0.02
Panel III: HVCCF										
Duration of the visit (minutes)	0.07	0.07	0.09*	0.07	0.12***	0.18***	0.11***	0.11***	0.09**	0.06
Extent to which environmental distractions (television, phone calls, visitors, pets, other children, noise, etc.) interfered with the home visit.	-0.12***	-0.12***	-0.15***	-0.07	-0.15***	-0.05	0.00	-0.14***	-0.03	-0.14***
Was the visit fun?	0.20***	0.25***	0.22***	0.34***	0.28***	0.13***	0.17***	0.15***	0.07	0.12***
Was the visit conflictive/tense?	-0.13***	-0.18***	-0.14***	-0.29***	-0.17***	-0.17***	-0.13***	-0.14***	0.00	-0.05

Notes: Pearson correlation coefficients between the 'constructs' of the checklist and stand-alone items but that did not fit any of the HOVRS-A+2 constructs, with other variables from three different sources: Panel I includes general information about the visit from the checklist, Panel II variables from the visitor and community surveys, and Panel III data from the Home Visit Content and Characteristics Form. The scores for each "construct" of the checklist were computed by factor analysis. Coefficients significant at * p<0.10, ** p<0.05, and *** p<0.01.

Appendix I. Sample selection for quality analysis

The home visiting component of Cuna Mas was subject to a rigorous experimental impact evaluation carried out during the first two years of Program activities, from mid-to-late 2013 to mid-to-late 2015. The impact evaluation sample includes a total of 5,859 children ages 1-24 months at baseline in 360 communities in 180 districts (120 treatments and 60 controls) in 12 departments. DGSE-MIDIS et al. (2015) offers more details on the evaluation design and in particular, on the selection of the evaluation sample and the allocation of treatment, which was randomized at the district level stratifying in blocks of three districts each. Within each district, the two communities with the largest number of kids 0-24 months old according to the information system for targeting social programs (*Sistema Focalización Hogares*, or SISFOH) were included in the sample. Next, within each block of three districts, the three communities (i.e. one per district) with the largest number of children were linked together, forming a ‘trio of communities’. Similarly, the other three communities (one per district) with the second largest number of children were also linked together, forming another ‘trio of communities’.

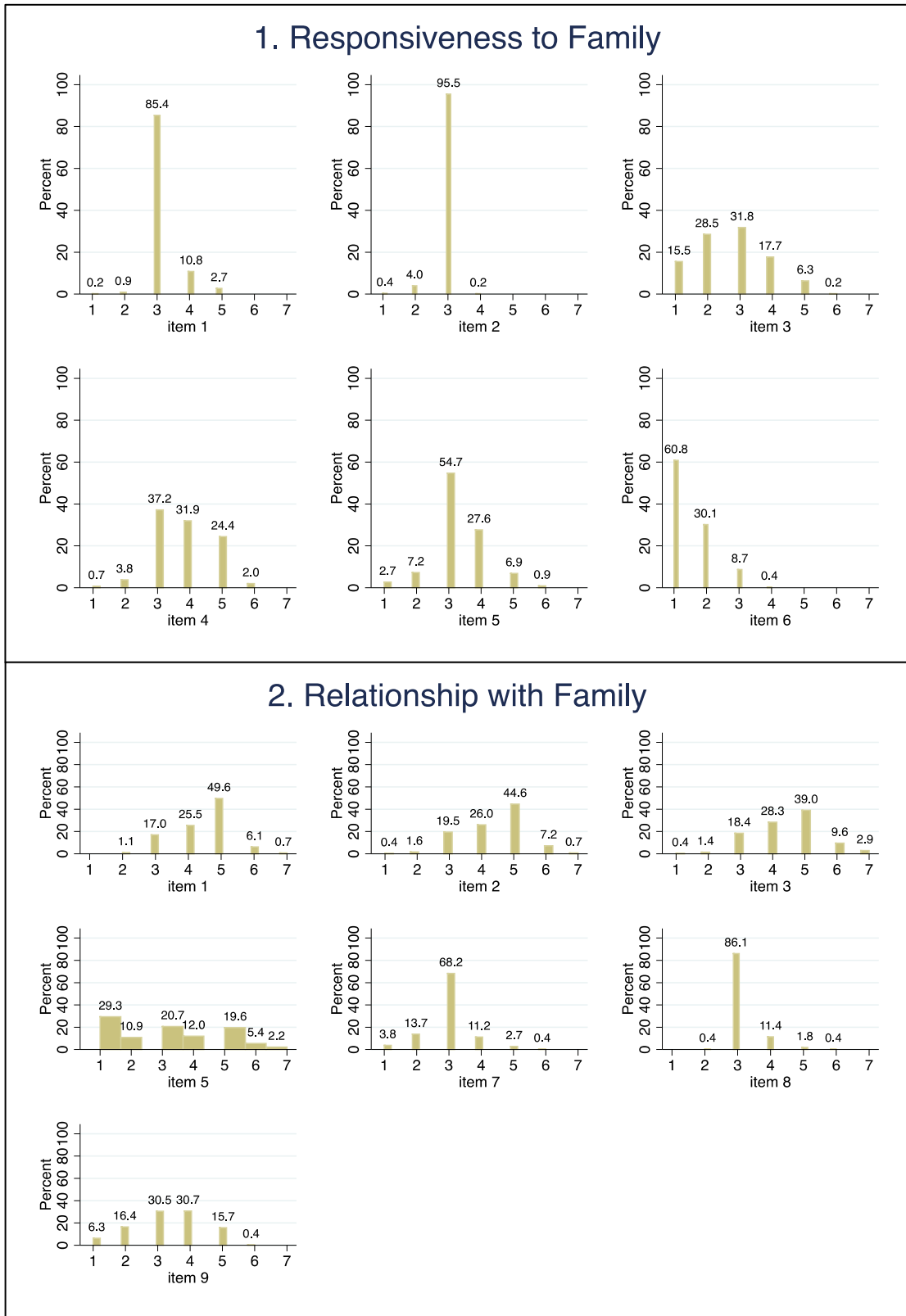
The set of communities in the quality sample were selected from the complete set of treatment communities in the impact evaluation sample taking into consideration budgetary and practical constraints, as described next. First, it excluded the trios of communities where at least one treatment community in the block had been initially assigned to the treatment group but later deemed ineligible by the Program to receive benefits because of social conflict, high dispersion, or due to the presence of Cuna Mas daycare program. In these communities, which represent 16.7 percent of those in the treatment group, no child was ever enrolled in the Program. Second, it excluded one trio of communities where one control community did receive the intervention. Third, it also excluded those trios of communities in which at least in one community over 30 percent of the households had responded the baseline survey in an indigenous language. This aimed to minimize the likelihood of a home visit taking place in a language different than Spanish and that of the child not understanding Spanish, which would pose difficulties for the assessment of developmental outcomes. As a result, the quality sample is distributed across nine of the 12 departments in the complete evaluation sample. Excluded departments were Amazonas, Loreto and Cusco since only a handful of communities in these areas satisfied all inclusion criteria for the quality analysis.

Within these communities, only children receiving visits—this is to say, younger than 36 months of age at the time of data collection—were included in the quality sample. We first selected eligible children within the complete evaluation sample and then identified their home visitor. We found a total of 176 home visitors. For each one of these home visitors, we collected information during their visits to 3 or 4 different children (4 children in 22% of cases), including the visit to the child in the complete evaluation sample.²³ The total quality sample includes information on the visits for 554 children, of which 229 were also part of the complete evaluation sample and 325 were children visited by the *same* home visitors but not included in the evaluation sample.

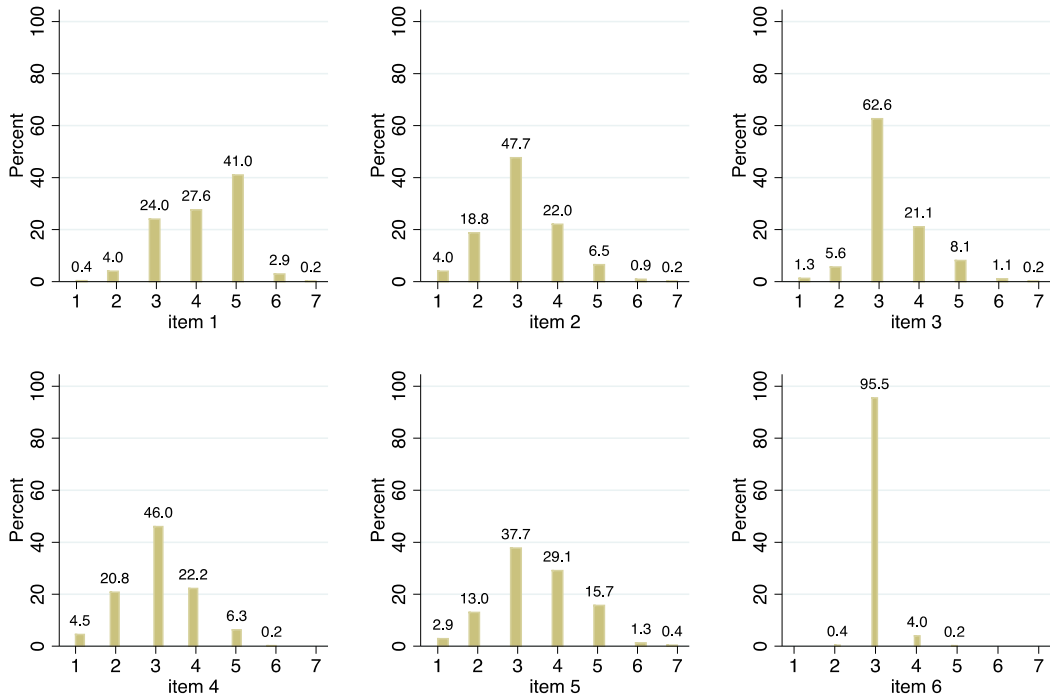
²³ For six home visitors (3% of the sample) we only have one filmed visit and for ten home visitors (6%) we only have video recordings from two visits.

Appendix II. Appendix Figures

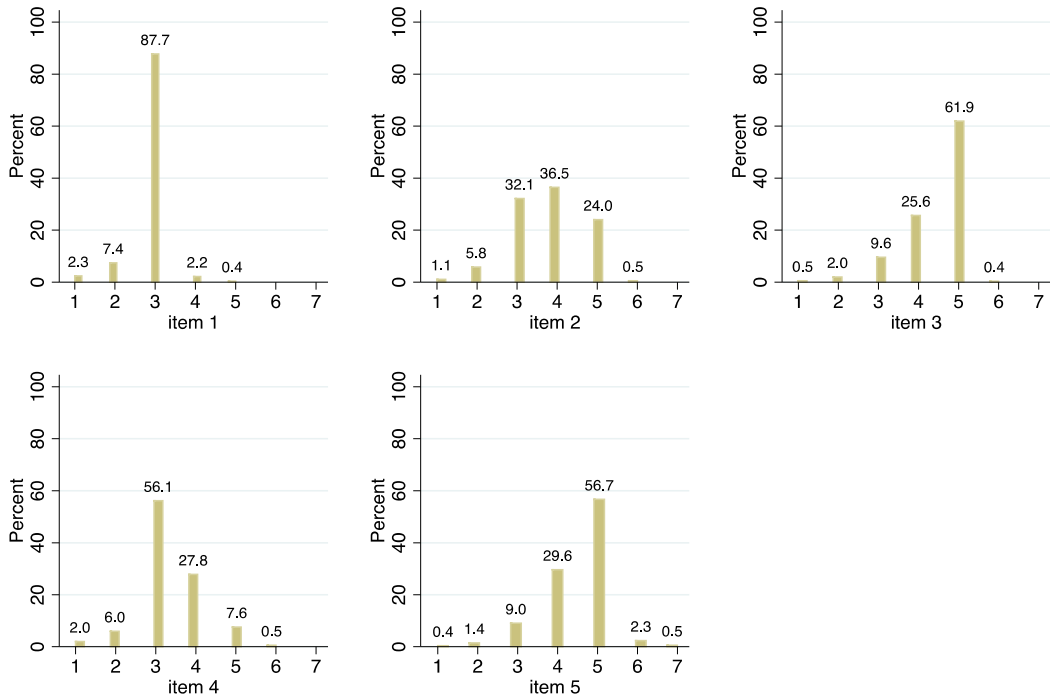
Figure A1: Item variability in the HOVRS-A+2, by scale



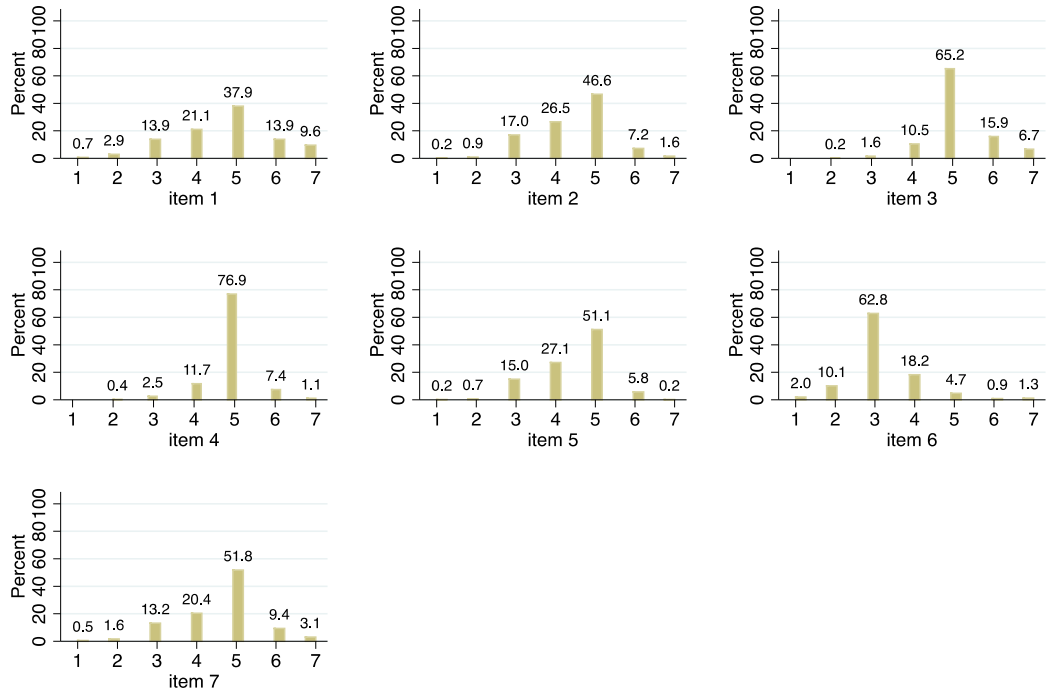
3. Facilitation of Parent–Child Interaction



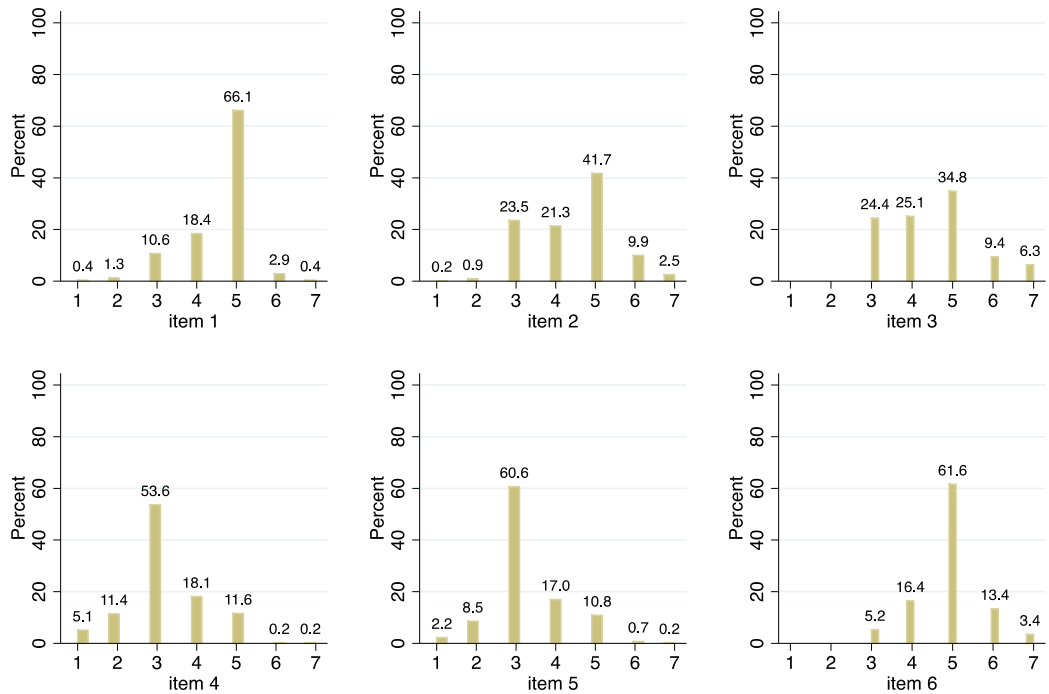
4. Non-Intrusiveness & Collaboration



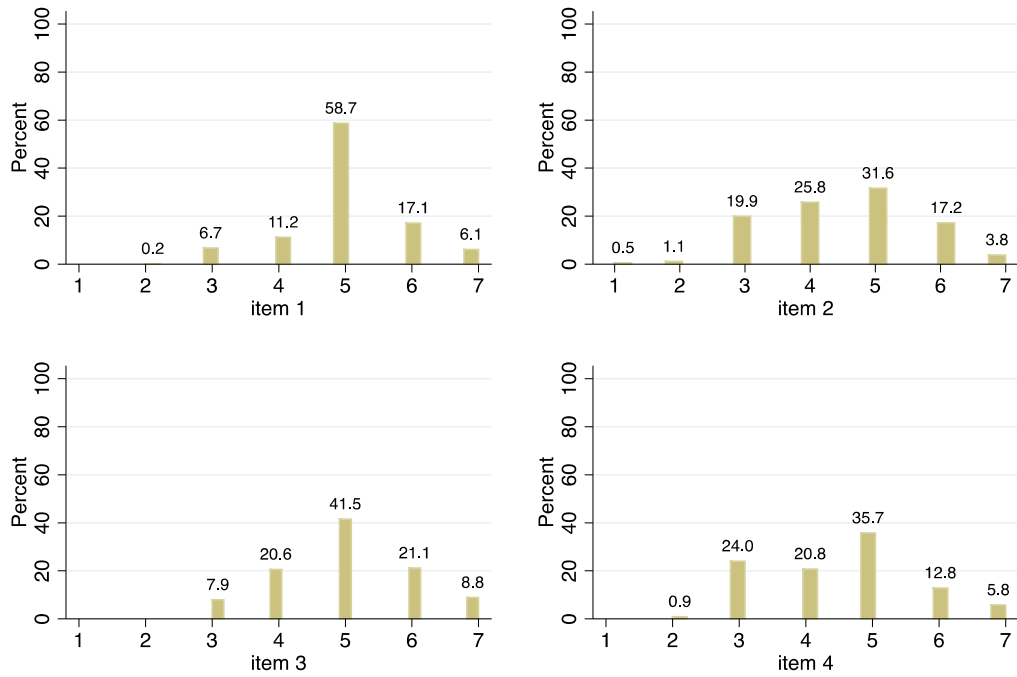
5. Parent-Child Interaction during Home Visit



6. Parent Engagement during Home Visit

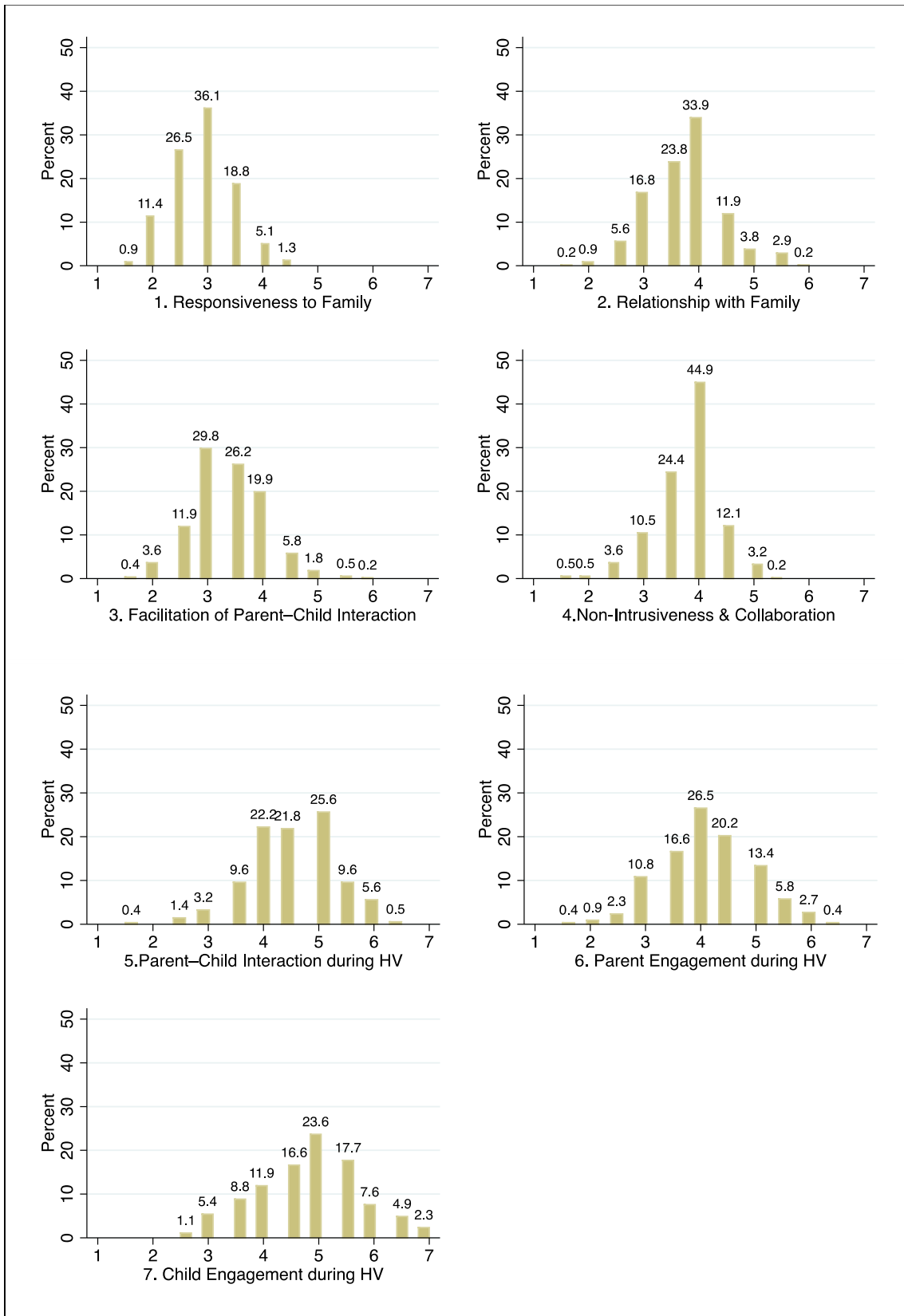


7. Child Engagement during Home Visit



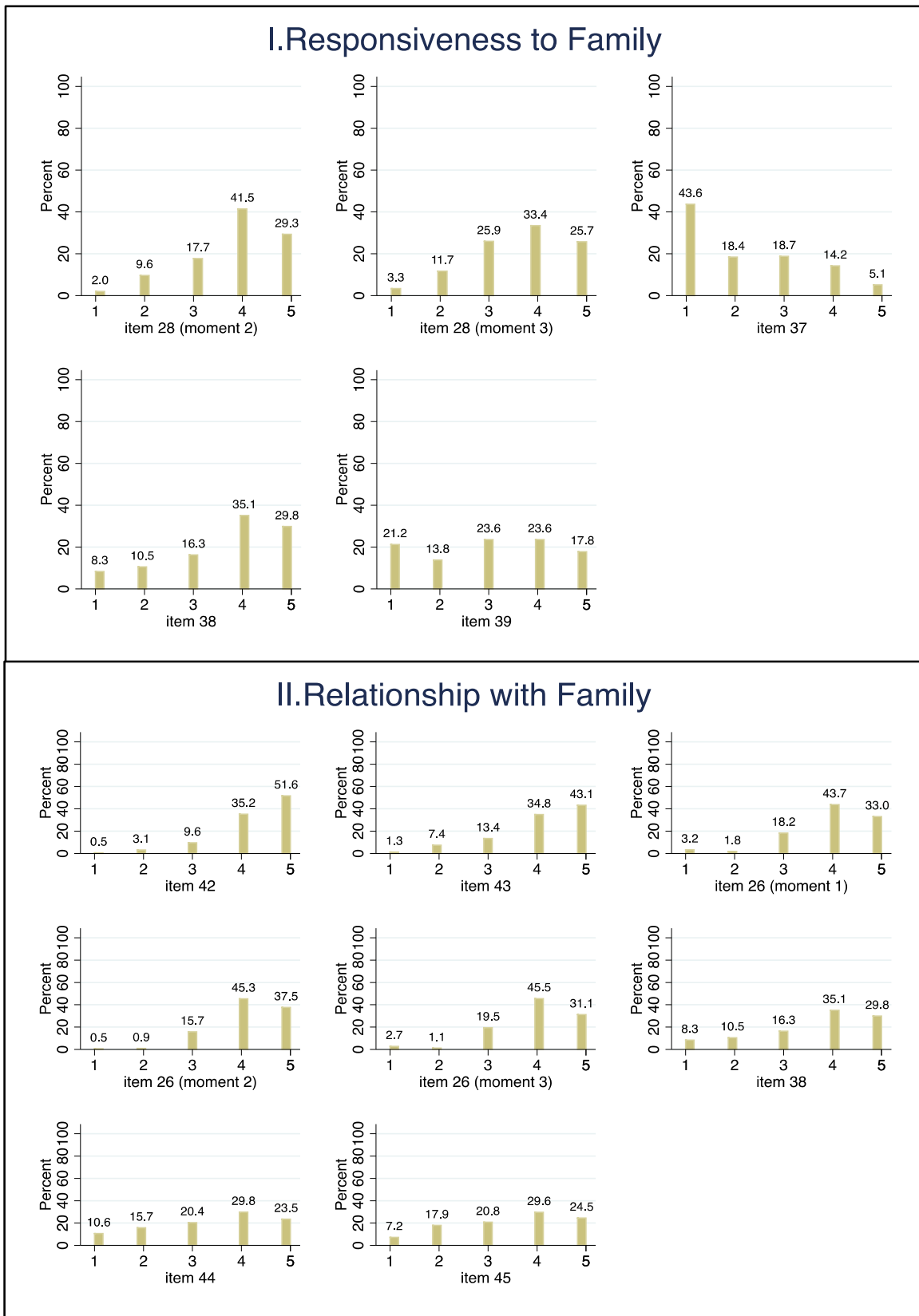
Notes: frequency densities for each item of the HOVRS-A+2 by scale. N=554 except for item 5 in scale 2 coded only for the 92 visits for which other family members were present; and item 2 in scale 7 coded only for the 547 visits with children 12 months or older. Each item is scored on a 7-point scale, with anchor points of 1 'needs training or support', 3 'adequate', 5 'good', and 7 'excellent'.

Figure A2: Variability of the scales' scores in the HOVRS-A+2

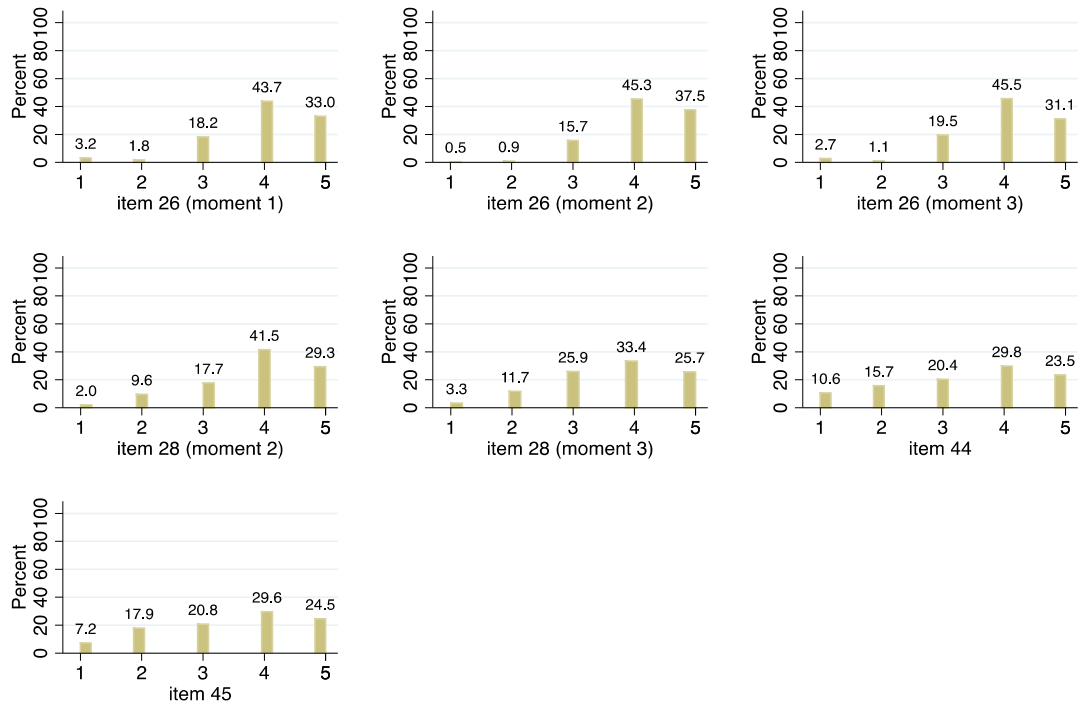


Notes: frequency densities for the scales' scores of the HOVRS-A+2.

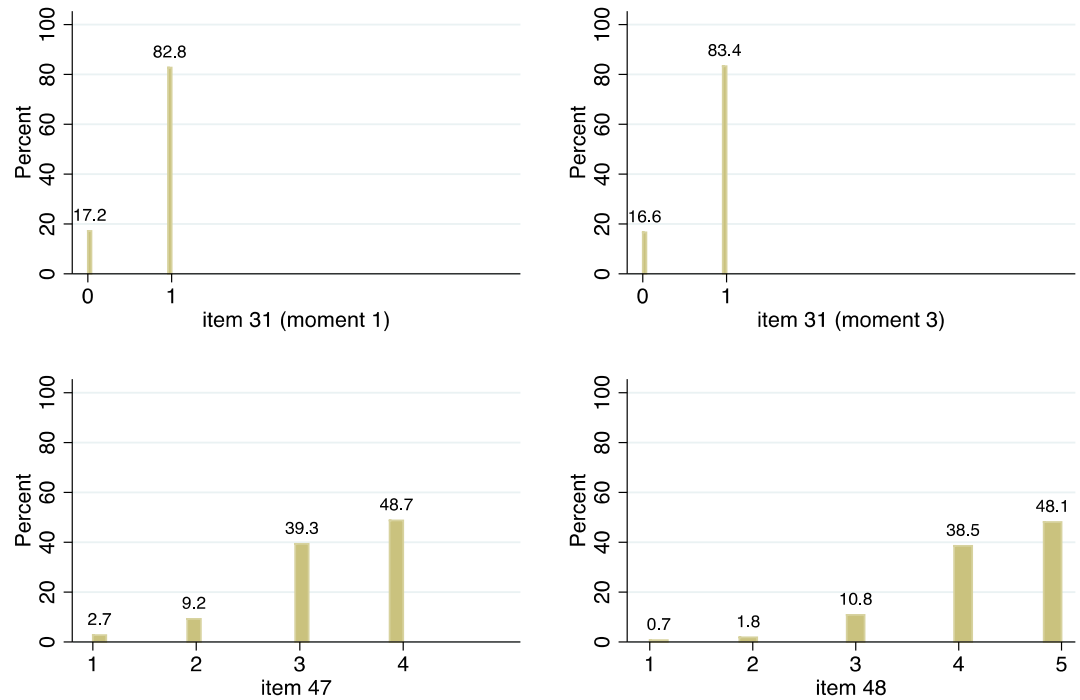
Figure A3: Item variability of the checklist, by 'construct'



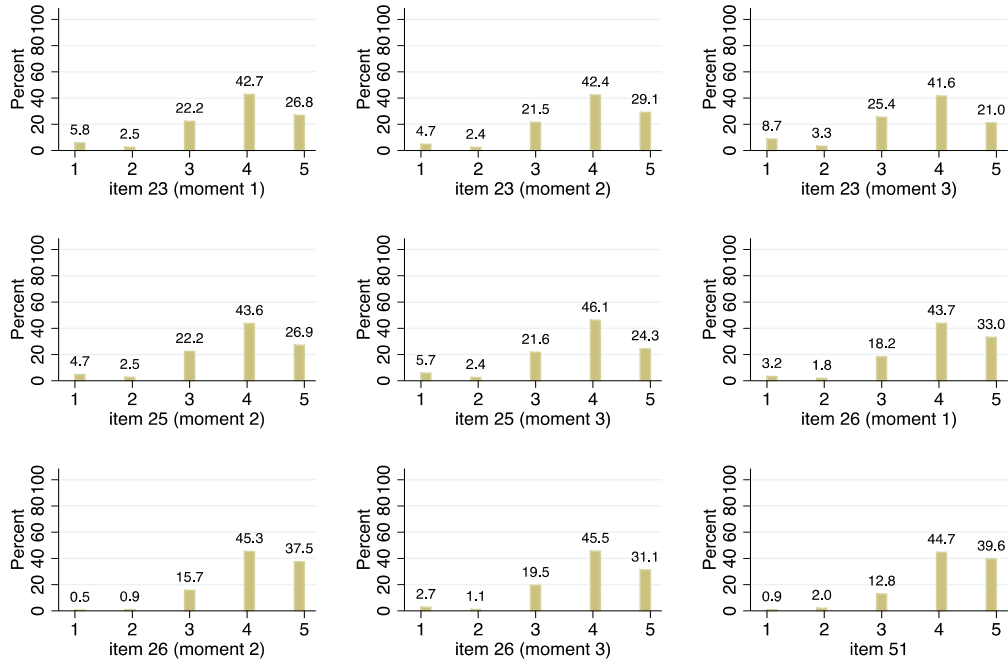
III. Facilitation of Parent–Child Interaction



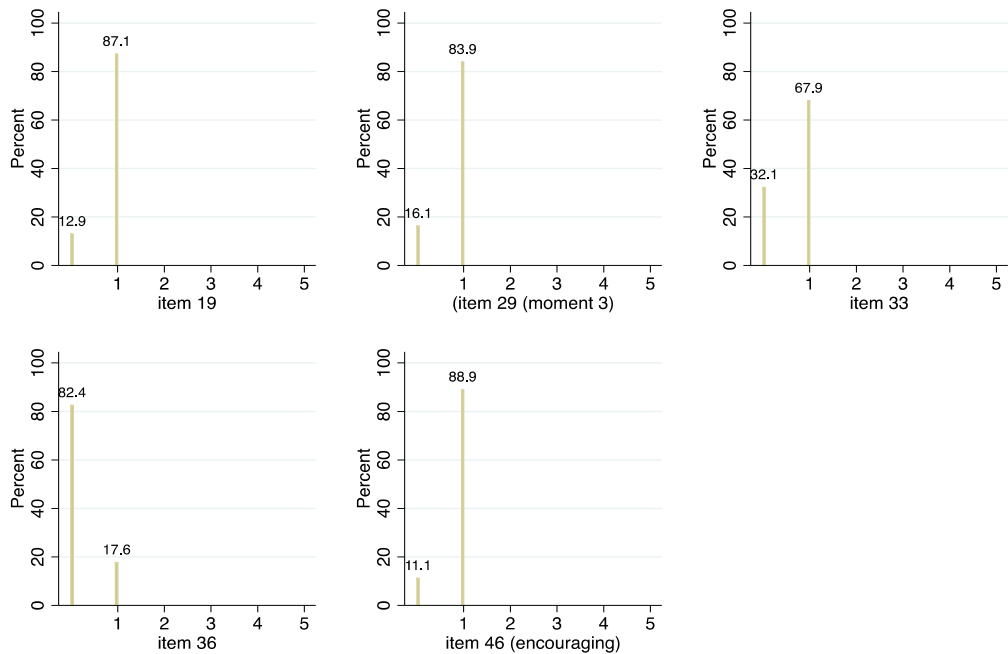
IV. Child and caregiver engagement



V. Key practices during visit moments



VI. Stand-alone items



Notes: Items mapped to constructs I, II, III are scored on a scale from 1 to 5 where 1 refers to “almost never” and 5 to “almost always”, except for items 26 in “constructs” II and III where the scores correspond to 1 “did not explain/demonstrate/engage”, 2 to “poor explanation/ demonstration/ engagement”, 3 “average”, 4 “good”, and 5 “very good explanation/demonstration/engagement.” For construct IV, items 31 are binary (0-1) and items 47 and 48 are scored on scale from 1 to 5 where 1 refers to “almost never” and 5 to “almost always.” All items in “construct” V are scored on a scale from 1 to 5, with the same categories as item 26 mentioned above. Stand-alone items are binary (0-1).

Appendix III. Appendix Tables

Table A1: Correlations between the items, scales, domains and total scores of the HOVRS-A+2

	Domain 1. Home Visitors Practices	1. Responsiveness to Family	2. Relationship with Family	3. Facilitation of Parent-Child Interaction	4. Non-Intrusiveness & Collaboration	Domain 2. Family Engagement	5. Parent-Child Interaction during Home Visit	6. Parent Engagement during Home Visit	7. Child Engagement during Home Visit	Total score
1. Responsiveness to Family	0.81***	1.00***	0.62***	0.64***	0.44***	0.47***	0.42***	0.44***	0.34***	0.69***
1. To plan activities and topics of the home visit with the parent	0.25***	0.37***	0.2***	0.18***	0.085**	0.13***	0.11**	0.1**	0.11**	0.2***
2. To prepare for the home visit using parent-selected activities	0.31***	0.25***	0.28***	0.27***	0.21***	0.17***	0.18***	0.22***	0.051	0.26***
3. To get information about the family's strengths and child's development	0.64***	0.77***	0.53***	0.53***	0.29***	0.33***	0.27***	0.33***	0.24***	0.52***
4. To provide feedback on family strengths for supporting child development	0.7***	0.73***	0.58***	0.62***	0.37***	0.42***	0.38***	0.36***	0.32***	0.6***
5. To adapt activities to the family's interests and needs	0.63***	0.64***	0.49***	0.48***	0.49***	0.45***	0.4***	0.36***	0.39***	0.59***
6. To respond to family input for the agenda and activities of the home visit	0.5***	0.55***	0.35***	0.42***	0.34***	0.42***	0.39***	0.41***	0.28***	0.5***
2. Relationship with Family	0.88***		1.00***	0.72***	0.53***	0.61***	0.58***	0.54***	0.44***	0.81***
1. To interact socially with parent(s), focusing on child development	0.75***	0.51***	0.79***	0.67***	0.49***	0.57***	0.55***	0.51***	0.4***	0.72***
2. To set the tone for positive interactions	0.73***	0.46***	0.8***	0.62***	0.5***	0.55***	0.56***	0.47***	0.39***	0.7***
3. To express positive emotions about the home visit	0.71***	0.48***	0.77***	0.65***	0.42***	0.56***	0.48***	0.44***	0.5***	0.69***
5. To engage other family members if present during the home visit	0.4***	0.2*	0.6***	0.19*	0.26**	0.24**	0.21**	0.22**	0.16	0.34***
7. To reflect on family's life and activities in relation to child's development	0.62***	0.52***	0.66***	0.5***	0.35***	0.36***	0.33***	0.37***	0.22***	0.53***
8. To show respect and acceptance of the family, home, culture, and lifestyle	0.5***	0.44***	0.46***	0.45***	0.32***	0.33***	0.3***	0.27***	0.26***	0.45***
9. To discuss sensitive issues respectfully and reflectively	0.57***	0.46***	0.64***	0.46***	0.3***	0.37***	0.34***	0.41***	0.2***	0.51***
3. Facilitation of Parent-Child Interaction	0.88***			1.00***	0.52***	0.61***	0.57***	0.54***	0.46***	0.81***
1. To elicit ongoing parent-child interactions during the home visit	0.7***	0.48***	0.6***	0.73***	0.5***	0.58***	0.54***	0.53***	0.41***	0.7***
2. To promote developmentally supportive interactions	0.72***	0.57***	0.58***	0.74***	0.49***	0.45***	0.44***	0.37***	0.34***	0.64***
3. To engage parent and child together	0.65***	0.44***	0.57***	0.71***	0.4***	0.53***	0.48***	0.49***	0.39***	0.64***
4. To support parent responsiveness to child cues	0.69***	0.57***	0.59***	0.74***	0.39***	0.4***	0.36***	0.33***	0.32***	0.59***
5. To directly encourage or reinforce positive parent-child interactions	0.67***	0.56***	0.55***	0.77***	0.33***	0.42***	0.35***	0.39***	0.34***	0.59***
6. To help parents use available resources to support child development	0.19***	0.17***	0.15***	0.22***	0.1**	0.15***	0.1**	0.16***	0.12***	0.19***
4. Non-Intrusiveness & Collaboration	0.74***				1.00***	0.57***	0.58***	0.5***	0.39***	0.71***
1. To encourage the parent's ideas and interests for interactions with child	0.41***	0.23***	0.36***	0.27***	0.5***	0.38***	0.37***	0.37***	0.23***	0.43***
2. To avoid intruding on or ignoring parent-child interactions	0.39***	0.12***	0.29***	0.25***	0.66***	0.41***	0.43***	0.38***	0.22***	0.43***
3. To keep parent in the "teacher" role	0.57***	0.37***	0.44***	0.46***	0.63***	0.41***	0.43***	0.41***	0.21***	0.53***
4. To follow the lead of parent and child in pace and activities	0.65***	0.57***	0.5***	0.53***	0.57***	0.46***	0.41***	0.35***	0.4***	0.6***
5. To allow parent-child interactions to continue uninterrupted	0.38***	0.14***	0.25***	0.23***	0.66***	0.34***	0.34***	0.25***	0.27***	0.39***
5. Parent-Child Interaction during Home Visit	0.65***					0.9***	1.00***	0.79***	0.53***	0.84***
1. To engage in interactions	0.59***	0.34***	0.57***	0.5***	0.53***	0.76***	0.84***	0.69***	0.44***	0.74***
2. To make contact with each other	0.53***	0.33***	0.51***	0.46***	0.44***	0.69***	0.77***	0.63***	0.39***	0.67***
3. To be available to engage in interactions	0.47***	0.29***	0.42***	0.43***	0.4***	0.74***	0.72***	0.62***	0.54***	0.66***
4. To observe and be ready to respond to the child's behavior	0.43***	0.31***	0.38***	0.39***	0.32***	0.55***	0.59***	0.59***	0.25***	0.53***
5. To respond to child and support child development	0.45***	0.31***	0.39***	0.39***	0.41***	0.68***	0.75***	0.65***	0.37***	0.62***
6. To adapt activities to child's interests and needs and encourage child engagement	0.54***	0.41***	0.46***	0.46***	0.48***	0.66***	0.69***	0.6***	0.4***	0.66***
7. To sustain positive interactions	0.58***	0.37***	0.5***	0.55***	0.5***	0.8***	0.79***	0.68***	0.57***	0.75***
6. Parent Engagement during Home Visit	0.61***					0.85***		1.00***	0.42***	0.80***
1. To show interest in materials and activities	0.5***	0.36***	0.44***	0.42***	0.43***	0.68***	0.68***	0.76***	0.32***	0.64***
2. To participate and focus on home visit topics and activities	0.54***	0.37***	0.47***	0.46***	0.46***	0.74***	0.71***	0.86***	0.36***	0.7***
3. To engage in play and activities with child	0.51***	0.36***	0.44***	0.45***	0.43***	0.74***	0.69***	0.85***	0.37***	0.68***
4. To initiate activities and conversations	0.49***	0.36***	0.43***	0.41***	0.45***	0.72***	0.68***	0.84***	0.34***	0.66***
5. To discuss questions and topics relevant to child and family	0.56***	0.45***	0.52***	0.49***	0.41***	0.63***	0.58***	0.77***	0.28***	0.65***
6. To be ready to interact with both child and home visitor	0.52***	0.35***	0.46***	0.5***	0.42***	0.71***	0.66***	0.7***	0.47***	0.68***
7. Child Engagement during Home Visit	0.5***					0.79***			1.00***	0.71***
1. To participate in home visit activities	0.45***	0.29***	0.41***	0.44***	0.35***	0.68***	0.49***	0.36***	0.83***	0.62***
2. To initiate successful activities or interactions	0.39***	0.31***	0.33***	0.36***	0.31***	0.6***	0.36***	0.28***	0.81***	0.54***
3. To sustain interactions with parent or home visitor	0.47***	0.29***	0.44***	0.42***	0.38***	0.75***	0.54***	0.47***	0.85***	0.67***
4. To show interest and enthusiasm about home visit activities	0.42***	0.28***	0.37***	0.41***	0.33***	0.66***	0.42***	0.33***	0.85***	0.59***

Notes: Pearson correlation coefficients between each item of the HOVRS-A+2 with the scales', domains' and total scores of the instrument, as well as the scales' scores between themselves and with the domains' and the total score. Coefficients significant at * p<0.10, ** p<0.05, and *** p<0.01.

Table A2: Items not selected for the shorter checklist and criterion that was not met

Construct and item	N	Proportion	Factor load	Criterion not met
Responsiveness to Family				
28. While main caregiver and child were doing the activity, how often did the home visitor comment/label/respond to what the main caregiver and child were doing (moment 1)	428	-	-	1
34. Did the home visitor ask the caregiver to provide some examples of how she could use the games and activities carried out during the visit in her daily life	353	-	-	1
35. Did the home visitor ask the main caregiver what might get in the way of doing the activities	354	-	-	1
Relationship with Family				
40. How often did the home visitor address the child by her name	553	0.93	-	2
41. How often did the home visitor get down to the child's level to interact with her	554	0.97	-	2
46. Overall, was the home visitor's attitude towards the main caregiver shy	552	0.03	-	2
46. Overall, was the home visitor's attitude towards the main caregiver indifferent	552	0.04	-	2
46. Overall, was the home visitor's attitude towards the main caregiver directive	552	0.03	-	2
24. Did the home visitor give the child the opportunity to explore the material before beginning the activity (moment 1)	244	-	-	1
24. Did the home visitor give the child the opportunity to explore the material before beginning the activity (moment 2)	554	0.93	-	2
Facilitation of Parent-Child Interaction				
28. While main caregiver and child were doing the activity, how often did the home visitor comment/label/respond to what the main caregiver and child were doing (moment 1)	428	-	-	1
46. Overall, was the home visitor's attitude towards the main caregiver shy	552	0.03	-	2
46. Overall, was the home visitor's attitude towards the main caregiver indifferent	552	0.04	-	2
46. Overall, was the home visitor's attitude towards the main caregiver directive	552	0.03	-	2
Child and caregiver engagement during visit				
30. Was the activity adequate to the developmental level of the child (moment 1)	553	0.95	-	2
30. Was the activity adequate to the developmental level of the child (moment 2)	554	1.00	-	2
30. Was the activity adequate to the developmental level of the child (moment 3)	548	0.98	-	2
31. Was the activity of interest to the child (moment 2)	551	0.95	-	2
50. How often was the home visitor taking care of the child alone during the visit	550	0.03	-	2
Key practices during visit moments				
25. How well did the home visitor demonstrate the activities to the main caregiver (moment 1)	451	-	-	1
52. How often was the visit conflictive/tense ¹	553	0.04	-	2
Knowledge of protocols				
18. Did the home visitor ask the main caregiver if any of the activities demonstrated in the previous week were done	554	0.97	-	2
22. Did the home visitor bring the toys and materials planned for the visit (moment 1)	553	0.93	-	2
22. Did the home visitor bring the toys and materials planned for the visit (moment 2)	552	1.00	-	2
22. Did the home visitor bring the toys and materials planned for the visit (moment 3)	546	0.99	-	2
24. Did the home visitor give the child the opportunity to explore the material before beginning the activity (moment 1)	244	-	-	1
24. Did the home visitor give the child the opportunity to explore the material before beginning the activity (moment 2)	554	0.93	-	2
29. Did the home visitor carry out all the steps of the activity (moment 1)	419	-	-	1
29. Did the home visitor carry out all the steps of the activity (moment 2)	553	0.91	-	2
33. Did the home visitor ask the main caregiver to recall the activities she would be carrying out during the week with the child	554	0.68	0.05	3
36. Did the home visitor ask the main caregiver what else they could do during the week	523	0.18	0.18	3
32. Did the home visitor transition easily from one activity to the other (moment 1 to 2)	552	0.94	-	2
32. Did the home visitor transition easily from one activity to the other (moment 2 to 3)	547	0.91	-	2
49. How often did the home visitor and caregiver discuss topics unrelated to visit's objectives ¹	553	0.02	-	2
Recap of activities				
18. Did the home visitor ask the main caregiver if any of the activities demonstrated in the previous week were done	554	0.97	-	2
33. Did the home visitor ask the main caregiver to recall the activities she would be carrying out during the week with the child	554	0.68	0.22	3
34. Did the home visitor ask the caregiver to provide some examples of how she could use the games and activities carried out during the visit in her daily life	353	-	-	1
35. Did the home visitor ask the main caregiver what might get in the way of doing the activities	354	-	-	1
36. Did the home visitor ask the main caregiver what else they could do during the week	523	0.18	0.22	3

Notes: criterion 1 = no more than 15 percent of missing observations; criterion 2 = at least 10 percent of responses different from 0 or from 1 for dichotomous items, and 10 percent of responses different from adjacent categories "always" and "almost always" combined for non-dichotomous items; criterion 3 = factor load ≥ 0.4 (only when criteria 1 and 2 are met and construct has an eigenvalue > 1).

¹ Non-dichotomous items.

Table A3: Correlations between the items and “constructs” of the checklist with the scales, domains and total scores of the HOVRS-A+-2

'Constructs' in the checklist	HOVRS-A+									
	Domain 1. Home Visitors Practices	1. Responsiveness to Family	2. Relationship with Family	3. Facilitation of Parent-Child Interaction	4. Non-Intrusiveness & Collaboration	Domain 2. Family Engagement	5. Parent-Child Interaction during Home Visit	6. Parent Engagement during Home Visit	7. Child Engagement during Home Visit	Total score
I. Responsiveness to Family	0.36***	0.3***	0.38***	0.34***	0.16***	0.23***	0.21***	0.23***	0.15***	0.32***
28. While main caregiver and child were doing the activity, how often did the home visitor comment/label/respond to what the main caregiver and child were doing (moment 2)	0.33***	0.28***	0.34***	0.31***	0.15***	0.2***	0.2***	0.17***	0.15***	0.29***
28. While main caregiver and child were doing the activity, how often did the home visitor comment/label/respond to what the main caregiver and child were doing (moment 3)	0.29***	0.23***	0.31***	0.29***	0.11**	0.13***	0.11**	0.12***	0.11**	0.23***
37. How often did the home visitor ask the main caregiver's opinion about things like: familiarity with the games and activities proposed, games the child usually plays, what the child can do, what the mother likes to do with child or the whole family enjoy doing together, etc.	0.21***	0.21***	0.22***	0.17***	0.08*	0.14***	0.12***	0.17***	0.059	0.19***
38. How often did the home visitor show interest in what the main caregiver answered	0.25***	0.21***	0.28***	0.23***	0.1**	0.22***	0.21***	0.22***	0.13***	0.25***
39. How often did the home visitor build on the information provided by the main caregiver during the visit	0.32***	0.26***	0.34***	0.3***	0.15***	0.26***	0.26***	0.24***	0.17***	0.32***
II. Relationship with Family	0.35***	0.26***	0.35***	0.35***	0.19***	0.32***	0.29***	0.26***	0.26***	0.36***
42. How often did the home visitor respond to the child's vocalizations and gestures	0.16***	0.11***	0.19***	0.15***	0.061	0.18***	0.12***	0.094**	0.23***	0.18***
43. How often did the home visitor praise/encourage the child for her attempts (to do things) during the activity	0.23***	0.17***	0.23***	0.23***	0.13***	0.2***	0.15***	0.13***	0.21***	0.23***
26. How well did the home visitor engage the main caregiver and child in the activities (moment 1)	0.27***	0.21***	0.27***	0.25***	0.16***	0.27***	0.25***	0.21***	0.23***	0.3***
26. How well did the home visitor engage the main caregiver and child in the activities (moment 2)	0.24***	0.18***	0.25***	0.23***	0.14***	0.24***	0.25***	0.23***	0.15***	0.27***
26. How well did the home visitor engage the main caregiver and child in the activities (moment 3)	0.24***	0.16***	0.21***	0.25***	0.17***	0.29***	0.25***	0.22***	0.25***	0.29***
38. How often did the home visitor show interest in what the main caregiver answered	0.25***	0.21***	0.28***	0.23***	0.1**	0.22***	0.21***	0.22***	0.13***	0.25***
44. How often did the home visitor prompt the caregiver to respond to and/or praise the child	0.27***	0.17***	0.26***	0.3***	0.14***	0.19***	0.19***	0.15***	0.15***	0.25***
45. How often did the home visitor praise/encourage the main caregiver during the activity for his/her efforts	0.33***	0.26***	0.33***	0.33***	0.16***	0.28***	0.26***	0.26***	0.2***	0.33***
III. Facilitation of Parent-Child Interaction	0.36***	0.28***	0.36***	0.36***	0.18***	0.27***	0.23***	0.23***	0.21***	0.34***
26. How well did the home visitor involve the caregiver and the child during the activity (moment 1)	0.27***	0.21***	0.27***	0.25***	0.16***	0.27***	0.25***	0.21***	0.23***	0.3***
26. How well did the home visitor involve the caregiver and the child during the activity (moment 2)	0.24***	0.18***	0.25***	0.23***	0.14***	0.24***	0.25***	0.23***	0.15***	0.27***
26. How well did the home visitor involve the caregiver and the child during the activity (moment 3)	0.24***	0.16***	0.21***	0.25***	0.17***	0.29***	0.25***	0.22***	0.25***	0.29***
28. While main caregiver and child were doing the activity, how often did the home visitor comment/label/respond to what the main caregiver and child were doing (moment 2)	0.33***	0.28***	0.34***	0.31***	0.15***	0.2***	0.2***	0.17***	0.15***	0.29***
28. While main caregiver and child were doing the activity, how often did the home visitor comment/label/respond to what the main caregiver and child were doing (moment 3)	0.29***	0.23***	0.31***	0.29***	0.11**	0.13***	0.11**	0.12***	0.11**	0.23***
44. How often did the home visitor prompt the caregiver to respond to and/or praise the child	0.27***	0.17***	0.26***	0.3***	0.14***	0.19***	0.19***	0.15***	0.15***	0.25***
45. How often did the home visitor praise/encourage the main caregiver during the activity for his/her efforts	0.33***	0.26***	0.33***	0.33***	0.16***	0.28***	0.26***	0.26***	0.2***	0.33***

IV. Child and caregiver engagement during visit (N=542)	0.22***	0.14***	0.21***	0.20***	0.18***	0.36***	0.27***	0.22***	0.39***	0.31***
31. Was the activity of interest to the child (moment 1)	0.21***	0.13***	0.22***	0.19***	0.16***	0.27***	0.21***	0.18***	0.28***	0.26***
31. Was the activity of interest to the child (moment 3)	0.09**	0.04	0.08*	0.08*	0.09**	0.15***	0.09**	0.04	0.22***	0.13***
47. How often did the child actively participate during the visit	0.17***	0.10**	0.16***	0.16***	0.13***	0.30***	0.20***	0.13***	0.40***	0.26***
48. How often did the caregiver actively participate during the visit	0.19***	0.13***	0.19***	0.15***	0.16***	0.30***	0.30***	0.35***	0.14***	0.27***
V. Key practices during visit moments (N=533)	0.38***	0.31***	0.38***	0.35***	0.21***	0.32***	0.30***	0.26***	0.25***	0.38***
23. How well did the home visitor explain the objective of the activities shown in the current week (moment 1)	0.26***	0.22***	0.26***	0.24***	0.15***	0.21***	0.22***	0.17***	0.16***	0.26***
23. How well did the home visitor explain the objective of the activities shown in the current week (moment 2)	0.33***	0.28***	0.34***	0.30***	0.16***	0.23***	0.23***	0.22***	0.13***	0.30***
23. How well did the home visitor explain the objective of the activities shown in the current week (moment 3)	0.32***	0.26***	0.31***	0.28***	0.20***	0.22***	0.22***	0.20***	0.15***	0.29***
25. How well did the home visitor demonstrate the activities to the main caregiver (moment 2)	0.26***	0.24***	0.27***	0.25***	0.08**	0.12***	0.12***	0.11***	0.08*	0.20***
25. How well did the home visitor demonstrate the activities to the main caregiver (moment 3)	0.22***	0.19***	0.23***	0.20***	0.10**	0.16***	0.12***	0.09**	0.17***	0.20***
26. How well did the home visitor engage the main caregiver and child in the activities (moment 1)	0.27***	0.21***	0.27***	0.25***	0.16***	0.27***	0.25***	0.21***	0.23***	0.30***
26. How well did the home visitor engage the main caregiver and child in the activities (moment 2)	0.24***	0.18***	0.25***	0.23***	0.14***	0.24***	0.25***	0.23***	0.15***	0.27***
26. How well did the home visitor engage the main caregiver and child in the activities (moment 3)	0.24***	0.16***	0.21***	0.25***	0.17***	0.29***	0.25***	0.22***	0.25***	0.29***
51. How often was the visit fun	0.21***	0.14***	0.23***	0.19***	0.13***	0.30***	0.22***	0.16***	0.35***	0.28***
VI. Stand-alone items										
19. The caregiver and the child demonstrated the activities from the previous week	0.19***	0.12***	0.18***	0.20***	0.14***	0.14***	0.13***	0.13***	0.11**	0.18***
29. The home visitor carried out all the steps of the activities to teach (moment 3)	0.16***	0.10**	0.15***	0.15***	0.11***	0.15***	0.14***	0.13***	0.12***	0.17***
33. The home visitor asked the main caregiver to recall the activities she would be carrying out during the week with the child	0.22***	0.22***	0.19***	0.17***	0.14***	0.16***	0.14***	0.16***	0.11***	0.21***
36. The home visitor asked the main caregiver what else they could do during the week	0.22***	0.26***	0.19***	0.21***	0.06	0.14***	0.13***	0.15***	0.08*	0.19***
46. Overall, the home visitor's attitude towards the main caregiver was encouraging	0.23***	0.19***	0.23***	0.20***	0.12***	0.18***	0.18***	0.19***	0.10**	0.22***

Notes: Pearson correlation coefficients between the “constructs” of the checklist and stand-alone items with the scales’, domains’ and total scores of the HOVRS-A+2. Items mapped to constructs I, II, III of the checklist are scored on a scale from 1 to 5 where 1 refers to “almost never” and 5 to “almost always”, except for items 26 in “constructs” II and III where the scores correspond to 1 “did not explain/demonstrate/engage”, 2 to “poor explanation/ demonstration/ engagement”, 3 “average”, 4 “good”, and 5 “very good explanation/demonstration/engagement.” For construct IV, items 31 are binary (0-1) and items 47 and 48 are scored on scale from 1 to 5 where 1 refers to “almost never” and 5 to “almost always.” All items under “construct” V are scored on a scale from 1 to 5, with the same categories as item 26 mentioned above. All items under VI are binary (0-1). The scores for each “construct” of the checklist were computed by factor analysis. Coefficients significant at * p<0.10, ** p<0.05, and *** p<0.

Table A4: Proposed checklist for monitoring home visiting quality at-scale

Panel I: Information about the visit, child and home visitor		
Visit		
Date		
Duration		
Number of adults (other than the caregiver) present		
Number of children (other than focus child) present		
Home visitor		
Completed years of education ¹		
Date of entry into program ¹		
Date when started working with family ¹		
Child		
Date of birth ¹		
Date of entry into program ¹		
	Scoring	Construct(s)
Panel II: Content and Planning		
Welcome and review of previous visit		
19. Did the caregiver and the child demonstrate the activities from the previous week	[0-1]	Stand-alone item
Activities²		
23. How well did the home visitor explain the objectives and achievements of the activity	[1-5]	Key practices
25. How well did the home visitor demonstrate the activity to the caregiver	[1-5]	Key practices
26. How well did the home visitor involve the caregiver and the child during the activity	[1-5]	Key practices/ Relationship/Facilitation
28. While main caregiver and child were doing the activity, how often did the home visitor comment/label/respond to what the main caregiver and child were doing	[1-5]	Responsiveness/Facilitation
29. Did the home visitor carry out all the steps of the activity	[0-1]	Stand-alone item
31. Was the activity of interest to the child	[0-1]	Child and caregiver engagement
Closure and recap		
33. Did the home visitor ask the main caregiver to recall the activities she would be carrying out during the week with the child	[0-1]	Stand-alone item
36. Did the home visitor ask the main caregiver what else they could do during the week	[0-1]	Stand-alone item
Panel II: Interactions		
38. How often did the home visitor show interest in what the main caregiver answered	[1-5]	Responsiveness/Relationship
39. How often did the home visitor build on the information provided by the main caregiver during the visit	[1-5]	Responsiveness
42. How often did the home visitor respond to the child's vocalizations and gestures	[1-5]	Relationship
43. How often did the home visitor praise/encourage the child for her attempts (to do things) during the activity	[1-5]	Relationship/Facilitation
44. How often did the home visitor prompt the caregiver to respond to and/or praise the child	[1-5]	Relationship/Facilitation
45. How often did the home visitor praise/encourage the main caregiver during the activity for his/her efforts	[1-5]	Relationship
46. Overall, the home visitor's attitude towards the main caregiver was encouraging	[0-1]	Stand-alone item
Panel III: Overall assessment of the child and caregiver's participation		
47. How often did the child actively participate during the visit	[0-1]	Child and caregiver engagement
48. How often did the caregiver actively participate during the visit	[0-1]	Child and caregiver engagement
51. How often was the visit fun	[1-5]	Key practices

Notes: ¹Information that should be collected only during the first home visit. ²Items applied at various moments of the visit for the Cuna Mas Program.