

Measuring the Quality of Home-Visiting Services

A Review of the Literature

Sara Schodt
James Parr
María Caridad Araujo
Marta Rubio-Codina

Social Protection and
Health Division

TECHNICAL
NOTE N°
IDB-TN-881

Measuring the Quality of Home-Visiting Services

A Review of the Literature

Sara Schodt
James Parr
María Caridad Araujo
Marta Rubio-Codina

October 2015



Cataloging-in-Publication data provided by the
Inter-American Development Bank

Felipe Herrera Library

Measuring the quality of home-visiting services: a review of the literature / Sara Schodt,
James Parr, María Caridad Araujo, Marta Rubio-Codina.

p. cm. — (IDB Technical Note ; 881)

Includes bibliographical references.

1. Child development. I. Schodt, Sara. II. Parr, James. III. Araujo, María Caridad. IV.
Rubio-Codina, Marta. V. Inter-American Development Bank. Social Protection and
Health Division. VI. Series.

IDB-TN-881

<http://www.iadb.org>

Copyright © 2015 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. 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.

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

<http://www.iadb.org/SocialProtection>

Measuring the Quality of Home-Visiting Services

A Review of the Literature

Sara Schodt, James Parr, María Caridad Araujo y Marta Rubio-Codina*

Abstract

This document reviews the literature on the definition and measurement of quality of home visiting programs designed to promote early childhood development, with a particular focus on those interventions aimed at enhancing child cognitive, language, and socio-emotional outcomes. After summarizing the evidence on home visiting programs in the United States and internationally, we discuss the key elements that define a high quality home visit. Next, we describe a range of instruments designed to measure both structural and process elements of quality of home visits, instances in which they have been used, and the results of their administration.

JEL Classification: J13, H53, H75, I38

Key Words: Early childhood development, quality, home visiting programs, measurement, process quality, structural quality.

* The preparation of this document was led by Sara Schodt (SCL/SCL), with valuable contributions from James Parr and under the guidance of M. Caridad Araujo and Marta Rubio-Codina. We thank Florencia López-Boo for her valuable comments and insights.

Contents

Introduction.....	2
Review of evidence-based home visiting programs and their evaluations.....	3
Quality home visits: a definition.....	8
Dosage.....	9
Content.....	10
Instruments to measure the quality of home visits.....	13
Use of instruments to measure home visit quality as part of program evaluations.....	19
<i>Better Beginnings: HOVRS-A and HVCCF</i>	19
<i>Baby FACES: HOVRS-A and HVCCF</i>	20
<i>Illinois Prevention Initiative - HOVRS-A+</i>	22
<i>Part C and Early Head Start in Central Iowa: HOVF</i>	22
<i>Evaluation of Family Check-Up Model: COACH</i>	23
Quality of home visits and its effects on child development.....	23
Concluding thoughts.....	27
Appendix.....	35

Introduction

Home visitation or home visiting, henceforth used interchangeably, is a strategy for the delivery of social services used to reach at-risk children and their caregivers, especially their mothers.¹ Through home visits, programs aim to change parental behavior and childrearing practices in ways that can improve children's or families' outcomes (Howard & Brooks-Gunn 2009). Different home-visiting models employ a diverse array of techniques and curricula focused on improving an equally varied range of outcomes across maternal and child health, family/parent functioning, and early childhood social, emotional, and cognitive development (Stoltzfus & Lynch 2009). Home visits are typically delivered through structured interactions involving a trained visitor, a caregiver, and a child. These interactions take place on a regular basis in the home of the caregiver.

Home visiting has grown in popularity in Latin America and the Caribbean (LAC) and in other regions of the world, bolstered by evidence of high and sustained impacts on child development. For the most part, this evidence comes from small-scale, rigorously evaluated programs, and for this reason, it is not clear if—when rolled out at scale—home visitation programs can be implemented with enough fidelity to achieve the same types of impacts. A key to ensuring fidelity and producing quality visits at any scale is collecting information on what exactly takes place during the home visit itself and whether the interactions that are at the heart of the visit are taking place as intended. Activities involved in collecting information about the home visit can be used for professional development and for training new and ongoing visitors. Additionally, collecting information on the home visit can provide a more nuanced and in-depth look at the processes that are most effective in improving child outcomes, allowing programs to continue to improve their services. It is for these reasons that tools for measuring the quality of home visits are a key element of consolidating effective programs at scale.

This note focuses on reviewing the literature on the definition and measurement of quality of home visits, with a particular focus on programs designed to enhance child cognitive or socio-emotional outcomes. To contextualize the subject, the first section presents a brief history of evidence-based home-visiting programs in the United States and internationally, with a focus on how these programs measured quality. The second section of this document reviews the literature on key components of the quality of home visiting. A third and final section describes the array of instruments that has been used for the measure of quality of home visits and the results of these instruments' administration.

¹ Throughout this document, we use the terms “father,” “mother,” and “caregiver” interchangeably to refer to the adults in charge of the children who receive home visits and who participate in them.

Review of evidence-based home-visiting programs and their evaluations

The goal of this section is to present a short summary of the most important evidence on home-visiting programs worldwide. The **Nurse-Family Partnership (NFP)** is, arguably, the best known and evaluated home-visiting program in the United States. It started in 1977 in Elmira, New York, to serve low-income, first-time mothers and their children. Under its model, mothers received bi-weekly visits from a registered nurse, from pregnancy until the child turned 24 months of age, with the goals of improved pregnancy outcomes, parenting skills, and maternal life course (Olds et al 2002). The NFP model has been tested in randomized control trials for over 35 years, not only in New York, but also, beginning in 1988, in Memphis, Tennessee, and in 1994 in Denver, Colorado. Consistent positive effects were found (in at least two of the three trials) on prenatal health and child school readiness at age 3, number of child injuries, number of subsequent mother pregnancies, spacing between births, and maternal employment. Important for the purposes of this review, children who received NFP had higher levels of achievement in Peabody Individual Achievement Tests (88.78 vs 85.70, $P = .009$), and higher scores on group-administered standardized language and math tests after 6 years of school (40.52 vs 34.85, $P = .02$; Kitzman et al 2010). The positive impacts from these initial NFP randomized trials led to an expansion of the program in 1996, and since that date, NFP has served over 212,000 low-income, first-time mothers in 43 states at over 100 program sites (Howard & Brooks-Gunn 2009; Daro 2006; Olds et al 1997).²

The impacts of other large-scale home-visiting programs on child outcomes have been less clear. For example, the **Hawaii Healthy Start (HHS)** program, begun in 1975 and employing paraprofessionals as home visitors (Duggan et al 2000), has shown mixed results. HHS was designed to prevent child abuse and neglect in the first three years of children's lives by improving the functioning of at-risk families, identified by a short questionnaire filled out by the mother at the time of her child's birth. It was evaluated in both randomized control and quasi-experimental trials. Early quasi-experimental evaluations suggested that the program was successful in the following: increasing pediatric health care use for illness or injury; improving maternal life skills, mental health, caregiver use of social support, the child's home learning environment, and parent-child interactions; and reducing caregiver substance use. In addition, a more rigorous impact evaluation on the program by Duggan et al (2004) found the impacts to be very limited. Specifically, the authors found HHS to be effective in linking families with pediatric medical care, improving some parenting skills, and decreasing injuries resulting from partner violence in the home, but they found no impacts on children's development or other outcomes (Duggan et al 2004). This model was later exported to a number of other states as the **Healthy Families America** program, which still screens mothers of newborn children for child-abuse risk factors.

The national **Early Head Start (EHS)** program has been brought to scale in the US and also employs at-scale home visiting as part of its service delivery; it has been evaluated numerous times. EHS targets both parents and young children, with the goal of improving the health and development of children under 3 from low-income backgrounds and improving their families' access to support services, as well as their well-being and self-sufficiency. As part of EHS home-based services, families receive weekly home

² Extensive information about the NFP can be found on the program's website, <http://www.nfp.org>.

visits focusing on developing parents' abilities to support a child's healthy development. Twice per month, group sessions are provided for parents and children to come together for learning, discussion, and social activity.

A number of randomized control evaluations of this program have yielded some outcomes in the area of child development that seem to vary with children's age at time of evaluation, ethnicity, and the quality of implementation of their particular EHS program. In three small, separate studies by Roggman & Cook (2010), Love et al (2001), and Roggman, Boyce & Cook (2009), no impacts were found in the area of child development³ at 24 months of age.

One of the largest randomized evaluations of EHS involved over 17 sites and about 3,000 children and parents; it measured children at 14 and 24 months of age during the program, at 36 months of age at the end of the program, and two years after the end of the program at age 5. The authors found that children participating in EHS home visitation services had improved outcomes in both the cognitive and language domains (Jones Harden et al 2012). At age 2, the effect magnitude was 0.12 of a standard deviation (SD) on the MacArthur Communication Development Inventory (CDI) and 0.15 SD on the Bayley Scales of Infant Development-II (Bayley-II). At age 3, there was an effect of 0.13 SD on the Peabody Picture Vocabulary Test-III (PPVT-III) and 0.15 SD on the Bayley-II. Finally, at age 5, two years after children graduated from the program, an effect of 0.09 SD on the PPVT was found, but no effect was found on the Woodcock-Johnson tests of Letter-Word Identification (English) or Applied Problems (math). Significant positive effects were also found on child socio-emotional development age 2 to 3 (between 0.1 and 0.2 SD), some of which were sustained until age 5. Child outcomes across the board, as well as some positive caregiver behaviors, were found to be stronger when the programs in which children participated were implemented according to federally-mandated guidelines (Jones Harden et al 2012). Raikes et al (2006) reexamined the same evaluation and found that when disaggregating the effects by race, impacts on cognition and language were all driven by program effects on African-American children and not on whites or Hispanics. It is worth noting that treatment children may have also been attending EHS center-based care services from infancy, and control children may also have been receiving center-based or home-visiting services from other providers, making it hard to isolate the effects of home visiting.

Despite the mixed evidence regarding the impact of home-visiting services on child development, there was enough interest generated from these programs that the US government increased resources for home-visiting interventions and research. As a result, in 2008, the Evidence-Based Home-Visiting (EBHV) program dedicated \$10 million toward expanding and improving home-visiting programs (Brookings 2015).

In order to facilitate the process of monitoring and evaluation and the distribution of EBHV funds, in 2009 the US Department of Health and Human Services contracted a review of the effectiveness of evidence-based home-visiting programs; pregnant women and families with children from birth to age 5 were targeted. **The objective of this review was to identify *high-quality* models. The review was called the Home**

³ Measures used included the Bayley Scales of Infant Development – Mental Development Index, the MacArthur Communicative Development Inventories, the Bayley Behavior Rating Scale, the Child Behavior Checklist, and parent-child structured play and child aggression observations.

Visiting Evidence of Effectiveness or HomVEE. The HomVEE defined a program as effective when it found significant positive outcomes in one or more of seven domains: (a) maternal health; (b) child health; (c) child development and school readiness; (d) child maltreatment; (e) juvenile delinquency, family violence, or crime; (f) parenting practices; (g) family economic factors; and (h) linkages and referrals (Paulsell et al 2010).⁴

The HomVEE identified 17 (out of 40) programs that met its standards for being both well-evaluated and also reporting positive impacts on the families and children they served. All 17 programs had some positive impact on primary measures of child development, school readiness, and positive parenting practices, although the magnitude of the outcomes was relatively small in size. None of the programs identified showed reductions in juvenile delinquency, family violence or crime (Paulsell et al 2010). The HomVEE review did not specify a time frame for the evaluations it included, so some programs were evaluated at earlier points in time than others. Appendix table A1 summarizes the results from a subgroup of programs that reported positive impacts specifically on child development outcomes. While the results of the HomVEE analysis were important, this review did not attempt to open the “black box” of home visiting in order to identify the key dimensions of quality or processes occurring during visits that are the main drivers of desired program impacts. This is an important lapse, because this kind of information—traditionally not collected—is now understood to be at the heart of understanding and improving home-visiting program effectiveness.

Outside of the US, there have been a few rigorous evaluations of home-visiting programs as well. The recent *Pro Kind Project in Germany*, based on the Nurse-Family Partnership model, is the first randomized control evaluation for a home-visiting program in that country and has shown positive child development results. Trained midwives, nurses or social pedagogues made weekly, bi-weekly and monthly visits to a randomized treatment group of 393 first-time, low-income and at-risk mothers, and offered health and social services to a control group of 362 comparable mothers (Sierau, Brand & Jungmann 2012). The intervention began during pregnancy and continued until the child’s second birthday; it focused primarily on improving maternal health and parenting skills. In a randomized control evaluation of the program, participant children had improved development (0.18 SD), as measured by the Bayley Scales of Infant Development - Mental Developmental Index (MDI) and Psychomotor Developmental Index (PDI) and a German Language Test (SETK-2) compared to those in the control group at 1 year of age (Sandner 2013); however, at 24 months, the effect size was indistinguishable from zero.

The UCD Geary Institute in **Ireland** employed a randomized control design to evaluate a preventive, early intervention called *Preparing for Life* (PFL). PFL focused on improving life outcomes of children and families in North Dublin, by visiting families and children from pregnancy through the start of formal schooling at age 5. Mothers randomly assigned to a treatment group received bi-weekly visits from trained mentors (with a

⁴ Specifically, HomVEE identified studies for each program model that met at least one of the following criteria: (a) at least one high- or moderate-quality impact study of the model finds favorable, statistically-significant impacts in two or more identified outcome domains; or (b) at least two high- or moderate-quality impact studies of the model using unique study samples find one or more favorable, statistically-significant impacts in the same domain. Studies with random assignment, low attrition and no reassignment of members across treatment and control groups were favored, as were studies with designs that involved regression discontinuity. (Paulsell et al 2014 p.7)

background in early education or psychology/social work), while mothers in the control group received standard medical and social services. The program recruited 233 pregnant women between 2008 and 2010 for participation in the evaluation and measured over 150 outcomes spanning child and family health, development, social conditions, and engagement. The authors report program impact primarily on parental behaviors and the home environment, with little impact on child development at 18 months (Doyle et al 2013); however, at 24 months, children who received home visits had significant positive impacts on many areas, including development⁵ (0.2 -0.24 SD), as compared to children in the control group that did not receive visits (Doyle et al 2013).

In Latin America, a few home-visiting interventions have been rigorously evaluated. **The best-known home-visiting evaluation in LAC is from Jamaica.** The program was a small pilot of a two-year intervention offering weekly home visits, nutritional supplements, or both, to children between 9 and 24 months of age who were stunted (Walker et al 2005). After two years, children who received both weekly stimulation-focused home visits and the nutrition supplements caught up to their non-stunted peers. By age 7, however, the effects of the home visits remained while the effects of the nutritional supplements had faded; the children receiving the combined treatment did not differ statistically from those who received only the home visits, though both groups still outperformed children who received only nutritional supplements. The authors found that at 7 and 11 years of age, those children who received early stimulation had significantly higher IQs and had higher scores on tests of reasoning and vocabulary than those who did not receive it. At 17 to 18 years of age, the same children had lower rates of school drop-out and scored significantly higher on 11 out of 12 cognitive tests, as compared to stunted children who only received nutritional supplements.

Attanasio et al (2014) adapted the curriculum from the Jamaica study to Colombia and implemented and evaluated it at larger scale using a cluster randomized design. The authors followed 1,420 children between 1 and 2 years of age from the poorest 20% of households in 96 municipalities across eight departments. They compared the effectiveness of a home-visiting-based, cognitive-stimulation intervention, micronutrient-supplementation intervention, or both, delivered over a period of 18 months. They found that the stimulation services improved children's cognitive scores by 0.26 SD and increased receptive language scores by 0.22 standard deviations, all measured with Bayley-III, while the nutritional supplementation in isolation had no significant impact on children's outcomes. In practical terms, the impact of the stimulation visits was enough to narrow the gap between poor and wealthy children's cognitive scores by almost one-third.

In **Nicaragua**, López-Boo, Urzua & Palloni (2013) estimated the average treatment effects on child development of a large-scale, integrated early childhood development program that included a home-visiting component. The impacts found were modest, although not that dissimilar in magnitude from those found in Early Head Start evaluations. Among a subset of the sample of children who were 36 to 60 months of age, the program improved verbal and numeric memory as measured by the McCarthy Scales of Children's Ability by 0.13 SD after a year and a half.

⁵ Standard deviations reported here are ranges for the statistically-significant outcomes from a battery of 41 instruments assessing child development. Fourteen of these instruments had statistically-significant outcomes, including the Short Early Development Instrument, the MacArthur-Bates Communicative Development Inventories, the Developmental Profile 3, and the Brief Infant-Toddler Social and Emotional Assessment. For more information see http://geary.ucd.ie/preparingforlife/wp-content/uploads/2014/03/24MoReport_final.pdf.

The Cuban *Educa a tu Hijo* program is a large-scale, cross-sectoral program that serves the majority of young children in Cuba. Home visitors are trained community members who may be educators, family doctors, nurses, or volunteers. These visitors provide hour-long home visits once to twice weekly for children under 2 years of age and their parents, as well as weekly two-hour group sessions for children ages 3 to 6. The themes covered focus on improving practices and activities that promote healthy child development. To date, the program has not been rigorously evaluated. Nonetheless, this model has been replicated across LAC in home-visiting programs in Brazil, Chile, Colombia, Ecuador, Guatemala, Mexico, and Venezuela (Tinajero 2010).

There are several other programs that use home visits to promote child development in Latin America. In Brazil, the Better Early Childhood (***Programa Infância Melhor - PIM***) home-visiting program, based on the Cuban model, is implemented in the state of Rio Grande do Sul to promote the comprehensive development of children from the prenatal stage to 6 years, with particular focus on the 0 to 3 period (UNESCO 2007). As of 2014, the program served 55,140 families, 8,271 pregnant mothers and 60,654 children in total (Schneider & Ramires 2007). The Peruvian ***Servicio de Acompañamiento a Familias (SAF) - Programa Nacional Cuna Más*** (PNCM), launched in 2013, reaches over 60,000 families in poor, rural districts with weekly home visits carried out by trained community personnel. It has an ongoing experimental evaluation built into its expansion (PNCM 2015). In Ecuador, ***Creciendo con Nuestros Hijos*** reaches over 240,000 children mainly in rural areas through weekly home visits (for children younger than 2) and group meetings (for children 2 to 3 years of age).⁶ In Colombia, the *Instituto Colombiano de Bienestar Familiar's Modalidad Familiar* combines weekly group meetings and monthly home visits for children younger than 2.⁷ To our knowledge, with the exception of SAF-Cuna Más, most of these programs have not been the subjects of experimental evaluations, nor has there been any systematic information collected on the key dimensions of quality of the home visits they deliver.

Outside of LAC, home visiting has been implemented and evaluated in other low- and middle-income countries. In Bangladesh, Hamadani et al (2006) assessed the effects of adding psychosocial stimulation to the treatment of undernourished children in a randomized controlled trial. Based in part on the Jamaican intervention described earlier, under the **Bangladesh Integrated Nutrition Program (BINP)**, undernourished children 6 to 24 months old received nutrition supplements through community nutrition centers (CNCs). The intervention incorporated psychosocial stimulation into the treatment of the same group of undernourished children, focusing on responsive parenting (including feeding) and child development education. The stimulation intervention provided additional home visits and group meetings for mothers and 107 children from 20 CNCs for 12 months. They later compared this group to a control group of 107 children from the same villages who received only the nutrition supplements. The authors found that after a year of the psychosocial-stimulation-based visiting program, children who received the intervention had significantly improved mental scores (0.33 SD) and ratings

⁶ Modalidad Creciendo con Nuestros Hijos Ecuador, atención domiciliaria, February 2014, retrieved on July 11, 2015, from <http://www.desarrollosocial.gob.ec/wp-content/uploads/2014/05/Presentacion-CNH-6.2.2014.pdf>

⁷ "Primera Infancia." Instituto Colombiano De Bienestar Familiar: Modalidad Familiar. Consulted on July 11, 2015 at <http://www.icbf.gov.co/portal/page/portal/PrimeraInfanciaICBF/Serviciosdeatencion/modalidadesdeeducacioninicial/ModalidadFamiliar>.

on vocalization (0.30 SD) and cooperation (0.41 SD), among other outcomes, as measured by the revised version of the Bayley Scales (BSID-II).

Also in Bangladesh, Nahar et al (2012) studied **malnourished children receiving either nutrition or play-based stimulation delivered through home visits**. The authors examined a sample of 507 hospitalized, severely-underweight children between 6 months and 2 years of age. Results on the Bayley and a battery of motor skills tests were measured at baseline and after three and six months of intervention. At the six-month follow-up, the treatment group had improved Bayley mental development scores (0.97 SD) and motor skills scores (0.56 SD).

Yousifzai et al (2014a) investigated the effects of home visiting on child development and growth outcomes for 1,489 mother–infant dyads in the **Lady Health Worker (LHW) program in rural Pakistan**. Randomizing at the community level, researchers assigned 80 clusters of children to receive either routine health and nutrition services or responsive stimulation, or both enriched interventions. The treatments were delivered through both group sessions and home visits. The Bayley-III was used to measure child development at 12 and 24 months of age. Children who received responsive stimulation had significantly higher development scores on the cognitive, language, and motor scales at 12 and 24 months of age and on the social-emotional scale at 12 months of age, than did those who did not receive the intervention. Specifically, at 24 months, the responsive stimulation intervention improved development outcomes with moderate to large effect sizes on cognition (0.6 SD), language (0.7 SD) and motor (0.5 SD) development as assessed with the Bayley-III (Yousafzai et al 2014). No additive benefits were found from the combined interventions.

In summary, there is a range of experiences in the implementation and evaluation of home-visiting programs in different contexts and countries. The evidence from high, long-term impacts of these interventions comes mainly from small-scale pilots; however, less is known about the impacts of home-visiting programs implemented at scale and especially on whether, in the process of bringing this type of intervention to scale, it is possible to implement it with the fidelity that is necessary to achieve impacts comparable to those demonstrated by pilot efforts.

Quality home visits: a definition

The evaluations discussed in the previous section provide information on programs that have had positive impacts on child outcomes; however, we know little about what kinds of practices and processes occurring during home visits contribute to those outcomes and, more broadly, to a program's success. This section aims to summarize what the literature identifies as the critical elements of a high-quality home visit.

Even when information is collected on what happens during a home visit, more often than not, it is not widely published or shared with outside audiences that might learn from it. As Yousefzai et al (2014b) point out, understanding more about the implementation of a program or project allows those in the home-visiting field to accurately identify what a program was doing that produced certain outcomes. A better understanding of what is happening during a home visit would help us ensure that programs are meeting their intended objectives; moreover, the knowledge gained would

be an important tool for linking ongoing research with practice and improving service delivery.

How has the quality of home visits been assessed? Paulsell et al (2010) suggest that there are **three dimensions that define home-visiting quality: dosage, content, and relationships**. Evidence and common sense suggest that the interplay between those three variables is just as important as each unique dimension; for example, dosage and content may have little impact on children's outcomes if the quality of the relationships fostered during the visits is low (Roggman et al 2008, Wasik & Bryant 2001).

Both dosage and content can be characterized as *structural* elements of quality; therefore, they can be measured relatively easily through a variety of checklist-type tools. A number of tools to measure and track dosage and content have been cited in several large-scale programs, such as the Evidence-Based Home-Visiting (EBHV) initiative, the Early Head Start Family and Child Experiences Study (Baby FACES) and the Partnering with Families for Early Learning (PFEL) studies. These tools are most commonly paper questionnaires or checklists that the visitor or a trained observer fills out at the end of each planned visit. These questionnaires register information such as whether the visit was completed, its length and location, the participants, the activities completed or material covered (Barrett, Zaveri & Strong 2010; Paulsell et al 2010). Many programs have only recently begun to register dosage and content information, so there is no single protocol or consensus on how often or with what detail the data is collected (Wasik et al 2013).

Relationships, on the other hand, are a *process* element of quality. Process variables focus on more dynamic aspects of quality such as the implementation of the content and on the quality of interactions between the visitor, the caregiver and the child (Thomassen & La Paro 2009). Relationship quality is a more complicated and lengthy process to register and quantify, and it requires trained observers.

Dosage

Measuring implementation dosage, or the frequency and duration of visits, is important for cost-effectiveness calculations and for planning for how programs might scale up or be replicated in different contexts (Wasik et al 2013). Measuring dosage also provides information about the fidelity of the implementation strategy, specifically, if the goals of the program are being met in terms of number and duration of visits carried out (Paulsell et al 2010). This is important, because even programs considered to be well-implemented often have large differences between the number of visits they plan to provide families and the number of visits that families actually receive (O'Brien et al 2012; Durlak & DuPre 2008). As an example of a different aspect of dosage, in a program considered to be very well-implemented (the Nurse-Family Partnership), researchers found that families, on average, completed less than 70% of the visits planned; that is, the actual dosage of visits was much less than the intended dosage (Ingoldsby 2013; Riley et al 2008). Dosage may vary for any number of reasons, for example, families may not be found at the agreed-upon time or may refuse visits; alternatively, programs may experience staff turnover or logistical complications that affect the regularity of their visits.

The relationship between quality and dosage may not be entirely straightforward. For example, Stoltzfus & Lynch (2005) find evidence that the relationship between

exposure and outcomes is not linear; greater exposure does not necessarily produce better outcomes; however, a number of studies *do* find that decreased dosage has an adverse impact on child development outcomes. For example, Brown & Liao (1999) find that families who participate in visits inconsistently or drop out early (diminishing their overall dosage) are less impacted by the program. While there is no evidence on the minimum dosage or the optimal interval between visits that is required to produce child outcomes, along those same lines, Boller et al (2004) do find evidence that parent-focused interventions require multiple experiences to be successful and, specifically, that one dose of treatment/visit/workshop has not proven to be effective in changing targeted outcomes. The lack of clear and consistent evidence of a positive relationship between dosage and outcomes may be explained by the quality (or lack thereof) of the home visits; simply increasing the frequency or duration of home visits, if they do not meet a certain threshold for quality, will likely not improve children's outcomes. On the other hand, as previously mentioned, the problem may stem from a discrepancy between a program's *intended* dosage and its *actual* dosage, or what actually happens on the ground. Clearly, if home-visiting programs do not track the effective dosage that families are receiving, they have very little information to understand the program outcomes or to improve their service delivery.

In one of the only evaluations of dosage, Powell & Grantham-McGregor (1989) examined the effect of different frequencies of home visiting on child development in two separate studies with underprivileged urban children from the same neighborhoods in Jamaica. In the first, 152 children ages 6 to 30 months old received no visits, biweekly visits, or monthly hour-long visits by a paraprofessional, supervised by a nurse from a local health center. Only the children who received the most frequent visits—the biweekly group—showed small but significant increases in scores on the Griffiths Mental Development Scales (0.20 SD); the other groups reported no impacts. In the second study, 58 children from 16 to 30 months old were randomly assigned to weekly hour-long visits or a control group. The group receiving weekly home visits had significant improvement (1.15 SD) on the Griffiths Mental Development Scales (developmental quotient) and on the performance subscale. The authors concluded that as the dosage of home visits that children receive increases—from monthly to biweekly to weekly—the impacts on their development increase as well (Powell & Grantham-McGregor 1989).

Fewer studies have attempted to measure another, more specific aspect of dosage, which refers not to the frequency of visits but to **the amount of time that caregivers spend carrying out or participating in the activities assigned to them by their home visitor in the period between visits**. Wallander et al (2014) examined the relationship between the time spent by caregivers on their assigned activities and the developmental outcomes of children in India, Pakistan and Zambia who received biweekly home visits during the first three years of life. Children whose caregivers implemented their assigned activities with greater frequency generally had better developmental outcomes (Wallander et al 2014).

Content

Content, a second dimension of quality, refers to the curriculum that the home visitors cover during the visit, the topics they facilitate, and the information they impart or leave behind for the caregiver. Content also captures aspects of program fidelity in that it measures the extent to which the content *actually delivered* during a particular visit corresponds to what was *meant to be delivered*, given the child's age and

developmental level, what the home visitor had planned with her supervisor, and what was dictated by the program's curriculum. Content varies significantly from program to program, depending on program goals and outcomes. Forms to measure content might include questions: Did the home visitor bring the right material and toys to deliver the intended activities? Did she use the materials correctly? Did she adjust the level of difficulty to make sure the content was appropriate for the child's developmental level? Did she cover all of the topics and activities planned for the visit? Content measurement can also include how much time the visitor spent facilitating child-focused activities as opposed to building relationships with the caregiver (Hallgren et al 2010).

Measuring content can also include recording how much time was *not* spent delivering content; that is, time lost to distractions. Here, it is clear that understanding content is key for a complete definition of dosage; even when the intended dosage of intervention was delivered, program impacts can only be expected when time was spent on content delivery. Peterson et al (2007) highlight the importance of measuring content in their study, finding that, like dosage, the planned content for two different home-visiting programs—one of which was Early Head Start—was not being delivered according to program goals and was not delivered consistently across different families. That is, home visitors were meeting with families for regularly-scheduled visits but were not realizing the activities planned for each visit as specified by the program.

Some tension exists between rigidly adhering to content and flexibly changing the focus of a visit, depending on caregivers' wants or needs. For example, O'Brien et al (2012) find that the more successful home visitors from the Nurse-Family Partnership—that is, those who maintain long-term relationships with the families they visit, thus having a greater chance to impact child outcomes—are the ones who adjust content based on caregivers' individual needs and interests. In particular, visitors with the lowest attrition rates among their families were those who very specifically adapted the visits to the caregiver's unique needs and aspirations *rather* than faithfully and strictly adhering to the program content or even to the prescribed number of visits. This ability to respond effectively requires a level of training and receptivity on the part of visitors that many programs struggle to provide. For most programs, then, adherence to the program's prescribed content, even at the cost of individually-tailored home visits, is a way to achieve a minimum level of service delivery and ensures that participating families are all getting the intended treatment.

Relationships

Evidence supported by researchers and practitioners of home visiting suggests that at the heart of effective programs lies the creation of ongoing relationships between the visitor and caregiver that are stable, respectful, warm, honest, open and responsive, and that empower the caregiver (Paulsell et al 2010; Roggman et al 2006; Miller & Rolnick 2002; Riley et al 2008). Higher-quality visits, characterized by warm, encouraging and supportive relationships, are more successful in changing caregiver behavior and positively impacting child development (Paulsell et al 2010; Peterson et al 2007; Roggman et al 2006).

Despite the importance of the visitor-caregiver relationship, there is little information about how this relationship is formed, how it works in practice, and how it influences outcomes. Collecting information about the quality of relationships as

well as the actual interactions during a visit—the “black box” of what actually happens during a visit—is therefore key to measuring visit quality and to positively impacting child development outcomes. Roggman et al (2001) recognize that there are few methods for measuring the quality of relationships during home visits, and the instruments available are relatively complex.

Measuring the quality of relationships in a visit is clearly more complicated than simply recording dosage or content, because it requires an understanding of the specific kinds of interactions and behaviors that promote caregivers to change their own behavior in relation to their children. The quality of home-visit relationships and interactions may be measured by a visitor self-report, although for obvious reasons, self-reports of visitor behavior can be biased. Nonetheless, in recent years as there has been a greater focus on accurately measuring the quality of this aspect of the visit, a number of observation-based measurement tools to be applied by a trained observer have been developed for use.

There is some interplay between the quality of relationships and dosage, as briefly discussed earlier. Research suggests that one issue affecting the potential impact of home visiting on children’s outcomes might be take-up and drop-out rates; that is, the number of families who agree to participate in home visiting but never follow through to participate in the program, and those that begin but do not receive the complete intended dosage because they drop out.

A few studies have addressed the concept of take-up and drop-out, or retention, rates, and have concluded that sustained warm, personal relationships with visitors may be vital to recruiting and retaining families. In particular, the EBHV study identified family-provider relationships and the manner in which family needs are identified and addressed as two key predictors of family take-up of services or retention. This confirms previous empirical findings that emphasize the importance of home visitor-caregiver relationships as being at the heart of the home-visiting theory of change (Roggman et al 2008; Prinz et al 2001). In one example, Barnes et al (2006) conducted a study of take-up rates among families with a new baby who had been previously screened and identified as vulnerable, but who did not take up an offer of home-visiting support, as compared with families that used the services. The authors found that families that did not take up the service were more socially, educationally and economically disadvantaged and lived in poorer neighborhoods. They hypothesized that individualizing services to meet the specific needs of these families could lead to higher rates of take-up (Barnes et al 2006). In another study, Roggman et al (2008) examined drop-out rates among 564 families participating in the national Early Head Start home-visit program that offers weekly visits to families with children less than 3 years of age. It found that families that dropped out of the program experienced visits that were focused less on child development, with visitors who were less successful at engaging parents, and had more distractions, for instance, engagement in activities that took time away from the visit like disciplining other children, dealing with pets or other people, ongoing loud noises, and others. Similar to the findings from the previous study, the more disadvantaged families were also more likely to drop out of the program (Roggman et al 2008). Findings such as these suggest that the quality of the visits, as reflected in the quality of the unique and individualized relationships between the home visitor and caregivers, may be a key factor in ensuring that families remain in the program and receive their intended visit dosage.

Instruments to measure the quality of home visits

After discussing what ought to be critical elements of a high-quality home visit, this section describes **the instruments that have been developed to measure home visit quality, and specifically those that focus on the kinds of interactions and processes that take place during the visit.** Quality measurement has been identified as key to good program implementation and to impacting children's outcomes. Since the interactive processes that take place during the home visit are at the heart of programs' successful service delivery (Paulsell 2010), the measures discussed place emphasis on measuring the overall quality of the home visitor, who has the important role of transmitting parenting and developmental lessons to the caregiver, with the aim of changing her day-to-day interactions with her child.⁸

The most well-known instruments used to measure home-visit quality primarily focus on visitor-caregiver relationships but also capture aspects of dosage and content. They are presented in table 1. The table presents each instrument's stated purpose along with the specific areas it assesses. The table includes a column that provides information on the scoring and point-scale utilized by the instrument. Another column describes the way in which the administration of the instrument is designed to be carried out (live observation, video recording, or both) and whether the instrument can be used for supervision purposes (such as a feedback tool for supervisors to record information about a visit they observe to later provide coaching and lessons to home visitors). The table also reports whether the instrument is designed for use by a trained observer, program supervisor, or the home visitor herself. Most instruments are meant to be more in-depth measures scored by a third party who analyzes the specific interactions that take place between the home visitor and the family, while a few serve as monitoring tools that can be filled out by the home visitor herself. Lastly, the table reports parameters of the instruments' psychometric properties, such as internal consistency and, when available, reliability.

⁸ There are few tools that measure other structural aspects of program quality, such as recruitment, administrative factors, and staff characteristics, among others. One recently developed is the Home Visiting Program Quality Rating Tool (HVPQRT). For a review of this tool, see Korfmacher et al 2013.

Table 1: Instruments for Measuring Quality

Instrument	Purpose	Content	Scoring	Method	User		Validity
					Obs	HV	
HOVRS Roggman et al 2006	Assess effectiveness of home visitor in engaging parent and child	1. Facilitation of parent-child interaction 2. Relationship with family 3. Responsiveness to family 4. Non-intrusiveness 5. Parent-child interaction 6. Parent engagement 7. Child engagement	7-point scale 1- Inadequate 3- Adequate 5- Good 7- Excellent	Live or video recording Supervision and research purposes	X		<i>alpha</i> =.78
HOVRS-A Roggman et al 2010	Same as HOVRS	Same as above with modifications to items 2 and 7	5-point scale 1- Inadequate 3- Adequate 5- Good	Same as HOVRS	X		<i>alpha</i> =.87
HOVRS-A+V2 Roggman et al 2010	Same as HOVRS	Same as HOVRS-A	Same as HOVRS	Same as HOVRS	X		<i>alpha</i> =.88

Instrument	Purpose	Content	Scoring	Method	User		Validity
					Obs	HV	
Home Visit Content & Characteristics Form Boller et al 2009	Observe visit content, participant characteristics and activities of home visits	1. Number of children/adults that participated 2. Use of interpreter 3. Identification of family strengths and challenges 4. Language of visit 5. Checklist of activities covered during visit 6. Extent of environmental distractions 7. Time allocation of activities	Environmental distractions recorded on 1-6 scale [1 very interfering, 5 not interfering, 6 NA] Time allocation recorded on 0-3 scale [0 not addressed; 1 touched on briefly; 2 discussed at least 10-15 minutes; 3 primary focus]	Originally developed to supplement the HOVRS form Live or video recording	X		N/A
Home Visit Encounter Form	Measure fidelity by recording time, frequency and content of home visits	1. Length and date of visit 2. Location of visit 3. If visit was completed 4. Topics covered and time allocation 5. Percentage of planned content completed	Time allocation of activities recorded as percentage of total time of visit	Typically completed in-person during every home visit by home visitor		X	N/A
Home Observation Visit Form McBride et al 1993	Assess the quality of home visits through interval coding	1. Individuals present 2. Primary interaction type 3. Content of interaction 4. Nature of home visit interaction	Observations are coded in 30-second intervals	Live or video recording Supervision and research purposes	X		Overall inter-rater reliability= 85%

Instrument	Purpose	Content	Scoring	Method	User		Validity
					Obs	HV	
Home Visit Assessment Instrument Wasik & Sparling 1995	Measure the overall knowledge and behavior of the home visitor	Section 1: Pre-visit interview (home visitor asked questions about the upcoming visit) Section 2 contains 10 categories: 1. Family needs 2. Child focus 3. Parent-child focus 4. Family 5. Health/safety 6. Parenting/problem solving 7. Case management 8. Closure & planning 9. Clinical/interviewing skills Section 3: Post-visit assessment	The categories in section 2 have a number of items that are scored on a scale from 0-3 with an N/A option. Sections 1 & 3 are interviews of the home visitor about the visited family, visit content, and dosage	Sections 1 and 3 are in-person and section 2 can be either in-person or video recording	X		N/A
COACH Dishion et al 2010	Assess the overall ability of the home visitor in providing services	1. Conceptual understanding of model 2. Observant and responsive to needs 3. Actively structuring sessions 4. Careful and appropriate 5. Hope- and motivation-inducing	1-3 – Needs work 5-7 – Good 7-9 – Exceptional	Used for live observation Developed for supervision		X	Inter-rater reliability = 67%

The most extensively validated observation instrument in home visiting to date is the Home Visit Rating Scale series (HOVRS, HOVRS-A and HOVRS-A+). All iterations of the HOVRS instrument are used to describe and assess strategies implemented during home-visiting interventions. With the home visitor as the unit of analysis, these measures are particularly effective in evaluating process aspects of the visit and the relationships observed, specifically, the home visitor's effectiveness in engaging the caregiver and child during activities and interactions throughout the course of the visit. One of the strengths of the HOVRS family of instruments is that it has proven adaptable to diverse program goals, visit formats, and different cultures, such as Spanish-only Latino families, rural Caucasian families, and urban African-American families.

The original HOVRS measure, developed in 2006, consists of a total score with 2 subscales divided among 7 items. The Home Visitor Strategies Quality subscale takes into account items such as facilitation of parent-child interactions, relationships with the family, responsiveness to the family, and non-intrusiveness. The Effectiveness Quality subscale incorporates the overall parent-child interactions and parent and child engagement during the visit. Each item is scored on a 7-point scale, with anchor points of 1 (inadequate), 3 (adequate), 5 (good), and 7 (excellent) (Roggman et al 2006).

HOVRS was revised and adapted by Mathematica Policy Research to both simplify its administration and to capture a larger range of attributes of visit quality. The revised version, called HOVRS-A, differs from the earlier one in three ways: (1) items are scored on a 1-5 point scale (with anchor points of 1-inadequate, 3-adequate, and 5-good), (2) indicators are aligned across each of the three anchor points to ensure consistency, and (3) two scale items (home visitor relationship with family and child engagement) are slightly modified (Roggman et al 2010).

In the past year, the HOVRS-A+, Version 2 was developed to combine elements of both the HOVRS and HOVRS-A. The HOVRS-A+ uses the 7-point scale of the original HOVRS with the adaptations made to the home visitor relationship with family and child engagement subsections. The HOVRS-A+ is currently the instrument with the highest internal consistency reliability that has been validated for home visiting quality measurement (Roggman et al 2010). As table 1 indicates, the HOVRS-A+ has achieved the highest internal consistency of the three HOVRS measures ($\alpha = .88$).

Several program models have used parts or all of the HOVRS-A+ to measure their own program quality. For example, the Parents as Teachers model created a measure known as the Personal Visit Observation Tool to assess program quality, which uses four of the seven scales from HOVRS-A+.⁹ Furthermore, the state of Idaho has provided guidelines for organizations competing for government funding on how to adapt HOVRS-A+ to the Nurse-Family Partnership, Parents as Teachers, and Early Head Start program models for measuring quality.¹⁰

⁹ The Personal Visit Observation Tool includes three sections: (1) Visit Preparation (2) Personal Visit Elements, and (3) Visit Quality. The HOVRS-A+ has been adapted in the latter section. The scales used are the Home Visitor Responsiveness to Family, Home Visitor-Family Relationship, Home Visitor Facilitation of Parent-Child Interaction and Home Visitor Non-Intrusiveness/Collaboration with Family. For more information, see http://www.roe26.net/pdfs/early_beginnings/ParentsasTeachersPersonalVisitObservationandTheHomeVisitRatingScale_HOVRS.pdf

¹⁰ See <http://healthandwelfare.idaho.gov/Portals/0/Children/HomeVisiting/MIECHV%20Assessment%20Guide%204-9-2013.pdf>

Complementing the HOVRS measures, the supplemental Home Visit Content and Characteristics Form (HVCCF) can also be used on its own to document a range of structural characteristics of the home visit, such as those mentioned in the dosage and content section above. The form also takes into account the number of participants in the visit and the language used during the home visit, measures time allocated for each activity, and records delivery of different content topics during each visit (Boller et al 2009).

The Home Visit Observation Form (HVOF) focuses primarily on content delivery and process evaluation, that is, the ways and degree to which the visitor is successful in engaging the family. To administer this instrument, observations are coded in 30-second intervals in which the observer is given 25 seconds to observe and 5 seconds to record the quality of the activities observed. Observations are made in the following four areas: (1) individuals present, (2) primary interaction type, (3) content of interaction, and (4) nature of home visit interaction (McBride & Peterson et al 1993).

The COACH measure was developed for the Family Check-Up (FCU) home-visiting program and focuses primarily on fidelity of program content delivery. FCU, developed in the state of Oregon, is aimed at reducing children's behavioral, academic, and internalizing problems, and making improvements in maternal depression, parental involvement, and positive parenting. The program model consists only of a three-session intervention (an initial contact, a home-based observational assessment, and a feedback session; Dishion, et al 2008). The theory behind COACH is that in order for home-visiting interventions to be successful, the visitor must strictly adhere to delivering the intended program content in the way it was designed. This measure evaluates the visitor across five dimensions, including (1) conceptual understanding of the program model, (2) observant and responsive to needs, (3) actively structuring sessions, (4) careful and appropriate teaching, and (5) hope- and motivation-inducing. The measure is administered throughout the visit. For supervision purposes, one 15-minute segment is rated for each completed visit, while the entire visit is observed only for research purposes. Items are scored on a 9-point scale, with anchors of needs work (1-3), good (4-6), and exceptional (7-9) (Dishion et al 2010).

In addition to the instruments described above, several others have been developed to measure quality of home-visiting programs. Fewer references were found in the literature documenting experiences with their use and administration. **The Supportive Interactions with Families (SIF), a self-rating scale,** measures the quality of interactions between the home visitor and the family. It can be used during live or video-recorded observations. Five dimensions are assessed on a 1-5 scale: home visitor focus, communication skills, support of parent-child interactions, problem-solving skills and professionalism (Twombly, Waddell & Harrison 2003). **The Home Visit Assessment Instrument (HVAI)** measures the overall behavior of the visitor in her interactions with caregivers and children. It is purely observational and can be used for both supervision and professional development purposes. The HVAI is divided into three sections: pre-visit details, observation of the home visit, and the post-visit details. The first and last sections are completed by interview and are only necessary for research purposes. The middle section evaluates home visitor effectiveness across ten areas: family needs, child focus, parent-child focus, family, health and safety, parenting and problem solving, case management, planning and closure, clinical skills, and post-

assessment. Each item is rated on a 0-3 point scale (Wasik & Sparling 1995). **The Home Visitation Developmental Assessment Scale (HVDAS)** is another instrument, separated into three different areas related to the home visitor: communication skills, problem solving, and self, including character, attitude, and values. The instrument uses a 5-point scale, in which a rating of 1 is lowest quality and 5 is highest (Keim 2011).

Use of instruments to measure home-visit quality as part of program evaluations

This section provides an overview of how some of the best-known instruments to measure home-visit quality have been used to evaluate home-visiting programs in the US. Important characteristics of each program evaluation related to the choice of instrument and its administration are summarized in table 2, at the end of the section.

Better Beginnings: HOVRS-A and HVCCF

The Better Beginnings study evaluated the quality and impact of home visits with families as part of the Early Learning Initiative, which was initiated by the Gates Foundation. A total of 35 families from suburban Seattle and central Washington state were included in the sample.

The HOVRS-A and Home Visit Characteristics and Content Form (HVCCF) were used to assess program quality for families with children between the ages of 1 week and 9 months. In order to assess the overall quality of home visits, trained observers carried out three rounds of live observations in which they accompanied each home visitor. A total of 35 observations were collected, whose average length was 61 minutes. The total HOVRS-A score was only computed for 31 of the 35 visits; in the remaining visits, the child was asleep for more than 75% of the total visit time (Hallgren et al 2010).

Scores obtained through the administration of HOVRS-A were high across all dimensions. Overall, the mean score of HOVRS-A was 4.2 out of 5 (SD= 0.59). The Home Visitor Strategies Quality subscale mean score was 4.1 (SD= 0.55), and the Effectiveness Quality subscale mean score was 4.3 (SD=0.79). On the whole, the results show that home visitors were successful at developing positive relationships with families and engaging parents and children. The area in which home visitors scored lowest was the facilitation of parent-child interaction (mean=3.8), suggesting that while parents did make efforts to engage in activities with their children, there were cases in which the home visitor only interacted with the parent rather than facilitating interactions among family members (Hallgren et al 2010).

The HVCCF was utilized with HOVRS-A. The HVCCF documented that 57% of all visits were conducted in English and 43% in Spanish. Additionally, 57% of all visits included another related adult present during the visit, and in 33% of cases there was another child present. On a similar note, distractions that interrupted the delivery of the visit content were reported in 33% of all observations. Twenty-six percent of distractions were due to other children and 11% were due to television, radio or computer games.

The time allocation of activities collected by the HVCCF provides important information on the extent to which certain topics are covered and the nature of home visits. In 51% of visits, home visitors and families discussed issues that had been identified during

previous visits (problems, progress), demonstrating continuity and building on experiences from past visits; however, home visitors provided feedback on child interactions in only 11% of observed visits. Sixty-three percent of the home visitors devoted 10 to 15 minutes of the visit to fostering relationships between the child and the caregiver (discussing infant cues and appropriate expectations). It is interesting to contrast this result with that of the HOVRS-A scale, in which the facilitation of parent-child interaction received the lowest score, suggesting that while this kind of interaction was taking place, it may not have been implemented in an engaging and high-quality manner.

The program staff in each community provided feedback about their experience with the instrument; in particular, they raised concerns about the use of HOVRS-A and HVCCF for monitoring and supervision purposes. Specifically, they stressed that modifications were necessary to determine how to appropriately weigh HOVRS-A scales and indicators and establish threshold scores for minimum acceptable quality levels. They also agreed more work was needed on how to use it for staff training purposes. Other worries related to the instrument's use include that it was challenging to assess interactions among families with young infants. Additionally, there were concerns that the instrument might not account for cultural differences in parent-home visitor engagement and might not fully differentiate among upper-end scores (Hallgren et al 2010).

The feedback provided by program staff suggested that the main challenge in administering the HVCCF was to align it with the specific program curriculum (Hallgren et al 2010); however, from an evaluation standpoint, this could also be considered a strength of the instrument, in that it is not aligned with any particular curriculum and can be used in any variety of program formats.

Baby FACES: HOVRS-A and HVCCF

The HOVRS-A and the HVCCF are also both currently being used for the Baby FACES Evaluation. Baby FACES, The Early Head Start Family and Child Experiences Survey, is an ongoing, longitudinal, descriptive evaluation of the Early Head Start (EHS) program, staff, services, and families intended to shape better policy and practice both locally and nationally. Starting in 2009, the study enrolled two cohorts of children from 89 EHS programs around the country: a newborn cohort (children less than two months old) of 194 pregnant women, new mothers and babies, and a 1-year-old cohort of 782 children between 10 and 15 month of age. Data on the children's experiences and outcomes was collected each spring until the children turned 3 or left the program.

Vogel et al (2015) focus their study exclusively on the 2-year-olds. The study examines a number of aspects of home visits, including staff turnover, family satisfaction, and program and home visitor quality. The study found that families received roughly 37 visits a year from trained home visitors, nearly 75% of whom held a higher degree in a field related to early childhood education. The authors report that the frequency of home visits fell slightly below an average of one visit per week during the summer and winter months, and while group sessions were also offered to families receiving home visits, fewer than 50% of families were reported to have attended these meetings.

For the HOVRS-A sample, a total of 242 families were observed. The HOVRS-A composite mean score was 3.4 with an SD of 0.09, and 41% of all visits scored in the medium quality, or 3 to 4 range. The Visitor Effectiveness subscale scored comparatively higher (mean=3.6; SD= 0.11) than the Visitor Strategies subscale (mean=3.2; SD = 0.09). For the latter subscale, home visits scored lowest in the areas of facilitation of parent-child interaction (mean=2.8; SD=0.12) and non-intrusiveness (mean=2.9; SD=0.12). These low scores imply that while there is substantial room for improvement, home visitors occasionally implement techniques that do aim to develop relationships between parents and children. In addition to this, the relatively higher score on the child engagement area (mean=4.2; SD=0.08) demonstrates that children often interacted with their parents and showed interest in the activities that occurred during home visits (Vogel et al 2015).

Based on the HVCCF, Vogel et al (2015) find that visits, on average, lasted 77 minutes, significantly less time than what was expected based on the program model (90 minutes). Only 35% of home visits lasted the full 90 minutes. Nearly all (97%) of home visits included in the sample took place in English, with the remainder in Spanish.

Regarding time allocation, the HVCCF found that almost half (49.7%) of the visits were spent on child-focused activities. This seems fitting given the high ratings that child engagement received on the HOVRS-A. Specifically, 86.1% of visits involved play, and the majority (67.1%) involved the provision of educational services. It is also interesting to note that 44% of visitors spent time modeling or facilitating parent-child interactions (Vogel et al 2015).

The use of HOVRS-A in the Baby FACES Evaluation also provides information on the linkage between home visitor quality and home visitor attributes, program characteristics, and home visit activities and characteristics (Vogel et al 2015). Using a sample of 185 home visitors, the authors find very low levels of correlation (never larger than 0.2) between home visitor characteristics and the HOVRS-A total score (and that of its Visitor Strategies and Visitor Effectiveness scales). Interestingly, the correlation between home-visit quality and the educational level of the home visitor was not statistically significant, although having a credential in home visiting (that is, some specific training) is positively associated with better visit quality. This is an interesting finding for developing countries implementing home-visiting programs in contexts where the overall level of formal schooling is substantially lower than in the US, but where practical and specific training for home visitors is a possibility. The correlation between HOVRS-A scores and whether the program has unfilled staff positions is negative but very low. This suggests that the quality of home visits decreases with higher rates of unfulfilled staff positions, high staff turnover and higher caseloads. These results are consistent with those of O'Brien et al (2012) who examined attrition rates among families in a nationwide sample of Nurse-Family Partnership programs and found that families were nearly eight times more likely to drop out of the program and had fewer completed visits if their original visitor left the program before their child's first birthday.

The Baby FACES evaluation by Vogel et al (2015) also finds that home visit length is positively correlated to overall quality and to quality as measured just by the Visitor Strategies scale. Furthermore, the number of children participating in the visit is negatively correlated to overall quality and also to both the Visitor Strategies and the Visitor Effectiveness scales.

Illinois Prevention Initiative - HOVRS-A+

In 2012, the Illinois Prevention Initiative (IPI) conducted a statewide evaluation of 30 home visiting programs that was designed and carried out by the Erikson Institute. A total of 85 families participated in the study, in which 73% received services based on the Parents as Teachers model, 23% on the BabyTalk model and 3% on the Healthy Families America model. The HOVRS-A+ was used to measure program quality.

Home visits were filmed and then scored based on the video recordings. Coders were rigorously trained. They were deemed reliable once 70% of scores for each indicator were within one point of the reference score. On average, video recordings lasted 32 minutes. The mean composite HOVRS-A+ score fell in the adequate-to-good quality range (mean=4.06, SD=0.99). In the Visitor Strategies subscale, home visits fell in the adequate-to-good range (mean=3.71, SD=1.03), as did Visitor Effectiveness subscale (M=4.59, SD=1.11). Home visitor strategies and effectiveness were significantly and positively related to one another ($r=0.77$).

On the Visitor Effectiveness subscale, home visitors were found to do better in developing positive relationships with the caregiver and the child and in encouraging collaboration than they did in facilitating interactions between the caregiver and the child. The relatively lower scores in the facilitation of caregiver-child interactions are consistent with findings from other studies (Korfmacher et al 2012).

Part C and Early Head Start in Central Iowa: HVOF

Peterson et al (2007) study home visits in central Iowa using the HVOF to collect data on the quality of home visits for participants of the Part C program and for families in an EHS program.¹¹ A total of 28 families were included in the Part C study and 92 families in the EHS study. A total of 1,131 visits were observed over four years. Graduate research assistants who had prior experience using the HVOF conducted live observations, making concerted efforts not to be intrusive. These observations began approximately five minutes after the visit started. Coding with the HVOF was typically done over cycles of 10 minutes with two-minute breaks until the home visit ended or until 60 minutes of data was coded. Observation time ranged from 21 to 44 minutes (mean=34) for Part C participants and from 11 to 60 minutes (mean=51) for EHS participants.

There were several important findings. EHS home visitors dedicated the majority of time during their visit to interactions with caregivers (between 59% and 86%) and very little time to interacting with the child alone (between 2% and 5%). Perhaps the most important finding was that very little time was devoted to enhancing caregiver behaviors through direct teaching (2% of time), discussing methods (6% of time), or coaching caregiver-child interactions (3% of time). Higher levels of engagement among caregivers were recorded on topics related to child development content (26% of time) and on

¹¹ The HVOF was subject to some revisions for its administration in the Early Head Start sample. The revisions included minor changes such as the addition of codes to reflect specific topics and behaviors believed to be unique to their EHS program model.

family member functions (10% of time). Caregivers were less engaged when home visitors addressed community resources and referrals (1% of time; Roggman et al 2008).

Evaluation of Family Check-Up Model: COACH

A randomized controlled trial study involving 79 high-risk families with children 2 to 4 years of age was conducted to assess the overall fidelity of visitors (who were trained therapists) using COACH and to correlate it with observed participant engagement. Subsequent evaluations were conducted with the families of this study when children turned 3 and 4 years old to track parents' parenting style (involvement, parent-child interactions and proactivity; Dishion et al 2008) and children's behavioral problems. The study was carried out in three states.

COACH coders were instructed to observe the feedback sessions that visitors carried out with families. The coders had advanced educational experience in child development and appropriate training on COACH. In addition to the COACH dimensions, observers simultaneously rated caregiver engagement with the same scoring scale they used to rate the home visitor. Coders focused on both content aspects of the visit (or adherence to the program curriculum) as well as the overall skill in the delivery of the contents (home visitor competence).

The results show that home visitors maintained strong fidelity in the delivery of the program; that is, they were able to capably deliver content as intended. The mean scores of the five COACH dimensions fell in a range of good quality (between 5.40-5.81). The study also documented a high correlation across the instrument's five dimensions (Dishion et al 2010).

V. Quality of home visits and its effects on child development

Two of the studies described in this section go a step further by examining not only the quality and characteristics of home visits, but also whether visit quality is associated with positive outcomes for children and caregivers. This section summarizes the findings from these studies.

In the Baby FACES study, Vogel et al (2015) explore the relationship between program quality as measured by HOVRS-A and key child development and family outcomes at ages 2 and 3 years. They find that average home visit quality does not appear to be related to any of the age 3 outcomes of interest, although there are a few significant age 2 findings. The authors find a significant, positive correlation between the HOVRS-A Visitor Effectiveness scale and children's engagement and emotional regulation scores as measured by the Bayley Behavior Rating Scale (0.13 and 0.14 SD). Negative associations were found between the Bayley's and home visitor reports of children's problem behaviors on the Brief Infant-Toddler Social Emotional Assessment, BITSEA (-0.17 SD). Perhaps most surprisingly, the Ages and Stages Questionnaire's (ASQ-3) gross motor, fine motor and problem-solving scales had significant negative correlations with home visit quality as measured by the HOVRS-A (between -.14 and -.17 SD). No significant relationships were found between home-visiting quality and quality of the home environment as measured by the HOME (Vogel et al 2015, p.65).

For the Family Check-Up program, Smith et al (2013) conducted a randomized prevention trial with a sub-sample of 79 participating families, to examine the relationship between FCU visit quality and caregiver and child outcomes. Families participating in the FCU were from a number of US States and screened to receive FCU services based on their participation in the national Women, Infants, and Children (WIC) Nutritional Supplement program. Specifically, the evaluation used the COACH instrument to measure the quality of home visit implementation and caregiver engagement, and the impact of those variables on improving caregivers' positive behavior support of their children and in reducing children's reported problem behavior. Using an observational coding system developed specifically for the intervention, the authors found that one year after the intervention, caregiving had improved a small but significant amount (0.06 SD), and two years after the intervention, children's problem behavior was significantly reduced (-0.24 SD) (Smith et al 2013).¹² They conclude that the FCU was effective for preventing early childhood problem behavior among families participating in the WIC program.

¹² Child behavior problems were measured by the Child Behavior Checklist, a 99-item questionnaire that was completed by the caregiver (Smith et al 2013).

Table 2: Instruments Used to Assess Program Quality in the United States

Program evaluation	Instrument	Sample size	Program Models	Context	Observer profile	References
Better Beginnings	HOVRS-A	31 families	Partnering with Families for Early Learning	Washington state Children 0-9 months of age from predominantly middle- and working-class white and Hispanic families	Trained observers with past education and experience in child development	Hallgren et al 2010
	HVCCF	35 families				
Baby FACES Cross-Site Evaluation	HOVRS-A	220-225 families	Early Head Start	Low-income population, with 49% of families belonging to minorities. Sample drawn from 89 EHS programs around the country, including pregnant mothers and children through 3 years of age	Child development professionals that completed training. Certification required coding three videos with 80% reliability	Vogel et al 2015
	HVCCF					
Illinois Prevention Initiative (PI) Birth to Three	HOVRS-A+ V. 2	85 families	Parents as Teachers (73%), BabyTalk (23%), Healthy Families America (3%)	Illinois representative sample of expecting parents and families with children ages birth to 3 from 30 sites. 32% Caucasian, 31% African-American and the rest Hispanic	Coders went through training and were deemed reliable once 70% of scores for each indicator were within one point of the reference score	Korfmacher et al 2012
Evaluation of Part C and EHS Services	Home Observation Visit Form	Part C - 28 families; EHS – 92 families	Part C Early Head Start	Central Iowa, the majority of participants were rural Caucasian mothers and children 6 mo-3 yrs with disabilities	Graduate research assistants with prior experience in use of instruments	Peterson et al 2007

Program evaluation	Instrument	Sample size	Program Models	Context	Observer profile	References
Evaluation of Family Check-Up	COACH	79 families	Family Check Up	Virginia, Oregon and Pennsylvania 2-year-old children High-risk indigent families previously identified as having behavior problems. 30% were African American and 7% Hispanic	University students and a professional social worker with 20 hours of training in COACH	Dishion et al 2008

Concluding thoughts

Home-visiting programs have grown in popularity in LAC in recent years; however, in the region, the knowledge of how to implement them effectively at scale is still in its earliest stages. As home-visiting programs continue to expand in LAC, practitioners, policy makers, and researchers will increasingly need tools to measure and understand the process that take place during the visit, in order to improve the service being delivered and ultimately developmental outcomes for the children being served.

Producing high-quality home-visiting services requires understanding of what the key aspects of quality are and how they can be consolidated across program staff and service delivery by means of training, mentoring, monitoring, and evaluation activities. It is with the objective of understanding the key characteristics that compose a high-quality, effective home visit that numerous researchers and practitioners have developed and administered instruments aimed at measuring key attributes of home visitors, home visits, and their content.

By and large, the studies described in this literature review provide descriptive information on home-visiting quality, by focusing on the home visitor as the unit of analysis, to determine whether interventions are implemented as intended and to document processes, including the activities and the quality of the relationships that take place during a home visit. These studies constitute a first attempt at describing what is occurring inside the “black box” of home visits, which may be important for their success in improving child outcomes. The evidence from programs discussed here documents enough heterogeneity in the findings to suggest the need to better understand what goes on during a home visit and, in particular, to understand which aspects are the most critical for the intervention to be successful.

As can be learned from the experiences reviewed, the administration of these instruments has its own challenges, particularly when used as periodic monitoring and mentoring tools in large-scale programs. Videotaping visits and carrying out reliable coding by highly-trained personnel can be costly at scale in many developing-country settings. This motivates the need to develop alternative instruments that can capture some key aspects of quality while being easy to collect and process routinely, in a cost-effective manner. The aim is for these instruments to guide program training, management, and mentoring activities on an ongoing basis. This emphasizes, once again, the crucial need to identify the key aspects of a home visit, as well as the easily-observable variables that can best capture them. This goal inevitably calls for more research on aspects related to the measurement of quality of home visits, preferably carried out in varied contexts and program settings. The IADB’s current research agenda on home visiting seeks to contribute to filling precisely these gaps in knowledge.

Bibliography

Attanasio, O. P., Fernández, C., Fitzsimons, E. O., Grantham-McGregor, S. M., Meghir, C., & Rubio-Codina, M. (2014). Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in Colombia: cluster randomized controlled trial. *BMJ*, 349, g5785.

Baker-Henningham, H., & López Boo, F. (2014). "Intervenciones de estimulación infantil temprana en los países en vías de desarrollo: lo que funciona, por qué y para quién." *Económica* 60: 120-186.

Barnes, J., MacPherson, K., & Senior, R. (2006). Factors influencing the acceptance of volunteer home-visiting support offered to families with new babies. *Child & Family Social Work*, 11(2), 107-117.

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.

Boller, K., Vogel, C., Cohen, R., Aikens, N., & Hallgren, K. (2009). "Home Visit Characteristics and Content Form." Princeton, NJ: Mathematica Policy Research.

Boller, K., Vogel, C., Johnson, A., Novak, T., James-Burdumy, S., Crozier, L., & Uhl, S. (2004). Using Television as a Teaching Tool: The Impacts of "Ready To Learn" Workshops on Parents, Educators, and the Children in Their Care. Final Report. MPR Reference No. 8778-750. Mathematica Policy Research, Inc.

Brown, C. H., & Liao, J. (1999). Principles for designing randomized preventive trials in mental health an emerging developmental epidemiology paradigm, *American Journal of Community Psychology* (Vol. 27, pp. 673-710). New York, USA: Kluwer Academic/Plenum Publishers.

Conhecendo o Programa Primeira Infância Melhor. Contribuições para políticas públicas na área do desenvolvimento infantil (2007) – 1a Edição. Porto Alegre: UNESCO.

Daro, D. (2006). Prenatal/postnatal home visiting programs and their impact on young children's psychosocial development (0-5): Commentary on Olds, Kitzman, Zercher and Spiker. Rev ed. Tremblay RE, Barr RG, Peters RDeV, eds. *Encyclopedia on Early Childhood Development*, 1-5.

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.

Dishion, T. J., Shaw, D., Connell, A., Gardner, F., Weaver, C., & Wilson, M. (2008). The family check-up with high-risk indigent families: Preventing problem behavior by increasing parents' positive behavior support in early childhood. *Child development*, 79(5), 1395-1414.

Doyle, O., Harmon, C., Heckman, J. J., Logue, C., & Moon, S. (2013). Measuring investment in human capital formation: An experimental analysis of early life outcomes (No. w19316). National Bureau of Economic Research.

Duggan, A., McFarlane, E., Fuddy, L., Burrell, L., Higman, S. M., Windham, A., & Sia, C. (2004). Randomized trial of a statewide home visiting program: Impact in preventing child abuse and neglect. *Child Abuse & Neglect*, 28(6), 597-622.

Duggan, A., Windham, A., McFarlane, E., Fuddy, L., Rohde, C., Buchbinder, S., & Sia, C. (2000). Hawaii's healthy start program of home visiting for at-risk families: evaluation of family identification, family engagement, and service delivery. *Pediatrics*, 105(Supplement 2), 250-259.

Durlak, J. A., & DuPre, E. P. (2008). Implementation Matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41, 327-350.

Engle, P. L., Black, M. M., Behrman, J. R., De Mello, M. C., Gertler, P. J., Kapiriri, L., ... & International Child Development Steering Group. (2007). Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *The Lancet*, 369(9557), 229-242.

Fernald, L., Gertler, P., & Neufeld, L. (2009). "Ten-year impact of Oportunidades—Mexico's Conditional cash transfer program—on child growth, cognition, language and behavior". *The Lancet*, 374(9706): 1997-2005.

Grantham-McGregor, S. M., Walker, S. P., Chang, S. M., & Powell, C. A. (1997). Effects of early childhood supplementation with and without stimulation on later development in stunted Jamaican children. *The American journal of clinical nutrition*, 66(2), 247-253.

Hallgren, K., Boller, K., & Paulsell, D. (2010.) Better beginnings: Partnering with families for early learning home visit observations. Princeton, NJ: Mathematica Policy Research. Document No. pp. 10-61.

Hamadani, J. D., Huda, S. N., Khatun, F., & Grantham-McGregor, S. M. (2006). Psychosocial stimulation improves the development of undernourished children in rural Bangladesh. *The Journal of nutrition*, 136(10), 2645-2652.

Hill, J., Brooks-Gunn, J., & Waldfogel, J. (2003). Sustained effects of high participation in an early intervention for low-birth-weight premature infants. *Developmental Psychology*, 39, 730- 744.

Home Visit Rating Scales-Adapted (HOVRS-A). Unpublished document, Princeton, NJ: Mathematica Policy Research.

Howard, K. S., & Brooks-Gunn, J. (2009). The role of home-visiting programs in preventing child abuse and neglect. *The Future of Children*, 19(2), 119-146.

Ingoldsby, E. M., Baca, P., McClatchey, M. W., Luckey, D. W., Ramsey, M. O., Loch, J. M., ... & Olds, D. L. (2013). Quasi-experimental pilot study of intervention to increase

participant retention and completed home visits in the nurse–family partnership. *Prevention science*, 14(6), 525-534.

Jones Harden, B., Chazan-Cohen, R., Raikes, H., & Vogel, C. (2012). Early head start home visitation: The role of implementation in bolstering program benefits. *Journal of Community Psychology*, 40(4), 438-455.

Keim, S. A., Daniels, J. L., Dole, N., Herring, A. H., Siega-Riz, A. M., & Scheidt, P. C. (2011). A prospective study of maternal anxiety, perceived stress, and depressive symptoms in relation to infant cognitive development. *Early human development*, 87(5), 373-380.

Kendrick, D., Elkan, R., Hewitt, M., Dewey, M., Blair, M., Robinson, J., ... & Brummell, K. (2000). Does home visiting improve parenting and the quality of the home environment? A systematic review and meta-analysis. *Archives of disease in childhood*, 82(6), 443-451.

Kitzman, H. J., Olds, D. L., Cole, R. E., Hanks, C. A., Anson, E. A., Arcolego, K. J., ... & Holmberg, J. R. (2010). Enduring effects of prenatal and infancy home visiting by nurses on children: follow-up of a randomized trial among children at age 12 years. *Archives of Pediatrics & Adolescent Medicine*, 164(5), 412-418.

Kitzman, H., Olds, D. L., Henderson, C. R., Hanks, C., Cole, R., Tatelbaum, R., ... & Barnard, K. (1997). Effect of prenatal and infancy home visitation by nurses on pregnancy outcomes, childhood injuries, and repeated childbearing: a randomized controlled trial. *Jama*, 278(8), 644-652.

Korfmacher, J., O'Brien, R., Hiatt, S., & Olds, D. (1999). Differences in program implementation between nurses and paraprofessionals providing home visits during pregnancy and infancy: a randomized trial. *American Journal of Public Health*, 89(12), 1847-1851.

Korfmacher, J., Sparr M., Chawla, N., Fulford, J., & Fleming, J. (2012). Illinois Prevention Initiative (PI) Birth to Three Program Evaluation. Chicago, IL: Erikson Institute.

López-Boo, F. L., Palloni, G., & Urzua, S. (2014). Cost–benefit analysis of a micronutrient supplementation and early childhood stimulation program in Nicaragua. *Annals of the New York Academy of Sciences*, 1308(1), 139-148.

Love, J., Kisker, E., Ross, C., Schochet, P., Brooks-Gunn, J., Boller, K., et al (2001). *Building their futures: How Early Head Start programs are enhancing the lives of infants and toddlers in low-income families. Summary report.* Report to Commissioner's Office of Research and Evaluation, Head Start Bureau, Administration on Children, Youth and Families, and Department of Health and Human Services. Princeton, NJ: Mathematica Policy Research.

McBride, S.L., & Peterson, C.A. (1993). Home visit observation form. Unpublished manuscript, Iowa State University.

Miller, W. R. & Rollnick, S. (2002). *Motivational Interviewing: preparing people for change* (2nd ed.). New York: Guilford Press.

Nahar, B., Hossain, M. I., Hamadani, J. D., Ahmed, T., Huda, S. N., Grantham-McGregor, S. M., & Persson, L. A. (2012). Effects of a community-based approach of food and psychosocial stimulation on growth and development of severely malnourished children in Bangladesh: a randomised trial. *European journal of clinical nutrition*, 66(6), 701-709.

Nievar, M. A., Van Egeren, L. A., & Pollard, S. (2010). A meta-analysis of home visiting programs: Moderators of improvements in maternal behavior. *Infant mental health Journal*, 31(5), 499-520.

Nurse-Family Partnership. (2015). Retrieved March 22, 2015, from <http://www.nursefamilypartnership.org/>.

O'Brien, R. A., Moritz, P., Luckey, D. W., McClatchey, M. W., Ingoldsby, E. M., & Olds, D. L. (2012). Mixed methods analysis of participant attrition in the nurse-family partnership. *Prevention Science*, 13(3), 219-228.

Olds, D. L. (2002). Prenatal and infancy home visiting by nurses: From randomized trials to community replication. *Prevention Science*, 3(3), 153-172.

Olds, D. L., Eckenrode, J., Henderson, C. R., Kitzman, H., Powers, J., Cole, R., ... & Luckey, D. (1997). Long-term effects of home visitation on maternal life course and child abuse and neglect: Fifteen-year follow-up of a randomized trial. *Journal of the American Medical Association*, 278(8), 637-643.

Olds, D. L., Henderson, C. R., Chamberlin, R., & Tatelbaum, R. (1986). Preventing child abuse and neglect: a randomized trial of nurse home visitation. *Pediatrics*, 78(1), 65-78.

Olds, D. L., Robinson, J., O'Brien, R., Luckey, D. W., Pettitt, L. M., Henderson, C. R., ... & Talmi, A. (2002). Home visiting by paraprofessionals and by nurses: a randomized, controlled trial. *Pediatrics*, 110(3), 486-496.

OPRE report no. 2015-10. Washington, DC: US Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation.

Paulsell, D., Avellar, S., Martin, E. S., & Del Grosso, P. (2010). Home visiting evidence of effectiveness review: Executive summary (No. 5254a2ab30e146ce900220dbc4f41900). Mathematica Policy Research.

Peterson, C. A., Luze, G. J., Eshbaugh, E. M., Jeon, H. J., & Kantz, K. R. (2007). Enhancing parent-child interactions through home visiting: Promising practice or unfulfilled promise? *Journal of Early Intervention*, 29(2), 119-140.

Powell, C., & Grantham-McGregor, S. (1989). Home visiting of varying frequency and child development. *Pediatrics*, 84(1), 157-164.

Prinz, R. J., Smith, E. P., Dumas, J. E., Laughlin, J. E., White, D. W., & Barrón, R. (2001). Recruitment and retention of participants in prevention trials involving family-based interventions. *American journal of preventive medicine*, 20(1), 31-37.

Programa Nacional Cuna Más. (2015). Retrieved May 10, 2015, from http://www.cunamas.gob.pe/?page_id=57.

Raikes, H., Green, B. L., Atwater, J., Kisker, E., Constantine, J., & Chazan-Cohen, R. (2006). Involvement in Early Head Start home visiting services: Demographic predictors and relations to child and parent outcomes. *Early Childhood Research Quarterly*, 21(1), 2-24.

Riley, S., Brady, A., Goldberg, J., Jacobs, F., & Easterbrooks, A. (2008). *Once the door closes: Understanding the parent-provider relationship*. *Children and Youth Services Review*, 30(5), 597-612.

Roggman, L. A., & Cook, G. A. (2010). Attachment, aggression, and family risk in a low-income sample. *Family Science*, 1(3), 191-204. doi:10.1080/19424620.2010.567829

Roggman, L. A., Boyce, L. K., & Cook, G. A. (2009). Keeping kids on track: Impacts of a parenting-focused Early Head Start program on attachment security and cognitive development. *Early Education and Development*, 20(6), 920-941.

Roggman, L. A., Boyce, L. K., Cook, G. A., & Jump, V. K. (2001). Inside home visits: A collaborative look at process and quality. *Early Childhood Research Quarterly*, 16, 53-71.

Roggman, L. A., Christiansen, K., Cook, G. A., Jump, V. K., Boyce, L. K., & Peterson, C. A. (2006). Home visits: Measuring how they work. Presentation in the Early Intervention Research Institute, Mini-Conference, Logan, UT.

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. A., Cook, G. A., Peterson, C. A., & Raikes, H. H. (2008). Who drops out of Early Head Start home visiting programs? *Early Education and Development*, 19(4), 574-599.

Rosero, J., & Oosterbeek, H. (2011). Trade-offs between different early childhood interventions: Evidence from Ecuador. Working paper 11-102/3, Timbergen Institute, Amsterdam, Holanda.

Sandner, M. (2013). Effects of early childhood intervention on child development and early skill formation: Evidence from a randomized controlled trial (No. 518). Discussion Paper, Wirtschaftswissenschaftliche Fakultät, Leibniz Universität Hannover.

Schneider, A., & Ramires, V. R. (2007). Primeira Infância Melhor: uma inovação em política pública. Brasília, Brazil: UNESCO.

Sierau, S., Brand, T., & Jungmann, T. (2012). Parental involvement in home visiting: Interpersonal predictors and correlates. *Infant Mental Health Journal*, 33(5), 489-495.

Smith, J. D., Dishion, T. J., Shaw, D. S., & Wilson, M. N. (2013). Indirect effects of fidelity to the family check-up on changes in parenting and early childhood problem behaviors. *Journal of consulting and clinical psychology*, 81(6), 962.

Stoltzfus, E., & Lynch, K. (2009). Home visitation for families with young children. R40705). Washington, DC: Congressional Research Service.

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

Tinajero, A. R. (2010). Scaling up early child development in Cuba. Washington, DC: Brookings Institution, Wolfensohn Center.

Twombly, L., Waddell, & Harrison. (2003). Supportive interactions with families: A self-rating scale. Unpublished manuscript, University of Oregon EIP.

UNESCO (2007). "Primera Infancia Melhor: Uma Inovacao En Politica Publica." Accessed 27 June 2015 from: <http://unesdoc.unesco.org/images/0015/001552/155250por.pdf>

Vogel, C. A., Caronongan, P., Thomas, J., Bandel, E., Xue, Y., Henke, J., ... & Murphy, L. (2015). Toddlers in Early Head Start: A Portrait of 2-Year-Olds, Their Families, and the Programs Serving Them (Executive Summary) (No. aeb78516d8c64039a9afdbece7d63290). Mathematica Policy Research.

Wallander, J. L., McClure, E., Biasini, F., Goudar, S. S., Pasha, O., Chomba, E., ... & Carlo, W. (2010). Brain research to ameliorate impaired neurodevelopment-home-based intervention trial (BRAIN-HIT). *BMC pediatrics*, 10(1), 27.

Walker, S. P., Chang, S. M., Powell, C. A., & Grantham-McGregor, S. M. (2005). Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: prospective cohort study. *The Lancet*, 366(9499), 1804-1807.

Walker, S. P., Wachs, T. D., Gardner, J. M., Lozoff, B., Wasserman, G. A., Pollitt, E., ... & International Child Development Steering Group. (2007). Child development: risk factors for adverse outcomes in developing countries. *The Lancet*, 369(9556), 145-157.

Wasik, B. A., Mattera, S. K., Lloyd, C. M., & Boller, K. (2013). Intervention dosage in early childhood care and education: It's complicated (OPRE Research Brief OPRE 2013-15). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, US Department of Health and Human Services.

Wasik, B. H., & Bryant, D. M. (2001). *Home visiting procedures for helping families* (2nd ed.). Thousand Oaks, CA: Sage.

Wasik, B.H., & Sparling, J.J. (1995). Home visit assessment instrument. Chapel Hill, NC: School of Education, University of North Carolina, Chapel Hill.

Yousafzai, A. K., Rasheed, M. A., Rizvi, A., Armstrong, R., & Bhutta, Z. A. (2014a). Effect of integrated responsive stimulation and nutrition interventions in the Lady Health Worker programme in Pakistan on child development, growth, and health outcomes: a cluster-randomised factorial effectiveness trial. *The Lancet*, 384(9950), 1282-1293.

Yousafzai, A. K., & Aboud, F. (2014b). Review of implementation processes for integrated nutrition and psychosocial stimulation interventions. *Annals of the New York Academy of Sciences*, 1308(1), 33-45.

Zaslow, M., Tout, K., Halle, T., Whittaker, J. V., & Lavelle, B. (2010). Toward the Identification of Features of Effective Professional Development for Early Childhood Educators. Literature Review. Office of Planning, Evaluation and Policy Development, US Department of Education.

Appendix

Table A1: Programs from HomVEE Review with Positive Impacts on Child Outcomes

Program	Reference	Age at evaluation	Effect size
Child FIRST	Lowell, D. I., Carter, A. S., Godoy, L., Paulicin, B., & Briggs-Gowan, M. J. (2011). A randomized controlled trial of Child FIRST: A comprehensive home-based intervention translating research into early childhood practice. <i>Child Development, 82</i> (1), 193-208.	12 months	Any ITSEA domain (proportion with clinically concerning problems) -.28 Child language (proportion with clinically concerning problems) -.88 ITSEA externalizing (proportion with clinically concerning problems) -.42
Child Parent Enrichment Project (CPEP)	Barth, R. P. (1991). An experimental evaluation of in-home child abuse prevention services. <i>Child Abuse and Neglect: The International Journal, 15</i> (4), 363-75. Additional Source: Barth, R. P., Hacking, S., & Ash, J. R. (1988). Preventing child abuse: An experimental evaluation of the Child Parent Enrichment Project. <i>Journal of Primary Prevention, 8</i> (4), 201-217.	6 months	Mood (ITQ subscale) -.39
Early Head Start Home Visiting	Jones Harden, B., Chazan-Cohen, R., Raikes, H., & Vogel, C. (2012). Early head start home visitation: The role of implementation in bolstering program benefits. <i>Journal of Community Psychology, 40</i> (4), 438-455.	Kindergarten entry	FACES positive approaches to learning .20 FACES social behavior problems -.15

Early Start	Fergusson, D. M., Horwood, L. J., Grant, H., & Ridder, E. M. (2005). Early start evaluation report. Christchurch, NZ: Early Start Project Ltd.	36 months	Total behavior score .36 Total internalizing score .26
Family Check-up	Shaw, D. S., Connell, A., Dishion, T. J., Wilson, M. N., & Gardner, F. (2009). Improvements in maternal depression as a mediator of intervention effects on early childhood problem behavior. <i>Development and Psychopathology</i> , 21, 417–439.	Age 3 and 4	CBCL Externalizing .23 CBCL Internalizing .21 Eyberg Problem Behavior .23
Healthy Families America (HFA)	Caldera, D., Burrell, L., Rodriguez, K., Crowne, S. S., Rohde, C., & Duggan, A. (2007). Impact of a statewide home visiting program on parenting and on child health and development. <i>Child Abuse & Neglect</i> , 31(8), 829–852. doi:10.1016/j.chiabu.2007.02.008	Age 2	BSID percentage within normal limits on cognitive .24 CBCL percentage with externalizing scores in normal range .19 CBCL percentage with internalizing scores in normal range .32
Healthy Families America (HFA)	Landsverk, J., Carrilio, T., Connelly, C. D., Ganger, W., Slymen, D., Newton, R., et al (2002). Healthy Families San Diego clinical trial: Technical report. San Diego, CA: The Stuart Foundation, California Wellness Foundation, State of California Department of Social Services: Office of Child Abuse Prevention.	Year 1 of program	BSID, Mental Development Index (MDI) .23 CBCL Somatic problems T score - .24
Home Instruction for Parents of Preschool Youngsters (HIPPY)	Necoechea, D. M. (2007). Children at risk for poor school readiness: The effect of an early intervention home visiting program on children and parents. Dissertation Abstracts International Section A: Humanities and Social Sciences, 68 (6-A), 2311. (Dissertation Abstract: 2007-99230-512)	At 16 weeks	Expressive One-Word Picture Vocabulary Test – Revised .34

Nurse-Family Partnership	Olds, D. L., Kitzman, H., Cole, R., Robinson, J., Sidora, K., Luckey, D. W., et al (2004). Effects of nurse home-visiting on maternal life course and child development: Age 6 follow-up results of a randomized trial. <i>Pediatrics</i> , 114(6), 1550–1559.	At 6-year follow up	CBCL (total problems) -0.37 KABC mental processing composite (arithmetic and reading) .18 PPVT-III receptive vocabulary 0.17
Nurse-Family Partnership	Olds, D. L., Robinson, J., O'Brien, R., Luckey, D. W., Pettitt, L. M., Henderson, C. R., et al (2002). Home visiting by paraprofessionals and by nurses: A randomized, controlled trial. <i>Pediatrics</i> , 110(3), 486.	21 months	PLS-3 (language delay) -0.45
Parents as Teachers	Drotar, D., Robinson, J., Jeavons, L., & Lester Kirchner, H. (2009). A randomized, controlled evaluation of early intervention: The Born to Learn curriculum. <i>Child: Care, Health & Development</i> , 35(5), 643–649.	At 36 months	Mastery Motivation – Task Competence 0.2
Parents as Teachers	Wagner, M., Clayton, S., Gerlach-Downie, S., & McElroy, M. (1999). An evaluation of the northern California Parents as Teachers demonstration. Menlo Park, CA: SRI International.	At 36 months	DP-II Self-Help Development Scale (mean months differential) 0.25
Play and Learning Strategies (PALS) Infant: Results for PALS Toddler (PALS II)	Landry, S. H., Smith, K. E., Swank, P. R., & Guttentag, C. (2008). A responsive parenting intervention: The optimal timing across early childhood for impacting maternal behaviors and child outcomes. <i>Developmental Psychology</i> , 44(5), 1335-1353.	At three months after program end date	PPVT-III receptive vocabulary 0.36 Cooperation 0.3 Social engagement 0.32 Use of words 0.37