ARTICLE

DEVELOPMENT AND INITIAL VALIDATION OF AN OUTCOME MEASURE FOR HOME VISITATION: THE HEALTHY FAMILIES PARENTING INVENTORY

JUDY KRYSIK AND CRAIG W. LECROY
Arizona State University School of Social Work

ABSTRACT: The Healthy Families Parenting Inventory (HFPI) is a 63-item outcome measure that was designed to examine change in nine parenting-related domains. The HFPI was developed to respond to the need for an outcome measure for home visitation programs that is relevant to the intervention, sensitive to change, and appropriate with a diverse participant base, and would produce data that are immediately useful in practice. The authors detail the steps in the development and initial validation of the HFPI. The pattern of inter-item and item-to-subscale correlations as well as an exploratory factor analysis and sensitivity to change analysis support the nine-factor model of the HFPI.

* * *

In 1990, a report from the U.S. Advisory Board on Child Abuse and Neglect called for immediate attention to the problem of child abuse. A separate report released the following year emphasized the need for prevention and suggested a voluntary program of universal home visitation as the key strategy to prevent child abuse (U.S Advisory Board on Child Abuse and Neglect, 1991). Some 20 years later, a national inventory of home visitation programs in the United States documented 119 home visiting program models in 46 states and the District of Columbia (The Pew Center on the States, 2011). Research conducted by Pew in 2010 reported annual spending on home visitation managed through the states at 1.4 billion, with the three most popular program models being Healthy Families America ($126.5 million), Parents as Teachers ($79.7 million), and Nurse–Family Partnership ($59.6 million) (The Pew Center on the States, 2011). Owing to initial evidence of program effectiveness, the federal Patient Protection and Affordable Care Act (2010) added an additional 1.5 billion to state budgets through the establishment of the Maternal, Infant, and Early Childhood Home Visiting Program (MIECHV; Paulsell, Avellar, Sama Martin, & Del Grosso, 2010). The Pew Report did not include home visiting programs for which federal funding is allocated directly to localities. Head Start and Early Head Start are by far the nation's largest and longest running programs in this category, followed in size by Healthy Start. Together, these two streams offunding mark substantial progress toward the goal of universal early home visitation.

The rapid expansion of home visitation and its associated cost has focused attention on the effectiveness of this strategy to achieve its goals. Although various models specify different program goals, the MIECHV legislation specified eight outcome domains. These eight domains included child outcomes such as health, development and school readiness, reduction in maltreatment and juvenile delinquency as well as parent outcomes including economic self sufficiency, linkages and referrals, maternal health, positive parenting practices, and reductions in family violence and crime (Paulsell et al., 2010).

In 2010, the U.S. Department of Health and Human Services sponsored a major review of the evidence on home visitation known as the *Home Visiting Evidence of Effectiveness (HomVEE) study.* The HomVEE reports on the eight outcome domains that were specified in the MIECHV legislation (Paulsell et al., 2010). Although the existing evidence was considered favorable in terms of expanding home visitation, it is perhaps best described as "mixed." The HomVEE staff conducted a thorough literature review on home visitation and prioritized for in-depth study the 11 program models with the most rigorous evidence; thus, it was focused on evidence of effectiveness rather than on amount of funding. The HomVEE study reviewed 162 experimental and quasi-experimental studies on the selected programs and reported the results across the eight outcome domains for each study (Paulsell et al., 2010). An examination of the outcome measures across the programs tended toward few favorable findings, many null findings, and few unfavorable or ambiguousfindings. For example, among the primary measures of

Direct correspondence to: Craig W. LeCroy, Arizona State University School of Social Work, Tucson Component, 340 N. Commerce Park Loop, Suite 250, Tucson, AZ 85745. e-mail: craig.lecroy@asu.edu.

INFANT MENTAL HEALTH JOURNAL, Vol. 33(5), 496–505 (2012)

DOI: 10.1002/imhj.21343

positive parenting practices reported in HomVEE, there were 16 favorable findings, 134 null findings, and 4 unfavorable findings (Paulsell et al., 2010).

The lack of consistent favorable outcomes reported in the HomVEE study occurred across program models and outcome domains, and were reported for single-item measures as well as scales and subscales common to home visitation. For instance, Wagner, Clayton, Gerlach-Downie, & McElroy (1999), in a study examining Parents as Teachers, reported results from the Knowledge of Infant Development Inventory (KIDI), the Home Observation for Measurement of the Environment (HOME) Inventory, and the Parenting Sense of Competence (PSOC) scale. They found no effect on 22 of 25 positive parenting practice measures and three unfavorable effects. In contrast, Olds et al. (2004), using the HOME as well as a measure of sensitive/responsive interaction, found no effect on two of two positive parenting outcomes for the Nurse Family Partnership program. Duggan et al. (2007), in a study of one Healthy Families America program, found no effect in the area of positive parenting practices through the use of a single item related to relinquishment of parental role. Sweet and Appelbaum's (2004) meta-analysis of home visiting programs for families with young children also found mixed results. In their analysis, three of the five child outcomes had average effect sizes significantly greater than zero, as did three of the five parent outcomes; however, some of the significant effect sizes were rather small.

In the past, when evaluations of home visitation have not achieved the anticipated results, two explanations typically have been advanced. One explanation is that the program was not implemented as planned. When Duggan et al. (2004), for example, found disappointing results in an experimental evaluation of home visitation, they concluded that "We believe that the program's implementation system contributed to its minimal impact on maltreatment" (p. 615). A second explanation has been that the program's theory or model was flawed. This explanation was suggested by Olds et al. (2002) in their comparison of nurse and paraprofessional home visitor services when they stated that ". . . .it is reasonable to ask whether paraprofessionals have legitimacy in the eyes of families during pregnancy and infancy" (p. 494). In light of the HomVEE study that reported on such a multitude of measures, a third plausible explanation in the face of a failed outcome study would seem to be that the measures were inadequate in detecting the change that occurred or that the wrong items were measured.

If a measure is not able to detect program-related changes, this might explain the unexpected or null findings. Recent recognition that there are limits to what is changeable leads to different criteria for selecting outcome measures than have typically been applied. In the past, clinical psychologists were driven by the assumption that most problems and disorders were treatable (Seligman, 2007). Instruments, therefore, have been developed without careful consideration as to how changeable various attitudes and behaviors might be. For example, questions related to a young child's mood may be more of a reflection on his or her personality than a reflection on positive parenting. What this suggests is that an outcome measure should be specifically designed to tap into behaviors and attitudes that are likely to change and should avoid items that focus on characteristics that are dictated by nature. Cultural bias also can lead to error in measurement. Many instruments were developed and validated on convenience samples that are very different from the populations served by home visitation programs. For instance, the Child Abuse Potential (CAP) Inventory, a common measure used in child abuse and neglect prevention programs, includes the item "I enjoy pets" (Milner, 1994). Household pets are not generally accepted in some cultures due to cleanliness-related concerns. Thus, it is likely that this item would contribute negatively toward the total score on all administrations when used with parents from such cultures (LeCroy & Krysik, 2010).

It is also important when selecting outcome measures to keep in mind the theory of the intervention and what the program is trying to change. For instance, the CAP Inventory item "I have an unlisted telephone number (Milner, 1994) is not likely to show change and would not be a focus of the intervention. The more items an outcome measure includes that relate to heritable traits, static conditions, aspects of living that are culturally based, and those that are tangential to the intervention, the less likely that the measure will yield data reflecting a positive program impact, even if there is one.

What the HomVEE review suggests is that for home visitation programs in general, there is a need to develop outcome measures that are tailored to the context of the intervention. In addition, beyond sensitivity to change, the data should provide information that is immediately useful to home visitors and supervisors in deciding how to help families. That is, the data collected should facilitate the identification of areas of need and concern as well as areas of strength, all of which can be used in case planning.

The remainder of this article outlines the process used to develop the Healthy Families Parenting Inventory (HFPI). The HFPI is a 63-item outcome measure that was developed specifically for use in assessing parent-focused outcomes in federal- and state-supported home visitation programs such as those profiled in the HomVEE study. The name of the measure reflects the desirable end goal of family-focused home visitation programs and emphasizes a strength-based intent over a deficit-focused one.

METHOD

The scale development processes outlined by Springer, Abell, and Hudson, 2002, DeVellis (2003), and Abell, Springer, and Kamata (2009) were used as guides to develop the HFPI. The first step in the scale development process was to clearly determine what should be measured. The next steps were item construction and determination of response format, followed by piloting for face and content validity, then data collection for testing empirical forms of validity and reliability, and finally, testing the sensitivity to change. IRB approval was obtained for the study.

Conceptualizing the Measure

The authors used a multifaceted approach to identify what an outcome measure for home visitation should be capturing. First, a

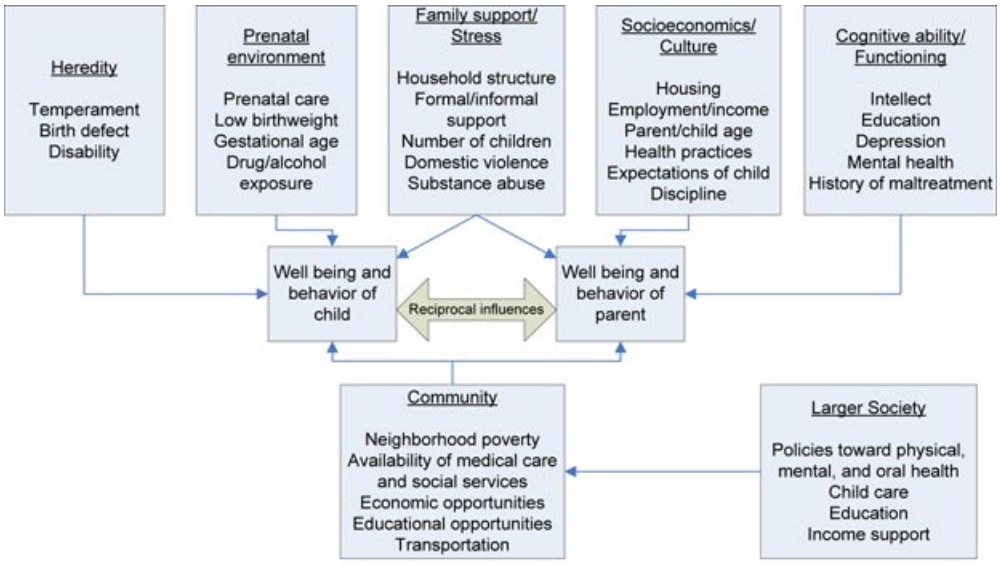

FIGURE 1. Ecological–transactional model of child maltreatment and healthy development.

literature review on the theory and research related to the goals of home visitation was conducted and is summarized later. The literature review was followed by an examination of the program logic model for one Healthy Families America program that served as the basis for later development of the national logic model and which was consistent with the direction of home visitation in general. Finally, an exploratory study of the outcomes of Healthy Families Arizona as perceived by the staff, supervisors, and participants of the program was conducted. An expert panel was formed and included two academics who specialized in the subject matter and in scale development, one clinician with experience in home visitation, and one Healthy Families supervisor. The expert panel provided feedback on the outcome domains and then contributed to the development of the overall set of items to assist with content and face validation, including examining the items for cultural specificity. For example, in examining cultural specificity, reviewers critiqued literacy level, how cultural traits might affect responses, the sequencing of questions, and the perceived intrusiveness of the questions.

Figure 1 was developed to graphically summarize the theory and research related to the early prevention of child maltreatment. In brief, theories of child maltreatment recognize that the root causes of early childhood maltreatment can be organized into a framework of four principal systems: (a) the individual parent and child, (b) the family, (c) the community, and (d) the larger societal system (Landy & Menna, 2006). Within each of these four systems, numerous factors have been found to increase a child's risk for maltreatment whereas other factors have been shown to protect against maltreatment. Researchers studying the etiology and effects of child maltreatment have argued for a simultaneous study of multiple risk and protective factors, suggesting that it is more than just one factor that makes certain segments of the population more likely to report child-abuse histories or experiences (Belsky, 1993; Brown, Cohen, Johnson, & Salzinger, 1998; Cicchetti, Toth, & Rogosch, 2000; Landy & Menna, 2006). In addition, the interaction between the various factors is considered transactional in nature (i.e., a series of actions and reactions). For instance, a prenatal experience characterized by risk is likely to negatively impact the well-being and behavior of the infant. In turn, the well- being and behavior of the infant will influence how the parent responds to the infant, and the parent's response will impact the continued well-being and behavior of the infant.

Studies have found that as the number of risk factors increase, a child's likelihood for abuse and neglect also increases (Brown et al., 1998). For instance, Brown et al. (1998) discovered that the prevalence of child abuse or neglect increased from 3% when no risk factors were present to 24% when four or more risk factors were present. Belsky and Isabella (1988) studied cumulative risks in predicting attachment security. Their results showed no single risk factor was as predictive as cumulative effects. More specifically, infants with zero, one, two, and three risk factors had a 92, 83, 38, and 17% respective probability of being securely attached. More recently, Appleyard, Egeland, van Dulmen, and Sroufe (2005) found a cumulative effect in a longitudinal study of at-risk urban children and showed that the number of risks in early childhood can predict problems in adolescence. These findings suggested that to effectively identify children who are at a greater risk for child maltreatment and developmental difficulties, a significant number of risk factors need to be considered and addressed. Given these findings, the expert panel needed to consider the multitude of personal, family, and environmental factors that strengthen families, reduce the risk of child abuse and neglect within families, and improve child outcomes such as health and development.

The Healthy Families Arizona program logic model was examined by the expert panel to investigate how the program theory

could be reflected in the development of the domains for the outcome measure. The logic model included 10 objectives, including increases in parental (a) support, (b) health behaviors, (c) problem-solving skills, (d) competence, and (e) positive interactions with the child. Also included in the logic model objectives were improvements in (f) maternal mental health; (g) human capital factors such as education, employment, and financial resources; (h) child health practices; (i) child development; and (j) positive discipline strategies. These 10 program objectives appeared consistent with the risk and resiliency understanding of child maltreatment represented in the literature and with the systems framework that focuses on the child, the parent, and the family's broader social and economic network.

To create a grounded outcome measure and understand how the program impacts program participants, a series of focus groups were conducted. The exploratory study included a series of four focus groups with 6 to 10 home visitors participating in each group. Individual interviews also were conducted with a mix of 25 current and former female Healthy Families Arizona participants who had varying lengths of involvement in the program. Notes from the focus groups and interviews were transcribed in a manner so that participants could not be identified.

The data-collection efforts for the focus groups and individual interviews were focused on describing the changes being sought or were made as a result of participation in the home visitation program. For example, in the home visitor focus groups, one of the questions asked: "If you had the opportunity to tell us what you think we should be measuring to capture the changes in participating families, what do you think we should be looking for?" In the participant interviews, one question was "What changes, if any, do you feel you have made as a result of your participation in the Healthy Families program?" Other questions examined aspects of the participants' experiences as parents. Illustrative examples of responses to the focus groups with home visitors and individual interviews with past and current participants are presented in Table 1.

Members of the expert panel reviewed the summary of the literature, the logic model, and the transcripts from the focus groups and individual interviews. The expert panel members met to discuss the data and agree on the themes that should be included in the measure. Expert panel review and discussion based on these three sources of data resulted in the identification of 10 themes or domains. The 10 domains were (a) obtaining social support, (b) using problem-solving skills, (c) satisfaction with the parental role, (d) quality of the home environment for child rearing, (e) increasing one's effectiveness as a parent, (f) improving parent/child interaction, (g) reducing maternal depression, (h) improving parental self-care, (i) mobilizing resources to meet need, and (j) feeling confident in one's parenting skills.

Item Construction and Format

The next step in the development of the measure was to develop an item pool. A major consideration in developing the item pool was whether to write items describing very specific behaviors or to rely on broader indicators of change. As noted by Sweet and Appelbaum (2004) in their meta-analytic review of 60 home visiting programs, families vary by family structure, age of the mother, ethnicity, and socioeconomic background, and typically have some combination of risk factors that can include low income, history of abuse and neglect, teen parent, domestic violence, substance abuse, and other problems related to mental health. Due to the diversity in the home visitation population, a decision was made to write broad rather than specific items. As items pertaining to each of the identified themes were developed, they were rewritten to conform to a Grade 5 reading level and then were examined for potential cultural bias related to age, gender roles, and ethnicity. Seventy-five items were developed by the expert panel and organized according to the 10 identified themes. As per the recommendation of DeVellis (2003), both positively and negatively worded items were included in the initial item pool as well as items that were worded differently, but involved redundant content. For instance, one item stated "I take care of my appearance" and was similar to a second, negatively worded item, "Life is so hectic that I can't find time to care for myself."

TABLE 1. *Parent and Home Visitor Responses to the Interview and Focus Group Questions*

Parent Responses
"I feel like I interact with my baby more, I feel more confident in how I interact. I pay more attention to what my baby needs."
"I get support from my home visitor that is what makes the biggest difference; I feel I need that support because what I am going through now is difficult."
"I just feel so much more capable, I'm more confident and I understand better how to care for my baby."
Home Visitor Responses
"You can see a difference in how they are taking care of themselves. In the past we might have seen the parent being attentive to the family but we didn't really see the parent attend to herself. So now you not only see them taking care of the child or children, but you see them also taking care of themselves, and realizing they need some mental health services."
"More caring and understanding about why the child is doing things. I saw this woman and her child was never dressed—either naked or in a diaper. I said, "You know what? If you start getting in a routine where she gets dressed every morning, then when she goes to school she'll have a much easier time." I would say to the child, "Show me how to put your shirt on." And finally the kid just started getting dressed, and the mother started understanding about the routine. . . ."
"I have one that way in the beginning she used drugs while she was pregnant, and she admitted all that to me and I connected her with the right resources where now she's drug free. But now, from the person that I saw when I first met her, she is so totally different now because she goes through that whole problem solving on her own."

The items for the HFPI were written as declarative statements rather than as questions (e.g., "I feel supported by others" rather than "How supported do you feel?"). In addition, a common set of response categories utilizing a 5-point Likert scale was used to ease the response burden on participants and facilitate variability in the responses. The response continuum selected for the HFPI reflects variations in frequency and includes "1 = *rarely or never*," "2 = *a little of the time*," "3 = *some of the time*," "4 = *good part of the time*," and "5 = *always or most of the time*."

The initial item pool was pilot-tested with a convenience sample of 100 parent volunteers not enrolled in home visitation. The initial piloting was used to test the items for clarity and ease of response. Piloting of the initial 75-item draft questionnaire with a convenience sample of 100 volunteer parents not enrolled in Healthy Families resulted in the revision of several items that were difficult to read or that created confusion as to what was being asked. This process led to the deletion of seven items. Demographic information was not collected from this sample.

Exploratory Factor Analysis

The next step in the process was to conduct an exploratory factor analysis (EFA) using the principal components extraction method with varimax rotation. The EFA was selected as the analytic strategy because the primary goal of the research was to determine the underlying factor structure and examine the maximum amount of variance and internal reliability of the instrument (Thompson, 2004). An a priori criterion loading of 0.30 was set for inclusion of items in the initial stage of item reduction, as per the recommendation of Waltz, Strickland, and Lenz (2004, p. 162). The results of the EFA and correlation analysis as well as the sensitivity to change analysis will be discussed later.

Data for the EFA came from the revised draft of the items that was administered to a convenience sample of 337 mothers of infants within their first 3 weeks of enrolling in Healthy Families Arizona. Families were only considered eligible for the Healthy Families program if the total family stress score met or exceeded the cutoff score of 25 for either parent, indicating risk for child abuse or neglect. Data from the 337 parents were entered in SPSS. Tinsley and Tinsley (1987) recommend a ratio of 5 to 10 participants per item up to about 300 participants for factor analysis; thus, the sample size was considered adequate. The characteristics of the sample were reflective of the overall population served by the Healthy Families Arizona program, and were marked by diversity, as are the participants of home visitation programs in general (Russell, Britner, & Woolard, 2007; Sweet & Appelbaum, 2004). For instance, the sample was entirely female and ethnically diverse, with 20.5% having a preference for services in Spanish and 79.5% with a preference for services in English. The women ranged in age from early teens to 44 years ($M = 23$ years, $SD = 5.97$). Sixty percent of participants had a gross annual household income less than $10,000. One half of the participants were first-time mothers, and the number of living children ranged from one to six. Most of the women were not employed at the time of birth (87.5%). Forty-one percent of the women had obtained at least a GED or high-school diploma, 12% were enrolled in school at the time of entry to the program, and 47% were not enrolled.

To evaluate the items, data from the negatively worded items were reverse-scored. Each item was then evaluated in terms of its mean score and variance. Item variance is valued in scale construction because if a majority of individuals similarly answer an item, the item will not discriminate among various levels of the construct being measured (DeVellis, 2003).

Examination of the response frequencies for each item showed acceptable variability that encompassed the entire response continuum from 1 through 5 for all items, except one that had responses ranging from 2 to 5. Ideally, the mean score on each item should be near the center of the range. If the mean is near either extreme of the response continuum, it suggests that the item is failing to detect the range of responses on the construct (DeVellis, 2003). The mean scores for the 67 items ranged from a low of 2.43 to a high of 4.53. No items were deleted as a result of the variability assessment.

Subscale scores were calculated by using a simple additive approach. That is, the responses to all of the items in each of the 10 thematic areas or subscales were summed. Each subscale was then subjected to an examination of inter-item and item-to-total subscale score correlations. Factorial validity was assessed by evaluating the way the data either conformed to or deviated from the following hypotheses: (a) The items of each subscale will correlate highest with the total score of the subscale formed by those items, (b) the items will be moderately correlated with one another, (c) the items will correlate to a lesser degree with the total scores of other subscales, and (d) the items will correlate poorly with variables or constructs related to the participants' backgrounds, such as income and age (Springer et al., 2002).

Sensitivity to Change Analysis

A useful measure is one that is good at detecting change resulting from the intervention for the population under study over time. Sensitivity to change can be considered in terms of the proportion of shifting items (i.e., those items that demonstrate improvement) to nonshifting items (i.e., those that show no change) (Hyland, 2003). To examine sensitivity to change, each item in the subscale was examined for significant change using paired *t* tests from entry to home visitation to 6 months and then to 12 months. The effect size or magnitude of change of each subscale was calculated using Cohen's (1988) *d*. Data were from the Healthy Families Arizona program for 2010 and included 408 participants with 6-month data and 213 participants with 12-month data.

RESULTS

The factor loadings for the 10-factor model were examined through an exploratory factor analysis using SPSS. Two items with factor loadings less than .30 were deleted. The two deleted items were "I do not know who I can turn to for help in making important decisions as a parent," and "Parenting comes naturally to me."

TABLE 2. *Factor Loadings and Subscale αs for the Nine-Factor Model*

Subscale (Cronbach's α)	Item	Factor Loading
Social Support (.84)	I feel supported by others.	.71
	I feel that others care about me.	.74
	I discuss my feelings with someone.	.54
	If I have trouble, I feel there is always someone I can turn to for help.	.85
	I have family or friends who I can turn to for help.	.80
Problem Solving (.92)	I learn new ways of doing things from solving problems.	.53
	I deal with setbacks without getting discouraged.	.69
	When I have a problem, I take steps to solve it.	.56
	When I am faced with a problem, I can think of several solutions.	.47
	I am good at dealing with unexpected problems.	.65
	I remain calm when new problems come up.	.75
Depression (.79)	I feel sad.	.50
	I feel positive about myself.	.68
	The future looks positive for me.	.72
	I feel unhappy about everything.	.68
	I feel hopeless about the future.	.70
	There isn't much happiness in my life.	.48
	I have so many problems I feel overwhelmed by them.	.51
	It is hard for me to get in a good mood.	.64
	My life is fulfilling and meaningful.	.53
Personal Care (.76)	I find ways to care for myself.	.54
	I take care of my appearance.	.57
	I get enough sleep.	.75
	I am a better parent because I take care of myself.	.79
	I take time for myself.	.58
Mobilizing Resources (.86)	I know where to find resources for my family.	.76
	I know where to find important medical information.	.70
	I can get help from the community if I need it.	.80
	I am comfortable in finding the help I need.	.67
	I know community agencies I can go to for help.	.76
	It is hard for me to ask for help from others.	.18[a]
Role Satisfaction (.76)	Because I'm a parent, I've had to give up much of my life.	.57
	I feel trapped by all the things I have to do for my child.	.69
	I feel drained dealing with my child.	.48
	There are times my child gets on my nerves.	.48
	I feel controlled by all the things I have to do as a parent.	.59
	I feel frustrated because my whole life seems to revolve around my child.	.30
Parent/Child Interaction (.77)	I have a hard time managing my child.	.67
	I can be patient with my child.	.67
	I respond quickly to my child's needs.	.60
	I do activities that help my child grow and develop.	.56
	When my child is upset, I'm not sure what to do.	.49
	I use positive words to encourage my child.	.46
	I can tell what my child wants.	.41
	I am able to increase my child's good behavior.	.37[a]
	I remain calm when my child is upset.	.61
	I praise my child everyday.	.55
Home Environment (.76)	My child has favorite things to comfort him/her.	.55
	I read to my child.	.39[a]
	I plan and do a variety of activities with my child every day.	.60
	I have made my home exciting and fun for my child.	.71
	I have organized my home for raising a child.	.58
	I check my home for safety.	.50
	My child has a schedule for eating and sleeping in my home.	.30[a]
	I set limits for my child consistently.	.26[a]

(Continued)

TABLE 2. *Continued*

	I make plans for our family to do things together.	.57
	I set rules for behavior in my home.	.45
Parenting Efficacy (.87)	I feel I'm doing an excellent job as a parent.	.81
	I am proud of myself as a parent.	.83
	I am more effective than most parents.	.72
	I have set goals about how I want to raise my child.	.58
	I am a good example to other parents.	.78
	I learn new parenting skills and use them with my child.	.60

[a] The item was revised as presented; however, the factor loading is for the original item.

The factor loadings for these two deleted items were .11 and .19, respectively. Two items, "I set limits for my child consistently (everyday)" and "It is hard for me to ask for help from an agency," also loaded below the .30 criterion, with factor loadings of .26 and .18, respectively. These items were revised and retained because they were judged by the expert panel members to be important to the intervention. The revised items read "I set limits for my child consistently," and "It is hard for me to ask for help from others." Three items that had factor loadings above the criterion, but less than .40, also were revised. For instance, "I know ways to increase my child's good behaviors" had a factor loading of.37 and was revised to "I am able to increase my child's good behavior" because the act was considered more consistent with the Parent/Child Interaction subscale than with the original item which related to knowledge. The item "I read to my child a lot" had a factor loading of .39 and was revised to "I read to my child." It is likely that the qualifier "a lot" was difficult to interpret given that the response categories asked for an estimate of frequency. The item "I have a schedule for eating, sleeping, and playing in my home" loaded at .30. The concept of play was deleted from this item because of the lack of perceived relevance for newborns. The factor loadings associated with the remaining items were all above 0.40 and are presented in Table 2. The extent to which these changes will result in improvements to the psychometrics will be determined through subsequent analysis.

Correlation Analysis

The pattern of item-to-item correlations within subscales and item-to-total subscale score correlations were generally as predicted. Based on the pattern of correlations, however, one of the 10 subscales was deleted. The parental competence subscale was highly correlated with three subscales: Parent/Child Interaction ($r = .84$), Home Environment ($r = .90$), and Parental Efficacy ($r = .86$). The high subscale correlations were due to three of the six items on the Parental Competence subscale correlating moderately with the total subscale score for Parent/Child Interaction and the remaining three items correlating moderately with the total subscale score for Home Environment. These six items were moved to the relevant subscales. Also based on the results of the correlation analysis, the item "I have so many problems I feel overwhelmed by them" was moved from the Problem Solving subscale to the Depression subscale to reduce the correlation between these two subscales and also to improve the face validity of the subscale. Feeling overwhelmed by problems does not appear to relate to the activity of problem solving.

Estimates of internal consistency calculated using Cronbach's α are presented in Table 2 for the nine remaining subscales. The subscale α coefficients ranged from a low of .76 to a high of .92, demonstrating good to excellent internal consistency. The subscale αs are expected to improve with the revised items.

Sensitivity to Change Analysis

An important feature in outcome measurement is the detection of important changes in areas that are meaningful to the program, such as improvements in parent/child interaction and mobilizing resources. The second column of Table 3 shows the percentage of items in each subscale that showed statistically significant gains from baseline to 6 months, and the fourth column shows the percentage of items with significant change from baseline to 12 months. The third and fifth columns in Table 3 report Cohen's d effect sizes to represent the magnitude of the change. Two of the nine subscales consistently performed the best over time: Both the Mobilizing Resources subscale and the Home Environment subscale had a high proportion of shifting items and medium effect sizes. The most problematic subscales were Social Support, Depression, and Personal Care. No more than 20% of the subscale items shifted significantly from baseline at either time point, and the effect sizes remained small. At least one third of the total scale items changed significantly over the first year of the intervention, and the overall effect size was in the medium range.

DISCUSSION

The purpose of this study was to report the psychometric properties of a new instrument designed to measure outcomes in home visitation programs. As Springer et al. (2002) noted, "More than ever, practitioners must be able to answer the questions about their work with clients . . . and have the tools to quickly and accurately assess and subsequently monitor change in client functioning" (p. 409). They suggested that instrument development is in its infancy and that more instruments need to be developed for the field.

This study began with the recognition that there was a need for more measures tailored specifically to the growing field of

TABLE 3. *Sensitivity to Change Analysis*

Subscale	Proportion of items shifting significantly from entry to 6 months	Standardized effect size entry to 6 months ($n = 408$)	Proportion of shifting items from entry to 1 year	Standardized effect size entry to 1 year ($n = 213$)
Social Support	20.0%	.05	0.0%	.05
Problem Solving	33.3%	.27	16.7%	.27
Depression	22.2%	.14	0.0%	.20
Personal Care	20.0%	−.14	20.0%	−.04
Mobilizing Resources	100%	.28	83.3%	.41
Commitment to Parental Role	0.0%	.10	50.0%	.15
Parent/Child Behavior	50.0%	.17	20.0%	.16
Home Environment	60.5%	.33	60.0%	.49
Parenting Efficacy	50.0%	.14	33.3%	.16
Total Score	41.3%	.23	31.2%	.30

Note. Effect sizes are Cohen's *d* and are generally interpreted as small .2, medium .5, and large .8 (Cohen, 1988).

home visitation for families with young children. Further, outcome instruments for home visitation need to be based on what is realistically changeable. Some measures tend to include items that are static over time and do not appear relevant to the intervention. As the population of families participating in home visitation tends to be diverse, a useful measure also should avoid items that may be based on culture, age, or gender. As a result, a multistep instrument-development process was embarked on that included an exploratory study, the use of an expert panel, a pilot process, exploratory factor analysis, and an analysis of sensitivity to change through the examination of longitudinal data.

The nine-factor model of healthy parenting in the HFPI is supported by the exploratory factor analysis and correlation analysis. Measures of internal consistency for the subscales were assessed as "good" to "very good," and are expected to improve with the adjustments that have occurred as a result of the initial empirical validation work reported in this study. The initial examination of sensitivity to change also is promising, with four of the nine subscales and the total scale score demonstrating small to moderate effect sizes over 1 year. Some subscales were not as sensitive to change as were others, which may be due to the program's weaker impact in these areas or the items being less able to detect change. However, the program might have limited impact on depression since home visitors are not directly trained to offer psychotherapeutic treatment that could impact depression, particularly in this sample with high levels of family stress. Further, social support is likely to show less change because in the early stages of parenting, infant families are typically more insular and also may be more neglecting of their personal care.

The resulting version of the HFPI (LeCroy, Krysik, & Milligan, 2007), is a 63-item instrument that measures change in nine parenting domains.[1] The nine domains are consistent with the risk and resiliency understanding of child maltreatment, and also reflect a systems framework. For instance, five subscales (Problem Solving, Depression, Personal Care, Role Satisfaction, and Parenting Efficacy) are closely aligned with the individual level (i.e., the parent). Two subscales, the Home Environment and Parent/Child Interaction, target the family level. The Social Support subscale is aligned with the community level whereas the Mobilizing Resources subscale deals with both the community level and the broader societal level. The majority of the HFPI items are worded positively and fit nicely with the strengths perspective of most home visitation programs. To be useful to the home visitor, certain items are targeted to major risk factors for child maltreatment, such as the items related to maternal depression and home safety. The subscales are independently useful because they fit well with the objectives of home visitation programs and are important outcomes as well as clinical indicators of program success.

During the home visit, responses to items that indicate problem levels should be the subject of further dialogue and may generate outside referrals for additional services. For instance, items related to parenting skills and those that promote child development such as reading can be particularly useful for case planning. If used as a practice tool, the HFPI should help the home visitor spend less time on areas where the parent already is performing well, enabling him or her to hone in on areas of need and concern. Items that reflect a high level of parenting competence and items that reflect positive growth when used over time provide areas for genuine positive reinforcement. At the program level, areas that do not reflect significant change over time can be identified for targeted home visitor training and resource development.

Note that the HFPI does not directly ask questions about discipline. When used as a practice tool, the HFPI could provide an opening into the discussion of sensitive topics such as corporal punishment. The HFPI also could be used in combination with other tools that directly screen for substance abuse, family violence, and mental health problems—three major risk factors for child maltreatment.

[1]The complete measure and scoring instructions are available from the authors; contact craig.lecroy@asu.edu to obtain this information. The instrument is available free of charge, and any questions can be directed to the authors.

Conclusion

The HFPI can be used for research and evaluation and as a practice tool in a variety of home visitation programs targeted at early childhood. It is easy to use because of its brevity and other design features such as a basic reading level. Another positive feature of the HFPI is that it is freely available for use. The entire 63 items can be used or it can be used in part by administering only those subscales of particular interest. Further examination of a broader array of parents, including those who do not meet the eligibility criteria for home visitation, could lead to a greater understanding regarding what is normative in parenting over the first year. Such an examination also could facilitate the examination of differential service use and problem areas across racial and ethnic lines as well as by income, family structure, and parental age.

Although only the English version was tested in this study, the HFPI has since been translated into Spanish using the back-translation method, and empirical validation of the Spanish version is under way. Further efforts to validate the HFPI should include a confirmatory factor analysis using a more geographically diverse sample. With a larger and more geographically diverse sample, the examination of cultural sensitivity also should be pursued. In sum, the HFPI is a response to the need for greater means to establish the accountability of home visitation programs that are currently the most promising primary prevention strategy to reduce child maltreatment and to promote the health and healthy development of children.

REFERENCES

Abell, N., Springer, D.W., & Kamata, A. (2009). Developing and validating rapid assessment instruments. New York: Oxford University Press.

Appleyard, K., Egeland, B., van Dulmen, M.H.M., & Sroufe, L.A. (2005). When more is not better: The role of cumulative risk in child behavior outcomes. Journal of Child Psychology and Psychiatry, 46, 235–245.

Belsky, J. (1993). Etiology of child maltreatment: A developmental–ecological analysis. Psychological Bulletin, 114, 413–434.

Belsky, J., & Isabella, R.A. (1988). Maternal, infant, and sociocontexual determinates of attachment security. In J. Belsky & A. Nezworski (Eds.), Clinical implications of attachment (pp. 41–94). Mahwah, NJ: Erlbaum.

Brown, J., Cohen, P., Johnson, J.G., & Salzinger, S. (1998). A longitudinal analysis of risk factors for child maltreatment: Findings of a 17-year prospective study of officially recorded and self-reported child abuse and neglect. Child Abuse Neglect, 22, 1065–1078.

Cicchetti, D., Toth, S.L., & Rogosch, F.A. (2000). The development of psychological wellness in maltreated children. In D. Cicchetti, J. Rappaport, I. Sandler, & R.P. Weissberg (Eds.), The promotion of wellness in children and adolescents (pp. 394–426). Washington, DC: Child Welfare League of America Press.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

DeVellis, R. (2003). Scale development: Theory and applications (2nd ed.). Applied Social Research Methods Series (Vol. 26). Thousand Oaks, CA: Sage.

Duggan, A., Fuddy, L., Burrell, L., Higman, S.M., McFarlane, E., Windham, A., & Sia, C. (2004). Randomized trial of a statewide home visiting program to prevent child abuse: Impact in reducing parental risk factors. Child Abuse & Neglect, 28, 623–643.

Duggan, A., Caldera, D., Rodriguez, K., Burrell, L., Rohde, C., & Crowne, S.S. (2007). Impact of a statewide home visiting program to prevent child abuse. Child Abuse & Neglect, 31, 829–852.

Hyland, M.E. (2003). A brief guide to the selection of quality of life instrument. Health and Quality of Life Outcomes, 1, 1–5.

Landy, S., & Menna, R. (2006). Early intervention with multi-risk families. Baltimore: Brookes.

LeCroy, C.W., & Krysik, J. (2010). Measurement issues in home visitation: A research note. Children and Youth Services Review, 32, 1483–1486.

LeCroy, C.W., Krysik, J., & Milligan, K. (2007). The Healthy Families Parenting Inventory. In J. Fisher & K. Corcoran (Eds.), Measures for clinical practice and research: A sourcebook (4th ed., Vol. 1, pp. 344–347). New York: Oxford University Press.

Milner, J.S. (1994). Assessing physical child abuse risk: The child abuse potential inventory. Clinical Psychology Review, 14, 547–583.

Olds, D.L., Robinson, J., O'Brien, R., Luckey, D.W., Pettitt, L.M., Henderson, C.R., Jr. et al. (2002). Home visiting by paraprofessionals and by nurses: A randomized, controlled trial. Pediatrics, 110, 486–496.

Olds, D.L., Robinson, J., Pettitt, L., Luckey, D.W., Holmberg, J., Ng, R.K. et al. (2004). Effects of home visits by paraprofessionals and by nurses: Age 4 follow-up results of a randomized trial. Pediatrics, 114, 1560–1568.

Patient Protection and Affordable Care Act (2010). Additional information on the efect of the Patient Protection and Affordable Care Act on the Hospital Insurance Trust Fund. Congressional Budget Office. Retrieved 2010-03-31.

Paulsell, D., Avellar, S., Sama Martin, E., & Del Grosso, P. (2010). Home visiting evidence of effectiveness review: Executive Summary. U.S. Department of Health and Human Services, Office of Planning, Research and Evaluation, Administration for Children and Families. Washington, DC: Government Printing Office.

The Pew Center on the States. (2011). Pew home visiting inventory: National overview. Retrieved from http://www.pewstates.org/research/reports/states-and-the-new-federal-home-visiting-initiative-85899377168.

Russell, B.S., Britner, P.A., & Woolard, J.L. (2007). The promise of primary prevention home visiting programs: A review of potential outcomes. Journal of Prevention and Intervention in the Community, 34, 129–147.

Seligman, M.P. (2007). What you can change and what you can't. New York: Vintage.

Springer, D.W., Abell, N., & Hudson, W.W. (2002). Creating and validating rapid assessment instruments for practice and research: Part 1. Research on Social Work Practice, 12, 408–439.

Sweet, M.A., & Appelbaum, M.I. (2004). Is home visiting an effective strategy? A meta-analytic review of home visiting programs for families with young children. Child Development, 75, 1435–1456.

Thompson, B. (2004). Exploratory and confirmatory factor analysis: Understanding concepts and applications. Washington, DC: American Psychological Association.

Tinsley, H.E.A., & Tinsley, D.J. (1987). Uses of factor analysis in counseling psychology research. Journal of Counseling Psychology, 34, 414–424.

U.S. Advisory Board on Child Abuse and Neglect. (1990). Child abuse and neglect: Critical first steps in response to a national emergency. Washington, DC: U.S. Government Printing Office.

U.S. Advisory Board on Child Abuse and Neglect. (1991). Creating caring communities: Blueprint for an effective federal policy for child abuse and neglect. Washington, DC: U.S. Government Printing Office.

Wagner, M., Clayton, S., Gerlach-Downie, S., & McElroy, M. (1999). An evaluation of Parents as Teachers demonstration. Menlo Park, CA: SRI International.

Waltz, C.F., Strickland, O., & Lenz, E.R. (2004). Measurement in nursing health research (3rd ed.). New York: Springer.