

FINAL REPORT:

A Randomized Trial of Healthy Families New York (HFNY):

Does Home Visiting Prevent Child Maltreatment?

NIJ Grant 2006-MU-MU-0002

October 31, 2010

Prepared by:

Kimberly DuMont
Kristen Kirkland
Susan Mitchell-Herzfeld
Susan Ehrhard-Dietzel

New York State Office of Children and Family Services

and

Monica L. Rodriguez
Eunju Lee
China Layne
Rose Greene

University at Albany, State University of New York

|

Acknowledgements

This research was supported by Award No: 2006-MU-MU-0002 awarded by the National Institute of Justice, U.S. Department of Justice. We thank the New York State Division of Temporary Disability and Assistance (NYS OTDA) for providing food stamps and public assistance data, the New York State Department of Health (NYS DOH) for providing additional information on the target child's birth weight, the New York State Office of Children and Family Services HFNY program coordinators, Joy Griffith and Bernadette Johnson, for providing access to HFNY's archived contracts and documents, and the Center for Human Services Research (CHSR) for extracting data pertaining to Healthy Families New York (HFNY) assessments, referrals, and home visits. We also acknowledge the assistance of and ongoing support from the management staff in the Erie, Rensselaer, and Ulster HFNY programs, whose support made the evaluation and follow-up data collection efforts possible.

We are especially grateful to Dorothy Baum for her competence and commitment in maintaining and organizing the interview data; Jeff Luks for his assistance in providing and explaining the data obtained from the HFNY Management Information System (MIS); and Ann-Margret Foley for her tireless and skilled supervision of the research interviewers. We also acknowledge the invaluable contributions of the interviewers, Maria Aviles, Gayle Bryant, Jenny Ciccone, Amber O'Dierno, Carol Erceg, Andrea Kowaleski, Veronica Salvas, and Arlene Tucker, whose perseverance helped us to locate and conduct interviews with over 900 families.

The opinions, findings, methods of analysis, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Institute of Justice.

TABLE OF CONTENTS

Executive Summary	3
Chapter 1: Project Description	1
Chapter 2: Methodology	22
Chapter 3: Description of Study Samples and Participants	34
Chapter 4: To What Extent is the HFNY Program Consistent with the HFA Model?	40
Chapter 5: Does HFNY Prevent or Reduce Child Maltreatment?	55
Chapter 6: Does HFNY Limit Precursors to Delinquency?	77
Chapter 7: Do the Benefits of HFNY Outweigh Its Costs?	88
Chapter 8: Conclusions and Implications	115
References	122

**A Randomized Trial of Healthy Families New York (HFNY):
Does it Prevent Child Maltreatment?
(NIJ Grant 2006-MU-MU-0002)**

EXECUTIVE SUMMARY

I. PROJECT OVERVIEW

The current study utilizes a seven-year longitudinal randomized controlled trial (RCT) to evaluate the effectiveness of a state-administered home visitation program in preventing child maltreatment and risks for delinquency. Healthy Families New York (HFNY), which is based on the Healthy Families America (HFA) model, was established as a strengths-based, intensive home visitation program with the explicit goals of 1) promoting positive parenting skills and parent-child interaction; 2) preventing child abuse and neglect; 3) supporting optimal prenatal care, and child health and development; and 4) improving parent's self-sufficiency.

In 2000, the New York State Office of Children and Family Services (OCFS) Bureau of Evaluation and Research, in partnership with the Center for Human Services Research at the University at Albany, initiated a RCT at three sites with the HFNY home visiting program. Families eligible for HFNY at each site were randomly assigned to either an intervention group that was offered HFNY services or to a control group that was given information on and referral to appropriate services other than home visiting. Baseline interviews were conducted with 1173 of the eligible women (intervention, n=579; control, n=594), and follow-up interviews at Years 1 and 2 achieved high rates of retention. In the third year, a reduced sample was assessed (n=522).

In 2006, additional funding from the National Institute of Justice (NIJ) was obtained to conduct a Year 7 follow-up in order to address the following four questions:

- 1) To what extent is the home visiting process of HFNY consistent with the HFA model?
- 2) Does home visiting effectively prevent or reduce child maltreatment?

- 3) Does home visiting limit the emergence of precursors to delinquency? and
- 4) Do the long-term benefits of an HFA-based home visiting program outweigh its costs?

Results will inform national child welfare policy about the effectiveness, costs, long-term benefits, and mechanisms through which a home visiting program achieves its impacts.

II. BACKGROUND & RESEARCH HYPOTHESES

Below, we present each of the study's goals, its method, and related literature.

To what extent is the home visiting process of HFNY consistent with the HFA model?

Researchers in the field of home visiting have called for increased understanding of the home visiting process (Gomby, 2007; Korfmacher, Kitzman & Olds, 1998; Olds, Sandler & Kitzman, 2007). Accordingly, several researchers have conducted case record reviews and fidelity assessments to better understand the mechanisms at work during a home visit (c.f., Harding, Reid, Oshana & Holton, 2004; Kessler, Nixon & Nelson, 2008). While some practices appear fairly constant across the programs (e.g., length of visits, variety of content discussed, and focus on parent-child interaction activities), other qualities of the model are context-dependent (e.g., characteristics of families served, service intensity, and service length).

- It is important to understand how HFNY's implementation of the HFA model may facilitate success in some areas and/or present obstacles to its effectiveness in others.

Does HFNY effectively reduce child maltreatment?

Evaluations of home visiting reveal that program effects on official reports of abuse and neglect during the early years of life have been scant (Duggan et al., 2004; Duggan et al., 2007; DuMont, Mitchell-Herzfeld et al., 2008; Olds, Henderson & Kitzman, 1994), but may emerge later in life (Olds, 1997). In contrast, results from several randomized controlled trials suggest that in the first one to three years of life, HFA-based programs have more favorable effects on

less severe forms of negative parenting, such as minor physical and psychological aggression (Duggan et al., 2004; Duggan et al., 2007; Landsverk et al., 2002; Mitchell-Herzfeld, Izzo, Greene, Lee & Lowenfels, 2005; DuMont, Mitchell-Herzfeld et al., 2008).

- The current study provides information regarding effectiveness of the program in preventing child maltreatment. It also assesses whether the early effects sustain and whether new effects emerge as the target children develop.
- The study also evaluates how who is offered home visitation services may affect program impacts on child maltreatment. Of particular interest are two policy-relevant subgroups: (1) the High Prevention Opportunity (HPO) subgroup, which consists of young, first-time mothers who initiate home visiting services prenatally; and (2) the Recurrence Reduction Opportunity (RRO) subgroup, which includes women who have had at least one substantiated child protective services report (as a non-victim) prior to random assignment (RA). Although the sample sizes for these two subgroups are limited, to the extent possible, we evaluate the effects of the program on their outcomes and, where appropriate, examine the potential mechanisms through which HFNY achieves its effects.

Does HFNY limit the emergence of precursors to delinquency?

A third major goal of the study is to examine HFNY's ability to prevent or limit behaviors and characteristics that are frequently associated with delinquency. These behaviors may represent risks for future transgressions (Broidy et al., 2003) or reveal skills and strategies that can play a protective role in the child's life (Reynolds, 2004). Despite results from several studies the potential for home visiting to effect children's early and later functioning remains unclear.

- We expected HFNY to positively impact children by promoting individual competencies, and encouraging home environments that recognize and reinforce strengths and achievements. These advantages may position children on a trajectory to avoid deviant and delinquent behaviors in adolescence. We hypothesized that at age 7, target children randomly assigned to the HFNY group would present with fewer problem behaviors, cognitive difficulties, and socio-emotional difficulties.

Do the long-term benefits of HFNY outweigh its costs?

Prior research has shown that the benefits of early childhood interventions exceed the costs of such programs (Karoly, Kilburn & Cannon, 2005). Home visitation programs in particular have been widely promoted as an efficient use of resources. Unfortunately, few evaluations of home visiting programs have included an economic component to support this level of confidence. In the current study, we conduct the first cost benefit analysis of the HFNY program, and adopt the perspective of the government to answer the following questions:

- What are the costs associated with the program?
- Does HFNY reduce spending for government-supported programs?
- Does HFNY increase tax revenues?
- Do the benefits of HFNY exceed the costs?
- Do the specific characteristics and/or experiences of HFNY participants influence the costs and benefits related to the program?

In combination, findings from the four study areas will enhance the field's understanding of whether, how, for whom, and at what cost a paraprofessional home visitation program effectively serves women at high risk for maltreating their children. Given the recent passage of the Maternal, Infant, and Early Childhood Visitation program, the results are both relevant and

timely, and may ultimately help home visiting play a more meaningful role in the lives of vulnerable families.

III. METHODOLOGY

Sample Selection

Recruitment for the RCT was conducted between March 2000 and August 2001. Randomization was conducted by way of a computer program. Women were selected for the study following the same screening and assessment procedures used to determine eligibility for HFNY. Family Assessment Workers (FAW) obtained informed consent from women prior to the administration of a well-established risk assessment tool. Following the sample selection period, 1173 eligible mothers completed baseline interviews (intervention, n=579; control, n=594).

Sample Description

About a third (34%) of the mothers in the study sample were White, non-Latina; 45% were African-American, non-Latina; and 18% were Latina. Like HFNY participants statewide, women in the study sample tended to lack a high school diploma or GED (47%), be young (31% under 19), or first-time mothers (55%). Women also had a high number of risk factors for child abuse and neglect, with an average of nearly 6 of 10 risks assessed as moderate to severe.

Data Sources

Data sources for the current study include both administrative databases and interviews with study respondents and their target children.

Baseline Covariates. Information gathered at baseline during interviews with the study mothers is used to identify appropriate covariates. These include dichotomous variables coded to represent the mother's race/ethnicity, the mother's age, the presence of a regular partner, at least

one move in the past year, being randomly assigned to the study prenatally or postnatally, receipt of at least a GED or high school diploma. The target child's gender and age were also assessed.

We used information collected on household composition and the number of prior pregnancies and births to create a variable that describes the total number of other biological children the respondent reported as of the baseline interview. In addition, we selected a series of measures to assess mothers' level of depressive symptoms, sense of mastery, and total parenting attitudes at baseline.

Mother's Earned Income. Mother's earned income was calculated as a sum of wages of various jobs that the respondent worked during the study period. In the baseline survey, wage data were collected on one job, but for survey years 1, 2, and 7, wage data were collected for up to five jobs within the period of time since the last interview.

HFNY Management Information System (MIS). We obtained MIS information on the characteristics and needs of the families served, the initial risk assessment, the frequency and content of home visits, the nature and outcome of service referrals, and worker demographic and training information for the 579 families who were randomly assigned to the HFNY arm of the study. We also created a count variable using items from the initial risk assessment to describe the level of moderate to severe risks present in families' lives as of random assignment.

NYS Administrative Databases. We used several NYS administrative databases to obtain data for the current study. We conducted searches of CONNECTIONS, the NYS Statewide Automated Child Welfare Information System, to determine whether respondents or their target children were ever the confirmed subject or confirmed victim in an indicated NYS Child Protective Services (CPS) investigation. Additionally, we performed a computerized search of the NYS Child Care Review Service (CCRS) and extracted information on service assessment

dates, service assessment choice, and start and end dates for each foster care placement for target children. Public assistance and food stamps data were obtained from the NYS Office of Temporary and Disability Assistance (OTDA), which provided the following benefit information: service eligibility, payment amount, date payment was issued, program type (e.g., food stamps, public assistance, etc.), and client identification information. Using the information available, we created a dummy variable to indicate public assistance eligibility at the time of random assignment. We also obtained birth weight data from the NYS Department of Health (DOH) for a subset of respondents who were randomly assigned to the study prenatally in order to corroborate the respondent's report of the target child's birth weight.

Year 7 Interviews. Respondents were re-interviewed at Year 7 if: (1) both the respondent and the target child were still living and (2) women in the control condition had not received the intervention at any time between random assignment and two weeks prior to the Year 7 assessment. Field staff completed 942 interviews with the original study participants. Interviews included information about parenting, the child, earnings, and household composition.

For the first time, we also conducted interviews with target children. Target children had to satisfy both the criteria established for the maternal interviews, had to live within driving distance of an interviewer in order to facilitate a face-to-face assessment, and had to be under the care and custody of the study respondent who could grant consent for the interview. Interviewers completed face-to-face assessments with 800 children. The target child interviews were designed to assess children's receptive vocabulary skills, socio-emotional health, self-regulatory abilities, and problem behaviors.

Analysis Plan

Prior to analyzing outcomes, we used Student's *t*-tests and Chi-square tests to assess the comparability of the intervention and control groups on selected baseline demographic and risk characteristics. This was repeated for the baseline sample, the Year 7 mother sample, and the Year 7 child sample. We also tested the representativeness of mothers who completed the follow-up interviews relative to those not included in the Year 7 assessment.

For analyses involving tests of effectiveness, all study respondents who had data were included in the analyses, regardless of their participation in the program. The dependent variables were analyzed using generalized linear models, applying the most appropriate distribution and link function.

In all tests of the program's effectiveness, the treatment condition was the primary independent variable, with the control condition (0) serving as the reference group. Covariates were included as necessary to maximize the equivalence of the two treatment arms overall or within subgroups. Where appropriate, we also controlled for relevant baseline variables or the target child's gender to further isolate the impact of the treatment.

IV. KEY RESEARCH FINDINGS

To what extent is the home visiting process of HFNY consistent with the HFA model?

The HFA program model is defined by a set of 12 research-based critical elements that reflect the primary objectives of the program model (Prevent Child Abuse America, 2001). We assessed HFNY's adherence to the elements of the national model to provide a meaningful context for interpreting findings from the outcome study.

- The average length of enrollment in HFNY was 20.68 months (SD=18.47). Just over half (52%) of the participants remained enrolled in the program by one year post-enrollment.

By two years post-enrollment, 33% of participants were still receiving home visiting services. Few families were still enrolled at three (22%) and five (4%) years.

- The average number of visits was 33.29 (SD=30.64).
- The majority of program participants received at least 75% of the expected visits congruent with the level of service to which they are assigned for five out of seven levels (Levels 1P, 1SS, 3, 4 & X). However, program participants did not receive the prescribed number of visits on Levels 1 and 2, when the frequency of the number of expected visits was high.
- On average, parent-child interaction activities occurred in 70% of each family's visits. Child development activities occurred in 63% of visits. Self-sufficiency activities occurred in 49% of visits, while Crisis Intervention activities occurred in 6% of visits.
- 80% of families who enrolled in the program had at least one referral for services other than HFNY, with an average of 10.79 (SD=16.46) referrals.
- 82% of the families who had a referral received at least one service as a result, with services being received for about 53% of the total number of referrals issued.

Does HFNY effectively reduce child maltreatment?

- Consistent with prior findings, HFNY mothers used serious physical abuse less frequently (.03 versus .15, $p < .01$) than mothers in the control group, and used non-violent discipline strategies more frequently (49.27 versus 45.27, $p < .05$). Target children also reported lower rates of minor physical aggression for HFNY mothers (70.8% versus 77.2%, $p < .05$).

- For the sample overall (n=1173), no differences were noted for cumulative rates or number of confirmed CPS reports for physical abuse or neglect or foster care between random assignment and the target child's seventh birthday.
- HFNY mothers in the HPO subgroup were less likely to engage in psychological aggression (79.7% versus 91.2%, $p < .10$) and less frequently used minor physical aggression tactics (3.7 versus 5.5, $p < .10$) than their counterparts in the control group. These findings are consistent with results found for this subgroup in years 2 and 3.
- For the HPO subgroup (n=179), differences in the cumulative rate of confirmed CPS reports for physical abuse or neglect were observed for the period from ages five through seven: 19.3% of the target children in the control group had a confirmed report versus 9.9% of the HFNY group ($p < .05$). This pattern of results did not extend to other areas of child welfare services such as having a track initiated for services or entering foster care.
- HFNY produced unexpected and unprecedented differences in rates of subsequent reports for HFNY mothers in the RRO subgroup (n=104). As compared to their counterparts in the control group, HFNY mothers had
 - lower rates of confirmed CPS reports for any type of abuse or neglect:
 - (41.5% versus 60.4%, $p < .10$);
 - lower rates of reports when the study mother was the confirmed subject:
 - (38.2% versus 57.4%, $p < .10$)
 - lower rates of confirmed reports involving physical abuse:
 - (3.3% versus 13.4%, $p < .10$)
 - a smaller number of total confirmed reports for mothers as the confirmed subject
 - (.8 versus 1.6, $p < .05$)
 - lower rates of preventive, protective, and placement services initiating:
 - (38.02 versus 60.02, $p < .05$)

Does HFNY limit the emergence of precursors to delinquency?

- More children from the HFNY group were reported to participate in gifted programs as compared to children in the control group (AOR: 2.80, $p < .01$). Fewer children in the HFNY group were receiving special education services (AOR: .70, $p < .10$) or self-reported skipping school (AOR: .35, $p < .01$); this latter finding was not supported by maternal reports (AOR: 1.08, ns).
- HFNY children in the HPO subgroup were less likely to score below average on the PPVT-IV (AOR: .43, $p < .05$); less likely to repeat a grade (AOR: .45, $p < .10$), and more likely to participate in a gifted program (5.8% versus 0%, $p < .10$).
- No significant differences were detected between the groups for the sample as a whole or within the HPO subgroup for problem behaviors, socio-emotional difficulties, and self-regulation.

Do the long-term benefits of HFNY outweigh its costs?

We conducted the cost benefit analysis from the perspective of government in order to examine monetizable costs and benefits generated by involvement in HFNY.

- Overall, women in HFNY generated a net savings of \$628 in government costs. This resulted in a recovery rate of 15% of the cost to provide HFNY services.
- For women in the RRO subgroup, investment in HFNY produced a net savings in government costs of \$12,395 per family and a return of \$3.16 for every dollar invested by the time the target child was 7 years old. This amounted to a 316% recovery of the initial \$3,920 net HFNY cost invested.

- HFNY women in the HPO subgroup generated a savings of \$1020 per family in the net cost to government, recovering 25% of the initial investment in the program by the target child's 7th birthday.

V. CONCLUSIONS & IMPLICATIONS

1. HFA-based programs can produce sustained effects with a diverse population.

Sustained effects were detected on parenting for the sample as a whole. Thus, the program produced significant differences among a very diverse group of families on both cross-sectional indicators of school engagement and longitudinal indicators of parenting. Given the concentration of significant non-monetized findings for parenting and the not yet assigned benefits for outcomes related to school, we consider the results of the cost benefit analysis to be an underestimate of the savings that might be expected to accrue following the initial investment in the program. Accordingly, we strongly recommend that the program continue to target a diverse group of mothers who are at considerable risk to maltreat, live in challenging communities, and depend on their home visitor to help their child chart a life course that averts risks for delinquency and promotes experiences associated with long-term school success.

A second important implication pertains to discussions regarding the effectiveness of different nationally-based home visiting models. Findings from the Year 7 follow-up suggest that HFA-based programs delivered by paraprofessionals can produce sustained effects on parenting that extend past the intended period of service. This pattern of effects helps to fill an important gap in the research on HFA-based home visiting programs.

2. Who is offered home visiting services matters.

We also observed significant differences in program effects depending on who was offered HFNY services. Analysis of two important subgroups revealed several effects of clinical

significance. With regards to research, we recommend that future evaluations of home visiting programs corroborate the current findings with larger samples to allow for statistical tests that are adequately powered to detect small to medium effects. In addition, study designs that are stratified from the outset would help to minimize potential differences across the treatment groups.

With respect to practice, the subgroup findings suggest potential ways to optimize resources.

Establish strong links between local department of social services and HFNY. Program effects on confirmed reports of child abuse and neglect were most robust among the group of HFNY women with prior confirmed CPS reports. These findings are particularly significant given the lack of evidence that other interventions can successfully lower rates of maltreatment recurrence. Administrative cost data also suggest that these families are less resource dependent. Thus, we strongly urge that the effect produced for HFNY's home visiting program be viewed as an opportunity to create meaningful differences in the lives of other families with prior histories of confirmed reports. We recommend encouraging local child protective services agencies to refer recently or actively indicated CPS cases to HFNY when the mother is expecting or has recently delivered a child.

Prioritize services for those entering during the prenatal period, especially women fitting the descriptions of the two subgroups. With regards to less severe indicators of harsh and punitive parenting and child outcomes, effects were most pronounced among young mothers enrolled prior to the birth of their first child. This subgroup also holds the potential for considerable long-term savings. The benefits estimated in the current study are likely underestimates of the percent recovery because many of the benefits realized for the HPO sample

are ones that are not readily monetizable. Specifically, there is the promise of greater returns resulting from the sustained impacts on rates of harsh parenting seen for this group of mothers at Years 2, 3, and 7, and the marked improvements in children’s receptive language skills by age 7 were we able to place a value on them.

In light of these findings, we recommend that programs focus screening efforts on all at-risk pregnant women in a community, rather than adhering to the universal focus on all *new mothers*. Prioritizing prenatal service initiation would also capitalize on the program’s effectiveness in helping mothers attain better birth outcomes (Lee et al., 2009). Furthermore, when young, first-time mothers or those with a prior substantiated report are referred during pregnancy, we recommend giving priority to these individuals, rather than limiting home visiting services to these groups. Home visiting services for women who recently gave birth would be offered whenever a slot is available. This recommendation preserves the opportunity for mothers with recent newborns to access services, takes full advantage of the opportunity for individual women to benefit from prenatal services (i.e., delivering a healthy weight baby), and maximizes the opportunity for the program to effect the greatest degree of change possible.

3. Examining patterns of effects on neglect may inform program practice.

Consistent with earlier findings from the trial, the HFNY home visiting program presented with both strengths and weaknesses. While significant results can inform our recommendations for policy and practice, so too can changes in the pattern of results over time. In the current study, administrative indicators and maternal reports of neglect provide a rich context for informing program practice.

In earlier waves of the RCT, self-reported parenting provided some evidence to suggest that HFNY might prevent program participants from neglecting their target child (DuMont,

Mitchell-Herzfeld et al., 2008). By Year 7, neglect indicators suggest that the early evidence of a program effect in this area has attenuated. However, results for the newly tested RRO subgroup suggest that the program can effectively prevent neglect for at-risk families, even at the level of administrative reports. This disparity creates an opportunity to examine how the program can be effective for one group, but not generalize to the entire sample.

A recent collaboration utilizing data from three separate longitudinal studies reported consistencies related to indicators of economic hardship and parent well-being when predicting neglect allegations (Shook-Slack et al., submitted). Thus, the potential for home visitors to promote parent well-being and lower the level of economic hardship for program participants may be of critical importance to the development of more effective service delivery strategies for preventing neglect. This may require intensifying or altering efforts to promote mothers' mental health and/or self-sufficiency activities. Thoughtful attention to this important area may benefit HFNY's participants more broadly.

Conclusion

Home visiting presents a unique opportunity for trained workers to forge enduring relationships with families at a time when parents are vulnerable and the developmental path of the newborn is particularly malleable. The current study presents timely evidence to suggest that involving families in home visiting services early on promotes positive experiences within the home during the initial years of life, for both the mother and the child. These benefits range from healthier birth outcomes (Lee et al., 2009) to healthy parenting (DuMont, Mitchell-Herzfeld et al., 2008; Rodriguez et al., forthcoming) to positive school experiences.

CHAPTER 1: PROJECT DESCRIPTION

Overview

Providing programs that promote positive outcomes in the lives of poor families is challenging. The current study utilizes a seven-year longitudinal randomized controlled trial (RCT) to evaluate the effectiveness of a state-administered home visitation program in preventing child maltreatment and risks for delinquency. Healthy Families New York (HFNY), which is based on the Healthy Families America (HFA) model, was established as a strengths-based, intensive home visitation program with the explicit goals of 1) promoting positive parenting skills and parent-child interaction; 2) preventing child abuse and neglect; 3) supporting optimal prenatal care, and child health and development; and 4) improving parent's self-sufficiency. In 2009, there were 39 HFNY program sites being operated in high need areas throughout New York State. To date, HFNY is the only HFA-based program recognized as a "proven program" by RAND's Promising Practices Network. In part, this accomplishment is due to its careful implementation, extensive data support system, and the methodological rigor and comprehensiveness of its RCT.

In 2000, the New York State Office of Children and Family Services (OCFS) Bureau of Evaluation and Research, in partnership with the Center for Human Services Research at the University at Albany, initiated a RCT at three sites with the HFNY home visiting program. Families eligible for HFNY at each site were randomly assigned to either an intervention group that was offered HFNY services or to a control group that was given information on and referral to appropriate services other than home visiting. Baseline interviews were conducted with 1173 of the eligible women (intervention, n=579; control, n=594), and follow-up interviews at Years 1 and 2 achieved high rates of retention. In the third year of the RCT, a reduced sample of mothers

were videotaped interacting with their three-year-old children (n=522). In addition to data gathered during the follow-up interviews, information regarding study participants' involvement in reports of child maltreatment also was extracted and coded from Child Protective Services records.

In 2006, funding from the National Institute of Justice (NIJ) made it possible to extend the trial to a seventh year to address four important questions:

- 1) To what extent is the home visiting process of HFNY consistent with the HFA model?;
- 2) Does home visiting effectively prevent or reduce child maltreatment?;
- 3) Does home visiting limit the emergence of precursors to delinquency?; and
- 4) Do the long-term benefits of an HFA-based home visiting program outweigh its costs?

Given the number and diversity of goals, the current study provides an initial look at each of these important topics. Subsequent topic-specific manuscripts will more fully explore each outcome area with the level of depth and detail they deserve.

Data were collected from multiple sources to respond to the above stated goals. Mothers in both the intervention and control groups were re-interviewed at the time of the target child's seventh birthday. Interviews included information about parenting, the child, earnings, and household composition. Field staff completed 942 interviews with the original study participants. For the first time, we also conducted interviews with target children. Interviewers completed face-to-face assessments with 800 of the children who were born and reached the age of seven. The target child interviews assessed children's receptive vocabulary skills, emotional health, self-regulatory abilities, and problem behaviors. The research team also extracted or obtained administrative data pertaining to Child Protective Services reports, foster care placements, federal and state-supported benefits, and program services and costs. Collectively,

results of our examination will inform national child welfare policy about the effectiveness, costs, long-term benefits, and mechanisms through which a home visiting program achieves its impacts.

The remainder of this chapter provides the context for the HFNY program, randomized controlled trial, and current study. Chapter 2 presents an overview of the methods adopted to address the key study questions, including a description of the overall sample, data sources, protocols used to obtain mother and child interviews, and an overall plan for analysis. Chapter 3 presents results describing the characteristics of respondents at the time of both the baseline and Year 7 interviews, the equivalence of the treatment conditions for each sample (i.e., baseline, mom interview, and target child interview). Subsequent to these descriptions, the report is organized topically to keep issue-specific methods, measures and plans for analyses in a close proximity to their corresponding findings. Chapter 4 describes the fidelity assessment; Chapter 5 presents the evaluation of the effectiveness of the program on parenting indicators related to abuse and neglect; Chapter 6 discusses data and results from the analyses of precursors to delinquency, and Chapter 7 presents the cost benefit analysis. The final chapter synthesizes these findings and discusses their implications for the practice and policy of home visiting

Background

In 2007, more than 3.5 million children were referred for investigation as alleged victims of child maltreatment, and nearly 800,000 were substantiated as victims; over 55% of the victimized children were age seven or younger (U.S. Department of Health and Human Services, 2010). The consequences of child abuse and neglect are of great concern, both for the children and society. When parents employ harsh, abusive, or neglectful parenting practices during the early years of life, children are at greater risk to engage in violence, substance abuse, juvenile

delinquency, and adult criminal behaviors (Eron, Huesmann & Zelli, 1991; Fergusson, Horwood & Lynskey, 1996; Herrenkohl, Huang, Tajima & Whitney, 2003; Smith & Thornberry, 1995; Widom, 1989). These early experiences may teach the child to model inappropriate or violent behaviors (Farrington, 1991), and to misinterpret people's motivation or intentions (Dodge, 1980; Dodge, Bates & Pettit, 1990). Child abuse and neglect may also cause biological, neurological, or cognitive problems that interfere with the child's ability to appropriately process (Feldman & Downey, 1994) or effectively regulate his or her responses to others or stressful situations (Schatz, Smith, Borkowski, Whitman & Keogh, 2008). Moreover, child abuse and neglect places considerable economic strains on child protective service systems, schools, hospitals, and taxpayers (Children's Safety Network Economics & Data Analysis Resource Center, 2000; Fromm, 2001; Mercy, Butchart, Farrington & Cerda, 2002).

Given the consequences of maltreatment and the influence of early experiences on children's development and behavior, a number of prevention programs have been developed to work directly with families during the initial years of a child's life (c.f., Daro & McCurdy, 2007; Eckenrode & Runyan, 2004; Reynolds, 2004; Yoshikawa, 1995). Preventive efforts such as home visiting attempt to reduce risk factors associated with child abuse and neglect, improve the situations in which families live, enhance children's newly developing abilities (Reynolds, 2004), and bolster protective factors that may buffer children from the effects of maltreatment (National Research Council, 1993). In addition, home visiting is proactive in its approach (Institute on Medicine, 1989) and brings services directly to families, who are often more willing to enroll when the barriers to participation are minimized (e.g., lack of transportation or childcare). Home visitors' ability to access families' everyday settings uniquely positions this service delivery strategy to integrate skills and information regarding healthy parenting practices,

child development, and self-sufficiency into the families' daily routines. Accordingly, home visitation is a highly recommended and broadly practiced strategy for promoting children's health and development and preventing child abuse and neglect (Guterman, 2001; U.S. Advisory Council, 1990). One especially popular home visiting model is Healthy Families America (HFA) which, since its inception in 1992, has become one of the most widely disseminated home visiting programs in the nation (Díaz, Oshana & Harding, 2004; Leventhal, 2005).

Healthy Families America (HFA) Model. The HFA program model is a nationwide initiative providing intensive home visiting services to expectant and new parents. The HFA model stipulates programs use trained paraprofessionals and professionals to provide voluntary services and referrals to families in an effort to promote positive parenting, enhance child health and development, and prevent child abuse and neglect (Díaz, Oshana & Harding, 2004). Home visitors typically live in the same communities as program participants and share their language and cultural background. The program model supports enrollment beginning early in the prenatal period through three months postnatally. The HFA model permits programs to serve mothers of all ages without regard to whether they have other children. Once enrolled, the model encourages the provision of long-term services, lasting throughout the first three to five years of the target child's life. It emphasizes a strength-based approach that includes promoting parent-child bonding and positive interactions, educating parents about child health and development, helping parents to access community resources, and using family and community supports to assist parents in addressing problems such as parental substance abuse or poor mental health.

Healthy Families New York. Modeled after the HFA initiative, the HFNY program was established in 1995 by the New York State Office of Children and Family Services to provide voluntary, comprehensive, and intensive home visiting services to expectant or new parents who

are identified as being at risk of abusing or maltreating their children. The program is primarily delivered by trained paraprofessionals who typically come from the communities that the program targets for service. Expectant parents and parents with an infant under three months of age who live in high risk target areas and who are considered to be at risk for child abuse or neglect are screened by various collaborative community agencies, such as prenatal care providers, hospitals, and other community service providers. Individual HFNY programs also engage in their own outreach efforts to identify and screen prospective families. Families who screen positive are referred to the HFNY program in their community, where they are systematically assessed by trained Family Assessment Workers (FAWs) for the presence of specific risk factors that place families at risk for child abuse or neglect, using the Kempe Family Stress Checklist (1976). Families who receive a “positive” assessment (a score of 25 or higher on the Kempe) and voluntarily accept services are enrolled in the program.

Families are provided intensive home visitation services bi-weekly during the prenatal period, weekly until the child is at least six months old, and periodically thereafter based on the needs of the family until the child begins school or Head Start. Home visits typically emphasize content that is appropriate to the particular service level on which the family is currently assigned. For example, visits on the prenatal level focus on promoting adequate prenatal care and providing information regarding fetal development, as well as preparing the family for childbirth and providing instruction on the care and safety of a newborn. Postnatal visits focus primarily on promoting positive parent-child interactions, educating parents about child growth and development, and enhancing family functioning and self-sufficiency.

HFNY programs determine the most appropriate curricula to use during home visits based on the specific needs and characteristics of individual families. At all stages, home

visitors provide support, education, information, and activities designed to promote healthy parenting behaviors and child growth, including proper nutrition, age-appropriate behaviors, and positive discipline strategies. Home visitors also help mothers access health care and other services as needed; identify and address issues regarding positive family functioning; and discuss childcare, education, training, and employment options.

HFNY's Randomized Controlled Trial and the Year 7 Follow-Up

As noted earlier, the current study addresses four critical issues in home visiting research. The RCT includes baseline data for the 1173 women who met the assessment criteria for HFNY and were randomly assigned to either the intervention or control group. Following the baseline interview, the research team obtained follow-up data from administrative records, the HFNY home visit data management information system, and in-depth interviews with mothers at the time of the target child's first, second, and seventh birthdays. Study retention rates were high, with 90% of the eligible women reinterviewed at Year 1, 85% reinterviewed at Year 2, and 80% at Year 7 (n=942). The Year 7 follow-up also included interviews with 800 target children.

Below, we describe literature related to each of the study's goals and briefly explain how the study will employ data from the evaluation to address the issues outlined.

To what extent is the home visiting process of HFNY consistent with the HFA model?

Researchers in the field of home visiting have called for increased understanding of the home visiting process (Gomby, 2007; Korfmacher et al., 1998; Olds, Sandler & Kitzman, 2007).

Accordingly, a handful of researchers have conducted case record reviews and fidelity assessments to better understand the mechanisms at work during a home visit (c.f., Duggan et al., 2004; Harding et al., 2004; Kessler et al., 2008). In the case of the HFA model, Harding and colleagues (2004) conducted a national survey of the model's implementation, which represented

about 100 sites from nine different states. In brief, researchers reported that home visits typically last one hour, cover a variety of topics, and consistently and frequently include activities that involve parent and child interaction, which is central to the HFA model. Families served across the programs are diverse in terms of race/ethnicity but consistently present with high levels of risk on the Kempe assessment. The majority of programs serve parents regardless of the number of prior children; about one-third provide services only to first-time mothers or teen parents. Once enrolled, the intensity and timing of visits delivered to participants varied considerably across sites. Several sites reported challenges associated with scheduling visits and consequently delivered considerably less than the intended dose. In addition, sites consistently reported that only about one half of the enrolled families continue to participate in the program past one year (Harding et al., 2004). Thus, while some practices appear fairly constant across the programs, other qualities of the model are context-dependent and site-specific.

To provide the most relevant context for interpreting results from the outcome portion of the HFNY evaluation, the current study explores how the HFNY program implements the twelve “critical elements” at the core of the HFA model, which are informed by research and theory and are intended to facilitate the implementation of high quality home visiting services (Prevent Child Abuse America, 2001). The twelve critical elements provide guidance on the general content, duration, and intensity of visits, but choices about particular curricula, emphasis, and styles of implementation are made at the level of the individual program or family. In sum, the HFA model both grants individual programs latitude to adapt or fit specific components of the model to the needs of the families and communities they serve (Kessler et al., 2008), and provides a well-articulated framework to promote consistent parameters across programs.

Evaluators and proponents of HFA suggest that the implementation of the model provides the context for interpreting a program's success or failure (Duggan et al., 2000; Gomby 2007; Harding, Galano, Martin, Huntington & Schellenbach, 2007; Olds, Sandler & Kitzman, 2007). Given the potential for differences in the intensity and timing of service delivery as well as the choices regarding the style and focus of implementation, it is important to understand how HFNY's implementation of the HFA model may facilitate success in some areas and/or present obstacles to its effectiveness in others. To date, HFNY has reported a number of positive outcomes for at-risk families (DuMont, Mitchell-Herzfeld et al., 2008; DuMont, Rodriguez et al., 2008; Lee et al., 2009; Mitchell-Herzfeld et al., 2005). Early findings from the randomized controlled trial produced results that are both consistent with findings from trials evaluating other HFA-based home visiting programs as well as suggestive of strengths that may be specific to New York's implementation of the model (Caldera et al., 2007; Duggan et al., 2007; Duggan et al., 2004; Landsverk et al., 2002). Accordingly, we seek to provide a description of the extent to which HFNY adheres to the expectations specified by the critical elements of the HFA model for the study respondents randomly assigned to the intervention condition. Aggregate descriptions of respondents' experiences with the intervention are informed by data extracted from the HFNY management information system. A presentation of the details of the intervention may help to explain why the program excelled in some areas but not in others (Korfmacher et al., 1998).

Does HFNY effectively reduce child maltreatment? The extension of the randomized controlled trial to a seventh year was initiated to provide a more comprehensive assessment of the long-term effectiveness of HFNY in preventing the maltreatment of children. Evaluations of home visiting reveal that program effects on official reports of abuse and neglect during the early years of life have been scant (Duggan et al., 2004; Duggan et al., 2007; DuMont, Mitchell-

Herzfeld et al., 2008; Olds, Henderson & Kitzman, 1994), but may, as Olds and colleagues reported (1997), emerge later in life. In contrast, results from several randomized controlled trials suggest that in the first one to three years of life, HFA-based programs have more favorable effects on less severe forms of negative parenting, such as minor physical and psychological aggression (Duggan et al., 2004; Duggan et al., 2007; Landsverk et al., 2002; Mitchell-Herzfeld, Izzo, Greene, Lee & Lowenfels, 2005; DuMont, Mitchell-Herzfeld et al., 2008). This is important, as harsh and coercive parenting behaviors are prominent risk factors for later child behavioral problems (Eddy, Leve & Fagot, 2001; Knutson, DeGarmo, Koepl & Reid, 2005; O'Connor, Deater, Rutter & Plomin, 1998). The prevention of negative parenting practices is also a critical element of the HFA model.

Consistent with other studies that have shown effects on less severe forms of negative parenting, program impacts on self-reported harsh and neglectful parenting emerged as the target child turned one, with significant reductions in the frequency with which HFNY parents reported committing acts of severe physical abuse, minor physical aggression, and psychological aggression against their children as compared to parents in the control group, and a trend showing that HFNY mothers were less likely to report neglecting their children. At Year 2, the impacts on frequency of abusive and neglectful parenting identified at Year 1 shifted to only the more serious acts, yet these behaviors were still not reflected in the rates of child protective services reports. HFNY mothers reported committing one-quarter as many acts of serious abuse at age 2 as control mothers (.01 versus .04, $p < .05$). Since previous research suggests that the strongest benefits of home visitation programs may not become evident for several years, an important question in the current study is whether these early effects will sustain and whether new effects emerge as the target children develop. The current study provides information

regarding effectiveness of the program in preventing child maltreatment using substantiated child protective services reports and follow-up interviews with 942 mothers seven years after they were randomly assigned to either the HFNY intervention condition or a control condition.

The study also evaluates how who is offered home visitation services may affect program impacts on child maltreatment. As mentioned previously, the HFA model permits home visiting programs to serve families from diverse backgrounds that vary on characteristics such as their pregnancy status, parity, age, and life experiences. This inclusive approach to services creates an unusual opportunity for the HFNY trial to isolate particular subgroups of families whose various life situations may differentially facilitate, reinforce, and/or challenge home visitors' efforts to affect parenting. Of particular interest are two policy-relevant subgroups that differ considerably in their demographic characteristics and life experiences. One group, which we refer to as the High Prevention Opportunity (HPO) subgroup, was identified in earlier work and consists of young, first-time mothers who have the opportunity to initiate home visiting services while still pregnant. We previously hypothesized that offering home visiting to first-time parents during adolescence presents an exceptional opportunity to promote healthy parenting practices before harmful patterns become entrenched (DuMont, Mitchell-Herzfeld et al., 2008; Rodriguez et al., 2010).

As an extension of our earlier work, we also evaluate the relationship of the program to confirmed child maltreatment reports within a group of women who have had at least one substantiated child protective services report (as a non-victim) prior to random assignment (RA); a subgroup that we refer to as the Recurrence Reduction Opportunity (RRO) subgroup. These women present home visitors with a very different challenge: to prevent the recurrence of a child protective services report. This subgroup is of particular interest to state agencies that administer

both child protective services and early intervention or prevention services. The HPO and RRO subgroups are described further below.

High Prevention Opportunity subgroup. As mentioned, we previously hypothesized that young, first-time mothers who engage in the program prenatally may be particularly receptive to services and instruction given their “heightened sense of vulnerability” (Olds et al., 1999, p. 46), and the increased malleability of the neural systems responsible for self-regulation and decision-making during adolescence (National Research Council, 2006). Thus, while teen parenthood may compromise youth’s ability to successfully cope with the stressors that arise from parenting (e.g., George & Lee, 1997; Stier, Leventhal, Berg, Johnson, & Mezger, 1993; Zuravin, 1988), early home visiting services may provide opportunities to both prevent negative practices from emerging and to implement healthy alternative strategies. Prior to the Year 7 follow-up, we tested our hypothesis at the Year 2 and Year 3 follow-ups using a number of different assessment methods. At Year 2, we tested a series of interactions to evaluate the potential moderating role of the HPO group on the effectiveness of services, and found that the impact of the program varied as a function of who was offered it. Specifically, young, first-time mothers in the HFNY group who were randomly assigned at 30 weeks of pregnancy or less were significantly less likely than their counterparts in the control group to engage in minor physical aggression in the past year (51% versus 70%) and harsh parenting in the past week (41% versus 62%). In contrast, the differences in the prevalence rates of minor physical aggression and harsh parenting between the intervention and control groups for the more heterogeneous group of mothers (those who varied in age, pregnancy status, prior involvement with the child protective services system, and presence of other offspring) were negligible (DuMont, Rodriguez et al., 2008).

This finding was replicated using data from the observational assessments at Year 3, which also showed that first-time mothers under age 19 who were offered HFNY early in pregnancy were less likely than similar mothers in the control group to engage in harsh parenting behaviors while interacting with their children during several structured tasks. Again, no differences between the HFNY and control groups were observed among the older or multiparous mothers (DuMont, Rodriguez et al., 2008; Rodriguez et al., 2010). Based on the consistent results across times and methods, we suggested that members of the HPO subgroup may be more receptive to education and guidance regarding avoidance of negative parenting behaviors than older or multiparous mothers, who may have already engaged in maladaptive parenting before program entry (DuMont, Mitchell-Herzfeld et al., 2008). Given the established nature of the HPO subgroup, the effectiveness of the program at Year 7 is evaluated using a within subgroup approach.

Recurrence Reduction Opportunity subgroup. Although the HFA model was not designed to intervene with abusive or neglectful parents in order to avert further maltreatment, almost ten percent of families participating in HFNY had engaged in abuse and/or neglect prior to enrolling in the program. We previously hypothesized that the participation of such women in the HFNY program is likely to dilute the impact of the program on child maltreatment by introducing the possibility that maltreatment of the current child may have already occurred prior to or concurrent with program entry, and that standing patterns of negative behavior need to be removed or replaced to avert subsequent maltreatment. The current study provides an opportunity to prospectively test this hypothesis by examining the accumulation of CPS reports where the mother and/or child were the confirmed subject or victim (respectively) between random assignment and the target child's seventh birthday. In line with previous research

showing mixed and often limited impacts of home visiting in preventing the recurrence of child maltreatment (c.f. DePanfilis & Zuravin, 2002; MacMillan et al., 2005), we expected that the lack of household resources or the poor parenting practices suggested by the presence of a prior substantiated report would either hinder the home visitor's ability to establish a trusting relationship with the respondent or present considerable challenges in reversing neglectful and abusive parenting practices.

The HFNY trial is one of the few evaluations of a home visiting program with a sample diverse enough to examine how the program affects two groups who, from the outset, present home visitors with vastly different challenges. Whereas the randomized trial of the program that inspired the HFA model, Hawaii's Healthy Start program (Duggan et al., 2004) was, as a function of the program's target population, limited to women who had already given birth, HFNY's evaluation included young women who were randomly assigned to the intervention or control groups prior to the birth of their first child, as well as those who had already given birth or were even confirmed in a child protective services report prior to random assignment. This variation allows us to identify and contrast a relatively homogeneous subgroup of our sample that has not yet had the opportunity to abuse or neglect their children, first-time mothers under the age of 19 who were randomly assigned at 30 weeks of pregnancy or less (HPO), with another fairly homogenous yet distinct subgroup of women, those who were already involved in a confirmed child protective services report (RRO). Together, the two subgroups represent approximately one-quarter of the sample, with 15% of the whole sample falling within the HPO subgroup and 9% comprising the RRO subgroup.

Although the sample sizes for these two subgroups are limited, to the extent possible, we evaluate the effects of the program on their outcomes and, where appropriate, examine the

potential mechanisms through which HFNY achieves its effects. These later analyses build on information gleaned from the fidelity assessment and impacts of the program identified during earlier interviews. Of particular interest is information regarding the early parent-child relationship, how it is supported by the home visitor, and whether impacts identified during earlier follow-ups, such as self-reported use of serious physical abuse and limit setting, have implications for cumulative rates of abuse and neglect.

Does HFNY limit the emergence of precursors to delinquency? A third major goal of the study is to look beyond the prevention of violence towards children to also examine HFNY's ability to prevent or limit behaviors and characteristics that are frequently associated with delinquency. These behaviors may represent risks for future transgressions and later violence (Broidy et al., 2003) or reveal skills and strategies that can play a protective role in the child's life (Reynolds, 2004).

A growing body of literature suggests that early childhood behaviors can provide important clues about later development (U.S. Department of Health and Human Services, 2001). For some children, delinquent behaviors emerge early and persist, while for others the effects are delayed and the onset does not occur until late adolescence or young adulthood. For children who initiate antisocial behaviors prior to puberty, the link between childhood risk factors and long-term involvement in violence is especially pronounced (Lipsey & Derzon, 1998; Hawkins et al., 1998; Smith & Thornberry, 1995). It is also well established that aggression, impulsivity, and oppositional behaviors in early childhood are distinct predictors of adolescent and adult criminality, and violence (Broidy et al., 2003; Farrington, 2005). Deviant youth are typically behaviorally troubled at younger ages and in multiple settings (Shaw & Gross, 2008). These problems may be displayed as poor self-control, emotional dysregulation, and aggression

as early as the toddler stage (Shaw & Gross, 2008). Moreover, once these problems reach the level of a diagnosis or intervention in middle childhood or early adolescence, they are often resistant to treatment (Campbell, 1995; Shaw & Gross, 2008). Conversely, when disruptive behaviors are absent throughout the toddler and preschool periods, these behaviors are unlikely to emerge in later childhood or adulthood (Shaw & Gross, 2008).

Home visitors have a unique opportunity to work with both parents and their children in the home environment during the initial years of the target child's life. This is a time when the child's behavior is particularly malleable, which increases the potential for the home visitor's relationship with the child to have a direct influence on his or her subsequent development. Alternatively, home visiting services may indirectly impact the child by influencing a range of parenting behaviors. Both effects have implications for the child's emotional, cognitive, and behavioral well-being and development. For example, maternal interactions characterized by positive affect, low levels of criticism, responsiveness, and the use of gentle guidance and support in mother-child play are associated with children's increased cognitive functioning (Pianta, Smith & Reeves, 1991) and higher performance on school readiness indicators, independent of factors such as SES and maternal IQ (Pianta et al., 1991). Childhood maltreatment and coercive parenting also are prominent risk factors for later behavioral problems, including oppositional and aggressive behaviors, self-regulatory deficits (Rodriguez et al., 2005; Sethi, Aber, Shoda, Rodriguez & Mischel, 2000), and child psychopathology (Egeland & Sroufe, 1981; Olson, Bates & Bayles, 1984). If patterns of negative parent-child interaction (Patterson, 1982) persist, they may eventually lead to delinquency, running away, teenage pregnancy, and alcohol and substance abuse (Maxfield & Widom, 1996; Widom, 1989; Widom & White, 1997). Conversely, if home visiting effectively prevents or reduces levels of

maladaptive parenting, it may diminish the opportunity for children to model abusive or non-empathic behaviors, and help them avoid engaging in a variety of problem behaviors.

While these early effects may kindle skills that play a protective role later in the child's life (Reynolds et al., 2004), evidence from randomized controlled trials of other home visiting programs presents an inconsistent patterns of results, with findings varying by developmental period, outcome of interest, and method of assessment. Three randomized controlled trials of home visiting programs reported gains in intellectual functioning during the first two years of life for children (Caldera et al., 2007; Landsverk et al., 2002; Olds et al., 2004), but no differences were detected at ages three and four (Landsverk et al., 2002; Olds et al., 1994; Olds et al., 2004). One of the few studies to follow recipients of home visiting until age six found a program effect on receptive vocabulary and mental processing skills (Olds et al., 2004a), but differences in mental ability were not present at age two (Kitzman et al., 1997) and appeared to dissipate by age nine (Olds, Sandler & Kitman, 2007).

Varied results have also been noted for problem behaviors. For example, one study reported that children who receive home visitation delivered by paraprofessionals experienced fewer internalizing symptoms at age two (Caldera et al., 2007), while a nurse home visiting program reported fewer behavior problems in the borderline or clinical range by age six (Olds et al., 2004); still others reported no early or lasting effects (McCarton et al. 1997; Olds et al. 2002). In the one study that has had the opportunity to conduct a longitudinal analysis, data revealed a sizable long-term program impact of home visiting (Olds et al., 1997). At fifteen years of age, youth born to teen mothers in the treatment group were less likely to run away or have an arrest or conviction, and reported lower levels of alcohol use and fewer sexual partners than their counterparts in the control group (Olds et al., 1997).

Thus, despite results from several studies, the potential for home visiting to impact children's early and later functioning remains unclear, particularly for HFA-based programs that have had little opportunity to evaluate children's long-term functioning. In the current study, we expected HFNY to positively impact youth by promoting individual competencies, such as cognitive skills or healthy relationships, while also encouraging home environments that recognize and reinforce strengths and achievements. In turn, these advantages may position children on a trajectory to avoid deviant and delinquent behaviors in adolescence. We hypothesized that at age seven, target children randomly assigned to the HFNY group would present with fewer problem behaviors, cognitive difficulties, and socio-emotional difficulties. To evaluate the impact of the program on child outcomes, child interviews at age seven involved interactive cognitive and behavioral assessments to estimate levels of problem behaviors, self-regulation and cognitive disabilities. Additionally, mothers provided reports of children's problem behaviors, psychological symptoms, and difficulties and successes in school. The use of multiple and independent assessment strategies is especially important since parent reports of children's problem behavior have been shown to produce biased results (Sternberg, Lamb, Guterman & Abbott, 2006).

Do the long-term benefits of HFNY outweigh its costs? Prior research has shown that the benefits of early childhood interventions exceed the costs of such programs (Karoly et al., 2005). Home visitation programs in particular have been widely promoted as an efficient use of resources. Unfortunately, few evaluations of home visiting programs have included an economic component to support this level of confidence. This issue is particularly relevant for New York State, considering its sizable investment in the delivery of HFNY to families throughout the state. In fiscal year 2007-2008, over 25 million dollars was budgeted for the program. While

HFNY has been shown to provide significant positive outcomes for families in areas such as birth weight, parenting attitudes and behaviors, and access to medical care and benefits (DuMont et al. 2008; DuMont et al. 2005; Lee et al. 2009; Mitchell-Herzfeld et al. 2005), an important yet currently unanswered question is whether the program results in cost savings to the government. An economic evaluation of a paraprofessional model for home visiting will not only be relevant to New York State, but will also be of national importance, given the use of the HFA program model throughout the country,

The Washington State Institute for Public Policy (WSIPP) (Aos, Lieb, Mayfield, Miller & Pennucci, 2004) examined the existing research on a number of prevention and early intervention programs to determine the costs and benefits of each program with regard to seven specific outcomes: educational attainment, criminal activity, alcohol, tobacco and other drug use, child abuse and neglect, teen pregnancy, and use of public assistance. The authors rated each study for methodological quality and used meta-analytic procedures to apply costs to the weighted average effect for each of the outcomes of interest as they were available. Of those examined were two home visiting programs, the Nurse Family Partnership (NFP) and HFA.

Aos and colleagues (2004) determined that the cost per family per year for the nurse home visitation program for low-income families was roughly \$9,118 (in 2003 dollars) and that benefits were substantial (\$26,298 in 2003 dollars). This works out to a return of approximately \$2.88 for every dollar invested in the program. These estimates were based on significant effect sizes and standardized outcomes found in three randomized trials, the study design valued most highly by the authors (Aos et al., 2004). In contrast, the costs and benefits presented for the paraprofessional home visitation model were almost one third of those associated with the nurse home visitation model (\$3,314 in 2003 dollars), with the benefits not quite covering the initial

investment in the program (\$2,052 in 2003 dollars), for a return of \$.62 for every dollar invested. The figures cited were based on the weighted average of outcomes produced by evaluations of paraprofessional home visitation programs. The effect sizes obtained from several of the studies were devalued due to the lack of a random assignment design, follow-ups of short duration, weakly measured outcomes, or the lack of significant effects on benefits that were monetized.

Findings from NFP and HFA program evaluations were also included in a later WSIPP publication examining the costs and benefits of evidence-based programs to prevent children from entering and remaining in the child welfare system (Lee, Aos & Miller, 2008). This study updated the outcomes valued in the previous study (e.g., child abuse and neglect, criminal activity, alcohol, tobacco and other drug use, educational attainment) and extended the focus to also include out-of-home placement. The results were similar to those obtained previously, with NFP returning \$3.02 for every dollar invested and HFA returning \$.57 for every dollar invested.

Care must be taken in comparing the two types of programs (nurse vs. paraprofessional) due to the likelihood that the net benefit and benefit-cost ratio for the paraprofessional home visitation models are lower as a result of the differences in follow-up duration, as well as limitations in the number and type of outcomes assessed (Karoly et al., 2005). In addition, one of the most important factors influencing the outcome of a cost benefit analysis is the decision regarding the benefits to which to assign a monetary value. Most cost benefit analyses focus on benefits that are easily valued and/or have been assessed in other cost benefit studies, such as use of government programs, tax revenues, and child abuse and neglect services. Unfortunately, this focus does not take into account observed effects that cannot be monetized, such as use of positive parenting strategies and improved social functioning. Other effects may not yet have been measured to a degree that would allow for proper assignment of value (e.g., increased

employment and earnings, and reductions in welfare use by children). As a result, the cost benefit analysis is likely to underestimate the true benefits of the program. Thus, results from economic analyses should be considered in tandem with these non-valued benefits, given that they are important in determining the effectiveness of a program and that program costs are often easier to value than program benefits (Plotnick & Deppman, 1999).

In the current study, the cost benefit analysis of HFNY was designed to address the following questions from the perspective of the government:

- What are the costs associated with the program?
- Does HFNY reduce spending for government supported programs?
- Does HFNY increase tax revenues?
- Do the benefits of HFNY exceed the costs?
- Do the specific characteristics and/or experiences of HFNY participants influence the costs and benefits related to the program?

In combination, findings from the study's different lines of inquiry will enhance the field's understanding of whether, how, for whom, and at what cost a paraprofessional home visitation program effectively serves women at high risk for maltreating their children. Given the recent passage of the Maternal, Infant, and Early Childhood Visitation program, the results are both relevant and timely, and may ultimately help home visiting play a more meaningful role in the lives of vulnerable families.

CHAPTER 2: METHODOLOGY

Research Design and Methodology: HFNY's Randomized Controlled Trial

Recruitment for the RCT was conducted between March 2000 and August 2001. Randomization was conducted by way of a computer program at three sites with home visiting programs that had been in operation since the HFNY's inception in 1995. All women were selected for the study following the same screening and assessment procedures used to determine eligibility for HFNY. Informed consent for the study was obtained by the Family Assessment Worker (FAW) prior to the administration of the Kempe Family Stress Checklist assessment.

During the sample selection period, 1254 mothers were deemed eligible for the study and 1173 (or 93.5% of those eligible) completed baseline interviews (intervention, n=579; control, n=594).

Baseline Sample Description. Of the 1173 mothers interviewed at enrollment, 49.8% were interviewed before the target child was born. Nearly all mothers, 99.7%, were interviewed at enrollment before six months post-birth. About a third (34%) of the mothers in the study sample were White, non-Latina; 45% were African-American, non-Latina; and 18% were Latina. Like HFNY participants statewide, women in the study sample were often young (31% under 19), and first-time mothers (55%), and had not yet completed high school or received a GED (47%). Women also self-reported or were assessed to have multiple risk factors for maltreatment. On average, women were assessed as having moderate to severe levels of risk on 5.7 of the Kempe's ten items.

Data Sources

Data for the current study come from a number of different sources, including administrative databases, interviews with study respondents, and interviews with their target

children. The Year 7 data sources provide the foundation from which study-specific measures were summarized.

Baseline Covariates. Information gathered at baseline during interviews with the study mothers is used to identify appropriate covariates and to evaluate if individual and family characteristics and resources differ across the treatment arms. Potential covariates include dichotomous variables coded to represent the mother's race/ethnicity, the mother's age, the presence of a regular partner, at least one move in the past year, being randomly assigned to the study prenatally or postnatally, and receipt of at least a GED or high school diploma. The target child's gender and age were also assessed at baseline or the next subsequent interview.

We also used data collected on the household composition and the number of prior pregnancies and births to create a variable that describes the total number of other biological children (excluding the designated target child) the respondent reported as of the baseline interview. Values for this variable range from 0 to 10 other biological children, with an average of .84 other biological children. More than half (55.2%) of the 1173 respondents reported having no other biological children at baseline.

In addition, we selected a series of summarized measures to assess mothers' level of depressive symptoms (Center for Epidemiologic Studies Depression Scale (CES-D), Radloff 1977), sense of mastery (Mastery of Psychological Coping Resources Scale; Pearlin & Schooler, 1978), and total parenting attitudes, as assessed on the Adolescent and Adult Parenting Inventory (AAPI-II, Bavolek & Keene, 1999).

Mother's Earned Income. Mother's earned income was calculated as a sum of wages of various jobs that the respondent worked during the study period. In the baseline survey, wage data were collected on one job, but for survey years 1, 2, and 7, wage data were collected for up

to five jobs. In the baseline survey, employment data were collected for the most recent job worked in the last three years prior to the survey. For survey years 1 and 2, employment data were collected on the most recent jobs worked since the last survey. In survey year 7, employment data were collected on the most recent jobs worked in the last five years prior to the survey.

HFNY Management Information System (MIS). The HFNY MIS is a centralized system used to collect and maintain comprehensive information from each HFNY site on the screen and Kempe risk assessment, characteristics and needs of the families served, the frequency and content of home visits, the nature and outcome of service referrals, progress toward program objectives, and worker demographic and training information. These data can be used to populate aggregate and individual program reports, which are used to support the quality assurance and improvement efforts employed by the Central Administrative team and individual sites. We obtained MIS information for the 579 families who were randomly assigned to the HFNY arm of the study to better understand how program components, such as visit frequency and content, relate to the outcomes experienced by families. We also created a count variable using items from the Kempe Family Stress Checklist to describe the level of moderate to severe risks present in families' lives as of random assignment. The Kempe taps a number of varied life domains and is a widely used tool for predicting parents' future risk of maltreating their children (Korfmacher 1999). The inventory was administered by a trained FAW just prior to random assignment.

NYS Administrative Databases. In order to maximize our ability to accurately identify respondents and their target children across the multiple administrative databases providing data for this study, we compiled a secure Master Person File containing all of the person information

available to us through the research study dataset. This included the respondents' and their target children's first and last names, dates of birth, sex, race/ethnicity, and social security numbers (for respondents only, if available). We used this information to conduct individual, manual searches of the NYS Welfare Management System (WMS) to obtain the unique client information numbers (CIN) for respondents and their target children, if available. WMS and CINs are used by multiple NYS agencies to track payments related to a range of state-administered services including Medicaid, public assistance, HEAP, food stamps, child care, and child welfare services. Using this method, we were able to obtain CINs for 99% of respondents and 95% of target children (including those who were never born). Various combinations of these study and system-based identifiers were then used to obtain administrative data on child protective, preventive and foster care services use, target child birth weight, and public assistance and food stamps eligibility and costs.

Child Protective Services: In order to determine whether respondents or their target children were ever the confirmed subject or confirmed victim in an indicated NYS Child Protective Services (CPS) investigation, person-based searches of CONNECTIONS, the NYS Statewide Automated Child Welfare Information System, were conducted. Designed to track calls made to the NYS child abuse and neglect hotline from intake through investigation conclusion, CONNECTIONS maintains information on all CPS investigations in a searchable database indexed by name and person identification number (PID).

To identify indicated CPS records involving respondents or their target children, a time and labor-intensive multi-stage search process was instituted. The name, sex, and date of birth for each respondent and target child were manually entered into CONNECTIONS via a phonetic-based search engine. A computer-generated list of potential system matches, rank-

ordered according to the goodness-of-fit observed between entered values and system records was generated for each individual. Potential system matches were then manually reviewed and evaluated by experienced CONNECTIONS users based on a detailed review of the information contained in the study's master person file. In order to be considered a viable system match, the generated records had to match on respondent or target child name, sex, and some combination of other key identifiers (date of birth, race, address, family members, street address, etc.).

Information was extracted from the CONNECTIONS system for indicated CPS reports occurring in the five years prior to random assignment through the target child's 7th birthday. This information included: start and end dates of CPS investigation, confirmed determinations, type of maltreatment alleged, subjects of the report, victims of the report, and severity of injury to victim.

Preventive, Protective, and Foster Care Services: The NYS Child Care Review Service (CCRS) is a legacy-based automated computer system that uses CINs to catalogue all service provision and legal activities pertaining to children in NYS who receive mandated child protective, child preventive, or foster care services. CCRS contain a cumulative record of all child preventive and protective services cases opened for services and all foster care entries, movements, and exits. We used target child CIN numbers to perform a computerized search of CCRS and extracted information for matching target children who received child preventive/protective or foster care services at any point from birth (or random assignment for those who were assigned postnatally) through the 7th birthday. This information included: service assessment dates, service assessment choice (preventive/protective/placement), start and end dates for each foster care placement, and start and end dates for trial discharges.

Public Assistance and Food Stamps: Public assistance and food stamps eligibility and cost data were obtained from the NYS Office of Temporary and Disability Assistance (OTDA) for the period from random assignment through the target child's 7th birthday. We provided OTDA with a data file containing the first and last names, dates of birth, and CINs of respondents and their target children. Using the information provided, OTDA was able to identify additional and, in some cases, multiple CIN numbers for participants. OTDA staff then used a computerized search process to extract the following benefit information for cases with matching CINs: service eligibility, payment amount, date payment was issued, program type (e.g., food stamps, public assistance), and client identification information. Using the information available, we created a dummy variable to indicate public assistance eligibility at the time of random assignment. This variable was frequently used as a covariate when evaluating the effectiveness of the program.

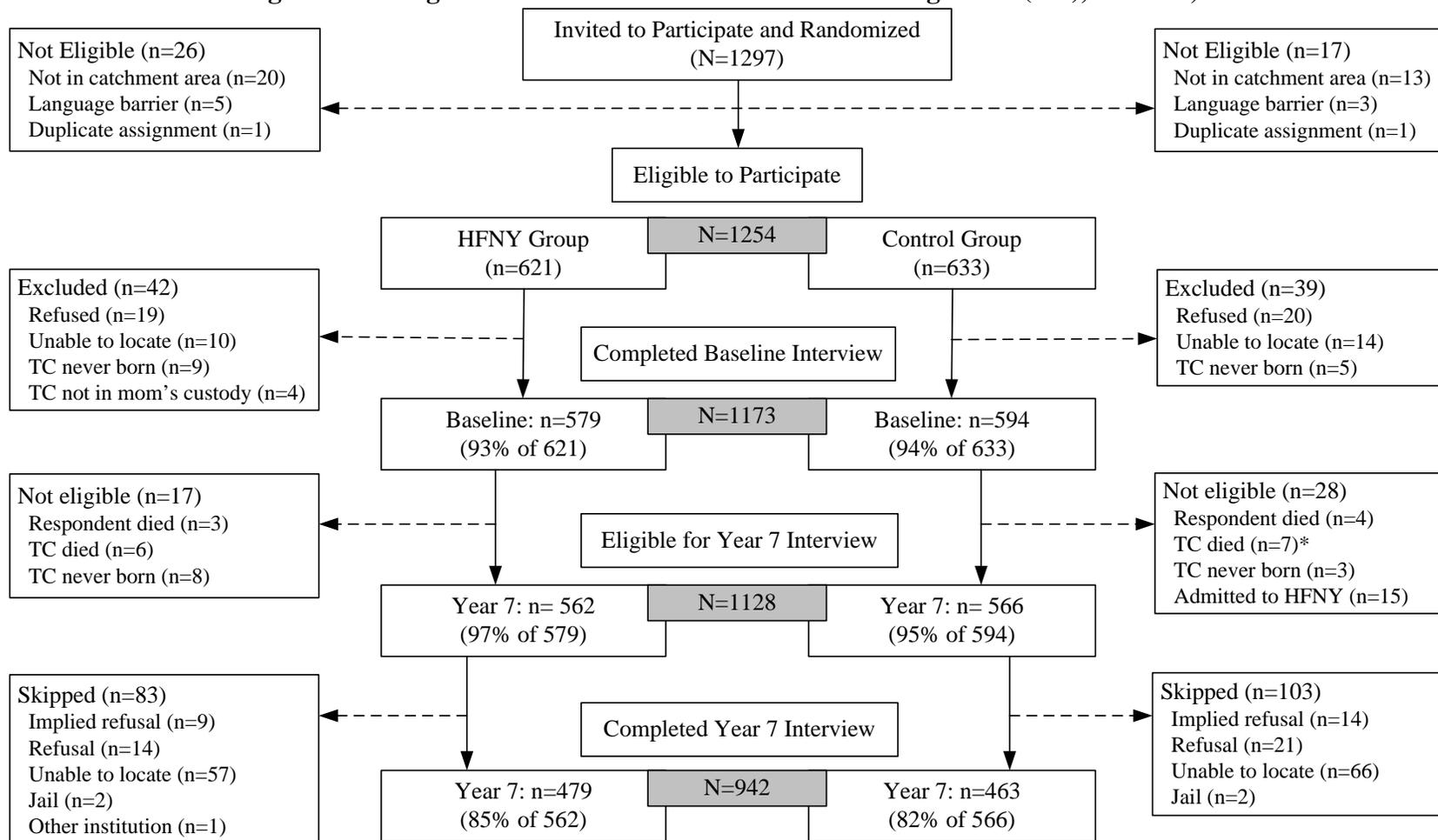
Birth Weight: We obtained birth weight data from the NYS Department of Health (DOH) for a subset of respondents who were randomly assigned to the study prenatally in order to corroborate the respondent's report of the target child's birth weight. We provided the DOH with a data file containing the county in which lived when randomly assigned, as well as the first and last names, and dates of birth of the respondents and their target children (if known). DOH staff then performed computerized searches to match the provided respondent and target child information to the birth certificate records maintained in their databases.

Year 7 Interviews

Mom Interview. As with the follow-up at Years 1, 2, and 3, two eligibility requirements applied for the Year 7 interviews: initial study respondents were eligible to participate in the mom interview at Year 7 if: (1) both the respondent and the target child were still living and (2)

women in the control condition had not received the intervention at any time between random assignment and two weeks prior to the Year 7 assessment. As shown in Exhibit 1, 1128 mothers were eligible to complete the Year 7 interview. Of the women still eligible, 942 women completed an interview at Year 7, including 479 from the intervention condition and 463 from the control condition. This represents 80.3% of the baseline sample and 83.5% of those still eligible. Given the exclusion of women from the control group who participated in the treatment, retention rates were slightly higher for those assigned to the intervention condition. The primary reasons for nonparticipation of the original study respondents at Year 7 included inability to locate and implicit or explicit refusals.

Exhibit 1. Consort diagram of biological mothers' involvement at random assignment (RA), baseline, and Year 7



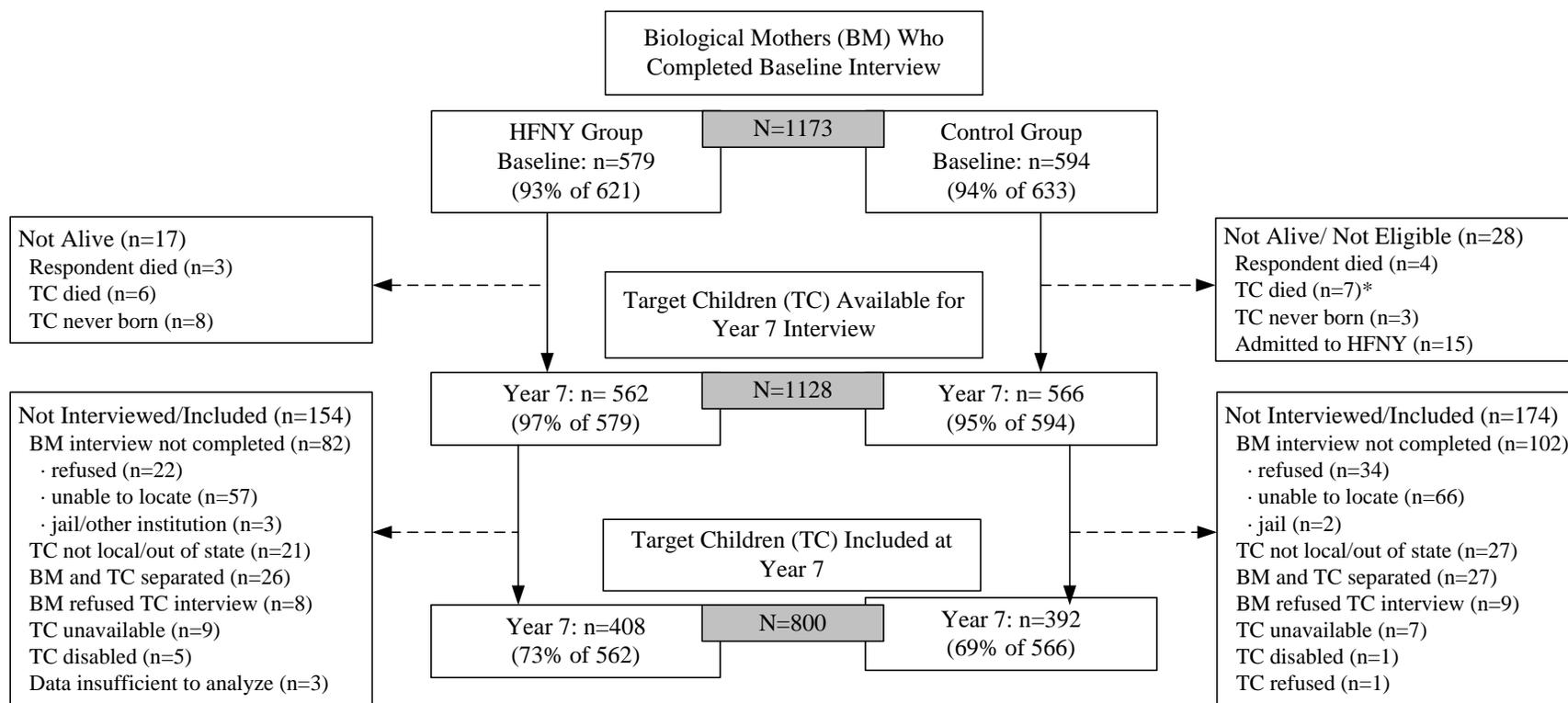
*For one family in the control group, both the respondent and the target child died. To avoid double counting this occurrence, the total number excluded from the sample includes only the respondent's death. However, the count of the number of target children who died includes the deceased child. Consequently, the number of control cases listed in the box sums to one more than the total number of cases not eligible for the Year 7 interview.

Child interview. Eligibility for participation in the child interview was more restrictive. In addition to both the criteria established for the mom interviews, the target children had to live within driving distance of an interviewer in order to facilitate a face-to-face assessment, and the study respondent had to have custody of the child in order to grant consent for the interview. Exhibit 2 displays the number of children eligible for a Year 7 assessment. Of the 1128 families with an eligible target child, 800 children were interviewed (70.9%). The primary reasons for nonparticipation for the target child included not being able to locate the mother, the mother's refusal, a separation of the mother and child, and out-of-state residence. Again, participation rates are slightly higher for children in the intervention group due to the exclusion of control cases who had received HFNY services. Specific reasons for non-participation were fairly consistent across the two groups.

Procedures

Participating mothers and their children were typically assessed in their homes by a trained interviewer, who was independent of the HFNY program and blind to group assignment. Following a lead letter, interviewers contacted mothers by telephone, described the interviews, and scheduled a time to meet. Once in the home, the interviewer answered mothers' questions about the study and obtained informed consent from the mother to conduct her interview and the child's. If applicable, the interviewer also obtained assent from the target child. Interview data were collected using laptop computers equipped with a Computer-Assisted Personal Interviewing (CAPI) system. Touch screens were used to collect sensitive information such as reports of parenting behaviors and to conduct the delay of gratification task, which was developed for the current study to measure this important precursor to delinquency.

Exhibit 2. Consort diagram of target children available for and included in Year 7 sample



*For one family in the control group, both the respondent and the target child died. To avoid double counting this occurrence, the total number excluded from the sample includes only the respondent's death. However, the count of the number of target children who died includes the deceased child. Consequently, the number of control cases listed in the box sums to one more than the total number of cases not eligible for the Year 7 interview.

The mom interview took about 60-75 minutes to complete. If the respondent was unable to complete the interview in a face-to-face setting or lived farther than a reasonable driving distance, interviews were conducted over the phone and sensitive questions, such as those concerning parenting and alcohol and drug involvement, were mailed with a prepaid postage envelope with the study's return address.

The target child interview was often collected next, although the order of the two interviews was not predetermined, and interviewers sometimes completed the visit in pairs. The target child assessment lasted approximately 30-45 minutes. Target children were offered a musical toothbrush, which was valued at about \$10.

Similar to earlier waves, the research protocol was approved by the Institutional Review Board of the University of Albany (IRB Approval #00-246).

Analysis Plan: General Approach

The information from the aforementioned data sources was used to conduct the fidelity assessment as well as to examine the program's effects on child maltreatment reports, parenting behaviors and risks for juvenile delinquency as reported by both the mother and child, and the cost benefit analyses.

Prior to analysis of the outcomes, we used Student's *t*-tests and Chi-square tests to assess the comparability of the intervention and control groups on a number of baseline demographic and risk characteristics. This was repeated for the baseline and Year 7 mother and child samples. The results of these analyses are presented in Chapter 3. At Year 7, we also tested the representativeness of mothers who completed the follow-up interviews relative to those not included in the Year 7 assessment.

All study respondents who had data were included in the analyses, regardless of their participation in the program. Consistent with the intention-to-treat approach (Hollis & Campbell, 1999; Sainani, 2010), we included respondents who were initially assigned to the treatment arm but received no treatment (9.8%) and, when analyzing administrative data sources, those participants who were assigned to the control condition but erroneously received the treatment (about 2.5%). Interview data were not available for cases from the control condition who received home visiting, but intervention cases involving no dose were included. Inclusion of the cross-over cases helps to maintain the integrity of random assignment design, but also has the potential to make it more difficult to detect a program effect.

Descriptive statistics were conducted prior to analyses to examine the distributional properties of the dependent variables and determine the most appropriate distribution and link function to apply. The dependent variables were then analyzed using generalized linear models in SAS 9.2. In all tests of the program's effectiveness, the intervention condition (1) was the primary independent variable, with the control condition (0) serving as the reference group. Covariates were included as necessary to maximize the equivalence of the two treatment arms or within subgroups. Additionally, where appropriate, we also controlled for relevant baseline variables or the target child's gender to further isolate the impact of the treatment.

CHAPTER 3: DESCRIPTION OF STUDY SAMPLES AND PARTICIPANTS

Description of Samples

Exhibit 3 shows the characteristics of the three samples we used to estimate effects on parenting, precursors to delinquency, and costs and benefits, including the baseline sample and the samples of study mothers and target children from Year 7. The exhibit reveals that the two Year 7 follow-up samples were fairly representative of the baseline, showing relatively consistent percentages and means for the majority of characteristics across all samples.

Respondents who participated in the Year 7 mother sample were comparable to study participants who were not included in the sample on a number of baseline characteristics, including treatment group assignment, age, parity, education, partner status, membership in the two subgroups, number of other biological children, levels of depressive symptoms, mastery, and parenting attitudes, earnings, and target child's gender. A few significant differences were detected, however. Compared to mothers not included in the Year 7 sample, those in the Year 7 sample were more likely to be African-American (47.8% versus 35.5%, $p < .01$) and less likely to be Latina (15.6% versus 27.8%, $p < .01$), while the percent of white women represented in each group was similar (34.6% versus 33.3%). A lower percentage of mothers in the Year 7 sample reported moving in the year prior to the baseline interview than those not included in the Year 7 sample (54.2% versus 62.8%, $p < .05$), and women completing the Year 7 interview initially presented with more moderate to severe risks on the Kempe than those not interviewed (5.77 versus 5.39, $p < .001$).

Exhibit 3: Baseline characteristics for baseline and Year 7 samples

Baseline characteristic	Baseline Sample (n=1,173)	Y7 Mom Sample (n=942)	Y7 Child Sample (n=800)
	Percentage		
Mother's race/ethnicity			
White, non-Latina	34.4	34.6	34.1
African-American, non-Latina	45.4	47.9	49.0
Latina	18.0	15.6	15.1
Mother < 19 years old	31.0	31.7	32.1
First-time mother	55.4	55.5	56.9
At least high school diploma or equivalent	47.4	47.1	46.9
Had partner	67.5	67.1	65.5
Moved in past 12 months	55.9	54.2	53.4
Receiving cash assistance at random assignment	36.5	37.8	36.3
Pregnant at random assignment	64.8	63.7	64.4
High Prevention Opportunity (HPO) subgroup	15.3	14.9	15.4
Recurrence Reduction Opportunity (RRO) subgroup	8.9	9.2	8.5
Target child female	46.1	46.5	47.3
	Mean (sd)		
Mean maternal age in years	22.45 (5.50)	22.42 (5.54)	22.20 (5.38)
Total number of other biological children	.84 (1.23)	.83 (1.22)	.77 (1.13)
Depressive symptoms (CESD)	15.64 (11.13)	15.61 (10.97)	15.53 (10.91)
Total mastery score	20.88 (2.89)	20.85 (2.89)	20.80 (2.84)
Total maternal parenting attitudes (AAPI)	134.85 (15.19)	134.51 (15.46)	134.31 (15.21)
Count of risk items (Kempe)	5.69 (1.36)	5.77 (1.38)	5.75 (1.40)
Estimated annual earnings (\$)	2679.10 (5656.41)	2750.98 (5807.73)	2603.93 (5507.06)

The children interviewed at Year 7 had mothers who generally were comparable to the study participants not represented in the child sample. Mothers of the children who were interviewed resembled those who did not have a child participating in the study on baseline characteristics such as treatment group assignment, parity, education, membership in the two subgroups, levels of depressive symptoms and mastery, earnings, and target child's gender. Similar to the sample of mothers, children in the Year 7 sample were more likely to be African-American (49.0% versus 37.8%, $p < .01$) and less likely to be Latina (15.1% versus 24.1%, $p < .01$) than those not included in the sample. Families of these children were also less likely to have moved in the year prior to the baseline interview than those not interviewed (53.4 versus 61.4, $p < .01$), and presented with a higher level of risk at the initial assessment (5.75 versus 5.56, $p < .05$). The requirement that mothers have custody of the child in order to grant consent for the child also likely contributed to a Year 7 child sample of slightly younger mothers (22.2 years versus 23.0 years) with fewer biological children (.77 versus .98) than those not followed.

Exhibit 4 presents the characteristics of the control and HFNY groups for each of the three samples. The characteristics of the two treatment arms are remarkably similar within each of the samples on most individual and family variables, demonstrating that the random assignment secured two equivalent groups and that the follow-up effort largely maintained the integrity of the initial design. Although the two groups were similar on the majority of attributes, there was evidence that they varied on a few characteristics. In two of the three samples, significantly more of the target children were female in the control than in the HFNY group. This difference has important implications for the child outcomes assessed: boys at this age consistently present with higher levels of externalizing behaviors than girls. Thus, the overrepresentation of boys in the intervention group's samples could potentially influence

outcomes. Consequently, we elected to include the target child's gender as a covariate for all outcome analyses. Additionally, while there were no differences between the groups in any of the three samples for receipt of cash assistance at random assignment, women in the control group in each of the samples had significantly higher annual earnings than women in the treatment group. Within the study mother and child samples, a smaller percentage of mothers in the HFNY group had obtained at least a high school diploma or equivalent as compared to mothers in the control group. Finally, the two groups differed within the baseline and study mother samples on the counts of Kempe items in the moderate to severe range, with mothers in the intervention condition being assessed with a slightly higher level of risk than mothers in the control group. Despite differences observed for the Kempe assessment, no significant differences were found for the self-reported levels of depressive symptoms, mastery, or parenting attitudes.

Exhibit 4: Characteristics of baseline and Year 7 samples by treatment arm

Characteristic	Baseline Sample(n=1173)		Y 7 Mom Sample (n=942)				Y 7 Child Sample (n=800)		
	Control	HFNY	Control	HFNY	Control	HFNY	Control	HFNY	
	(n=594)	(n=579)	(n=463)	(n=479)	(n=392)	(n=408)			
	Percentage		p	Percentage		p	Percentage	p	
Mother's race/ethnicity									
White, non-Latina	34.3	34.4		34.6	34.7		33.9	34.3	
African-American, non-Latina	46.5	44.4	.39	49.2	46.6	.49	50.3	47.8	.84
Latina	17.7	18.3		14.9	16.3		14.3	15.9	
Mother < 19 years old	29.8	32.3	.36	29.4	34.0	.13	30.1	34.1	.23
First-time mother	54.4	56.5	.47	53.1	57.8	.15	54.6	59.1	.20
At least high school diploma or equivalent	49.3	45.4	.18	50.8	43.6	.03	50.5	43.4	.04
Had partner	65.8	69.2	.22	65.0	69.2	.17	63.0	68.3	.12
Moved in past 12 months	56.2	55.6	.83	54.9	53.7	.71	54.1	52.7	.70
Cash assistance at random assignment	35.2	37.8	.35	36.1	39.5	.28	34.9	38.2	.34
Pregnant at random assignment	66.7	62.9	.17	65.0	62.4	.41	66.8	62.0	.15
HPO subgroup	14.8	15.7	.67	14.5	15.2	.74	15.6	15.2	.89
RRO subgroup	8.8	9.0	.89	9.7	8.8	.61	8.7	8.3	.86
Target child female	50.0	42.1	.01	49.9	43.2	.04	49.7	44.9	.17

(Exhibit 4 continues)

Exhibit 4 continued.

Characteristic	Baseline Sample(n=1173)		Y 7 Mom Sample (n=942)		Y 7 Child Sample (n=800)				
	Control (n=594)	HFNY (n=579)	Control (n=463)	HFNY (n=479)	Control (n=392)	HFNY (n=408)			
	Mean (sd)		Mean (sd)		Mean (sd)				
Mean maternal age in years	22.53 (5.43)	22.37 (5.56)	.60	22.63 (5.45)	22.22 (5.62)	.26	22.45 (5.36)	21.97 (5.40)	.21
Total number of other biological children	.83 (1.15)	.85 (1.31)	.77	.85 (1.16)	.82 (1.28)	.66	.82 (1.15)	.73 (1.12)	.27
Count of depressive symptoms (CESD)	15.61 (10.98)	15.68 (11.30)	.92	15.65 (10.74)	15.57 (11.20)	.90	15.64 (10.71)	15.43 (11.11)	.78
Total mastery score	20.85 (2.92)	20.92 (2.86)	.70	20.75 (2.95)	20.95 (2.82)	.30	20.70 (2.90)	20.89 (2.78)	.34
Total maternal parenting attitudes (AAPI)	134.45 (15.55)	135.26 (14.81)	.36	134.06 (16.01)	134.95 (14.91)	.38	133.87 (15.80)	134.73 (14.62)	.42
Count of risk items (Kempe)	5.60 (1.37)	5.79 (1.34)	.02	5.68 (1.38)	5.86 (1.37)	.04	5.67 (1.40)	5.83 (1.40)	.11
Estimated annual earnings (\$)	2998.74 (6298.98)	2351.17 (4894.49)	.05	3196.33 (6602.76)	2320.50 (4887.33)	.02	3014.80 (6190.61)	2209.17 (4733.04)	.04

CHAPTER 4: TO WHAT EXTENT IS THE HFNY PROGRAM CONSISTENT WITH THE HFA MODEL?

As described previously, the HFA program model is defined by a set of 12 research-based critical elements that reflect the primary objectives of the program model (Prevent Child Abuse America, 2001). These elements inform three main content areas: participant identification and assessment; service content and intensity; and the selection, training, and supervision of program staff. Programs are required to work within these critical elements, and use curricula approved by HFA to meet these goals.

Assessing HFNY's adherence to the 12 critical elements of the national model facilitates an understanding of the utility, application, and adaptability of the framework. Importantly, information generated from such a study is instructive in revealing areas where the program has excelled as well as areas where it could be improved. Documenting the extent to which HFNY adhered to the intended program design for the families randomly assigned to the treatment arm of the study also provides a meaningful context for interpreting findings from the outcome study.

The current examination includes only those three sites that are part of the evaluation study and only those women who were randomly assigned to the intervention group and participated in the research study's baseline interview (n=579). Additional data were extracted from the HFNY centralized Management Information System (MIS) for each respondent from their initial assessment through the date of their final attempted or actual visit. The HFNY MIS houses comprehensive information from each site on the characteristics and needs of the families served, the frequency and content of home visits, the nature and outcome of service referrals, progress toward program objectives, and worker demographic and training information. Aggregate and individual program reports are generated on a regular basis to assess program

accountability, to monitor the program as a whole, and to support the quality assurance and improvement efforts employed by individual sites. For the purposes of the current study, we extracted information on the risk factors and needs of the women served, the frequency and content of home visits, the nature and outcome of service referrals, progress toward program objectives, and worker demographics.

The fidelity assessment is presented according to the 12 critical elements, grouped within the three main content areas noted above (participant identification and assessment; service content and intensity; and the selection, training, and supervision of program staff).

Participant Identification & Assessment

Use a standardized assessment tool to systematically identify families. The target population for HFNY services is expectant parents and parents with an infant under three months of age who live in high risk target areas. Communities are considered high risk based on factors such as high rates of teen pregnancy, low birth weight babies, infant mortality, welfare receipt, and late or no prenatal care. Expectant parents and parents are screened for program referral by various collaborative community agencies, such as prenatal care providers and hospitals. Individual HFNY programs also engage in their own outreach efforts to identify and screen prospective families. Families who screen positive are referred to the HFNY program in their community, where they are assessed by trained Family Assessment Workers (FAWs) using the Kempe Family Stress Checklist (1976).

The Kempe is a 10-item standardized, semi-structured inventory designed to assess families' risk of engaging in child abuse or neglect. The FAW uses the instrument's semi-structured format to evaluate a variety of domains, including history of childhood abuse, substance abuse, mental illness or criminality, the presence of life stressors, and attitudes and

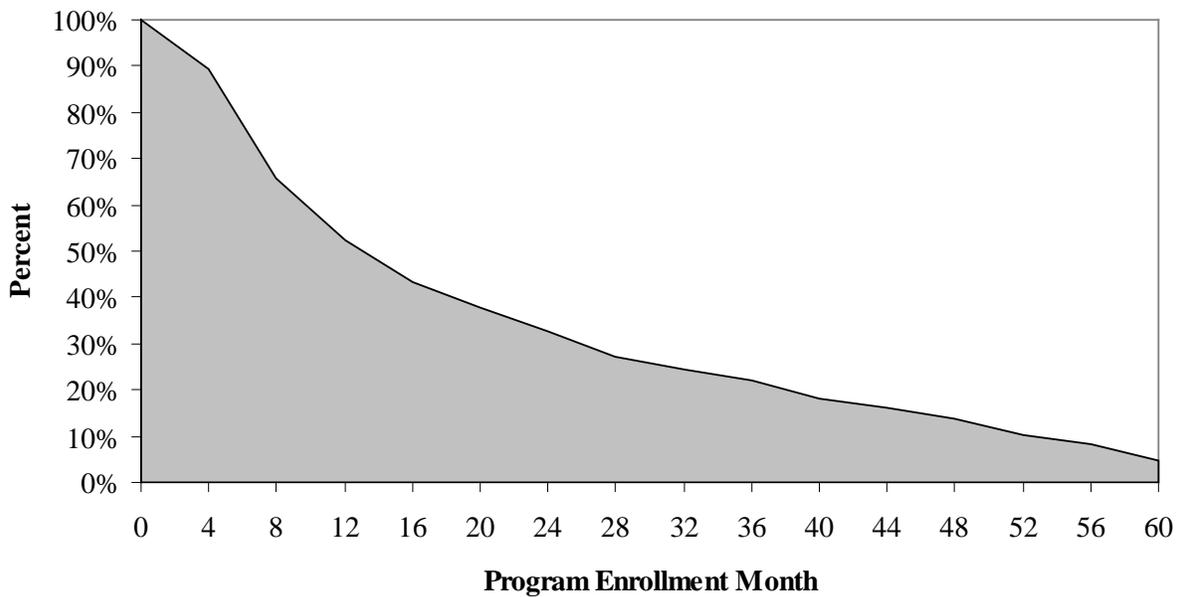
expectations regarding children. The instrument is also used to systematically assess whether or not families are likely to benefit from receiving home visiting services by examining their strengths and needs. Items are scored as being “no problem” (0), “mild problem” (5) or “severe problem” (10). Scores can range from 0 to 100. Families are deemed eligible for home visiting services if either parent receives a score of 25 or higher. Nearly all families referred to HFNY (99%) meet the prescribed cutoff on the Kempe (Mitchell-Herzfeld et al., 2005). The average score for the 579 women assigned to the treatment arm of the study was 42.94 (SD=13.27).

Offer services voluntarily and use positive, persistent outreach efforts to build family trust. HFNY offers a voluntary program that respects the rights of families to decline or disengage from home visiting at will. According to data from the MIS system, following the assessment and determination of eligibility, the Family Support Workers (FSWs) engaged in a variety of activities to develop trusting relationships with families and encourage them to enroll in home visiting services, without compromising the voluntary nature of the program. These activities included phone calls, letters, provision of program materials, and visits. FSWs were encouraged to continue their initial outreach efforts until the target child was three months old.

Initiate services prenatally or at birth. HFNY places great emphasis on the early initiation of services. Consistent with the program’s policy, data reveal that an FAW assessed 78% of the families involved in the program prenatally or within two weeks of the target child’s birth, and all of the first home visits (99%) were conducted in a timely fashion, either prenatally or within 3 months of the target child’s birth. Program staff successfully engaged the majority of respondents assessed; only 10 percent (n=57) of the families assigned to the intervention arm never engaged in or accepted the services offered. Thus, the following sections describe the experiences of the 522 study respondents who accepted and received at least one home visit.

Exhibit 5 depicts the retention rate for all participants. The average length of enrollment in HFNY was 20.68 months (SD=18.47). Just over half (52%) of the participants remained enrolled in the program by one year post-enrollment. By two years post-enrollment, 33% of participants were still receiving home visiting services. Few families were still enrolled at three (22%) and five (4%) years. Part of the drop-off between ages three and five may be due to children’s enrollment in Head Start and Kindergarten, which are natural endpoints for graduation from HFNY. However, only about 16% of families who enrolled in HFNY officially graduated from the program. On average, families who graduated from HFNY had been in the program for about 4.5 years (SD=.73).

Exhibit 5. Retention rate in HFNY for all enrolled participants (n=522)



Service Content & Intensity

Offer services intensely with well-defined criteria for increasing or decreasing intensity of service over the long term. Home visits are scheduled according to service level and typically last for 60 minutes. Most families will begin on either the Prenatal Level or Level 1, depending

on whether they enrolled pre- or postnatally. Prenatal visits are scheduled bimonthly, and use curricula such as “Partners for a Healthy Baby.” Prenatal visits generally focus on promoting healthy behaviors (e.g., eating nutritious food), discouraging risky behaviors (e.g., tobacco and alcohol use), developing ways to cope with stress, encouraging compliance with prenatal appointments, and providing information about the development of the fetus.

Visits increase to weekly on Level 1 once the mother gives birth, and generally continue for about six months. Fewer visits occur as families progress through the remaining levels, from biweekly (Level 2), to monthly (Level 3), to quarterly (Level 4). Two additional service levels are available on an as needed basis, providing greater than weekly visits or contacts in cases of exceptional need (Level 1-Special Services) and creative outreach to attempt to re-engage families who have lost contact (Level X). Families move to more or less intensive levels of service depending on their individual needs (e.g., level of risk, quality of parent-child interaction, family problem-solving skills, family crises, etc.). At all of these levels, FSWs utilize curricula such as "Parents as Teachers" and “Helping Babies Learn” to help instruct, model, and educate parents about the child’s development and the importance of their role.

Exhibit 6 shows the average number of visits for program enrollees throughout the five-year period. The average number of visits over all time periods combined was 33.29 (SD=30.64). The greatest number of visits occurred between birth and age 1 (M=18.71, SD=12.24), and, as prescribed, the remaining averages decreased in intensity as the families’ functioning improved.

Exhibit 6. Average number of visits by time period

		Prenatal	Birth to Age 1	Age 1 to Age 2	Age 2 to Age 3	Age 3 to Age 4	Age 4 to Age 5	All Time Periods Combined
Enrollees	<u>M</u>	5.73	18.71	14.91	11.68	8.75	5.49	33.29
	<u>SD</u>	4.27	12.24	8.96	7.06	6.75	4.90	30.64
	<u>N</u>	332	482	238	145	101	59	522

Exhibit 7 provides a comparison of the number of visits expected to occur on each level with the average number of visits that were completed and attempted. As shown in the exhibit, for most levels, the average number of completed visits neared or surpassed the number of expected visits and satisfied the requirements for the critical element.

Exhibit 7. Comparison of expected visits to actual visits (n=522)

Level	Expected Visits	Completed Visits (Mean)	Attempted Visits (Mean)	Participants Receiving 75% of Expected Visits	Participants Receiving 50% of Expected Visits
Level 1P (n=309)	2 per month	3.00	0.53	79%	94%
Level 1 (n=479)	4 per month	2.61	0.92	29%	70%
Level 1SS (n=9) ¹	4 per month	3.75	0.72	78%	89%
Level 2 (n=210)	2 per month	1.59	0.55	57%	93%
Level 3 (n=131)	1 per month	0.93	0.36	76%	92%
Level 4 (n=74)	1 per quarter	1.92	0.65	96%	96%
Level X (n=264)	1 attempt per month	0.81	2.27	92%	99%

¹The expected visit frequency on Level 1SS (Special Services) is more than one home visit weekly or weekly visits plus other contacts. We were able to evaluate whether one or more visits were attempted or completed for the 9 participants who were assigned to this level; however, we were unable to evaluate whether or not other contacts were attempted or completed. For the sake of comparison to other levels, we have provided values for a minimum of 1 visit per week.

Exhibit 7 also shows the percent of participants who received 75% and 50% of the expected visits. HFNY policy requires that 75% of program participants receive a minimum of 75% of the expected visits congruent with the level of service to which they are assigned. As seen in the table, the program met this goal for participants for five of seven levels. With the exceptions of the prenatal period and Level 1SS, which served only 9 families, a pattern of increasing success with requirements for fewer home visits was observed. On Level 1, however, only 29% of participants received the expected 75% of visits. This finding is consistent with

patterns reported by other HFA programs (Harding et al., 2004). Of the approximately 100 sites surveyed across nine states or Healthy Family America systems, typically 31% of families received 75% of the expected visits per month on Level 1 (Harding et al., 2004). Accordingly, the expectation that workers provide four home visits per month (Level 1) may not be feasible for a number of reasons. Scheduling conflicts with work and school may increase immediately following the birth of a child, some families may not have the time or be motivated to accept such frequent home visits (Williams, Stern & Associates, 2005), and other families may need more time to engage fully as they initiate the program (Harding et al., 2004). Additional research is needed to investigate the possible explanations for the varying degrees of success on the different levels.

Services are culturally sensitive. Cultural context and respect for diversity are incorporated into the HFNY program design. In order for services to be effective, it is important that staff acknowledge, respect, and understand the cultural and socio-demographic backgrounds of those they serve. The program is committed to recruiting staff that are representative of the language and culture of participants. FSWs are typically selected from the communities they serve and are representative of the demographics of HFNY clients.

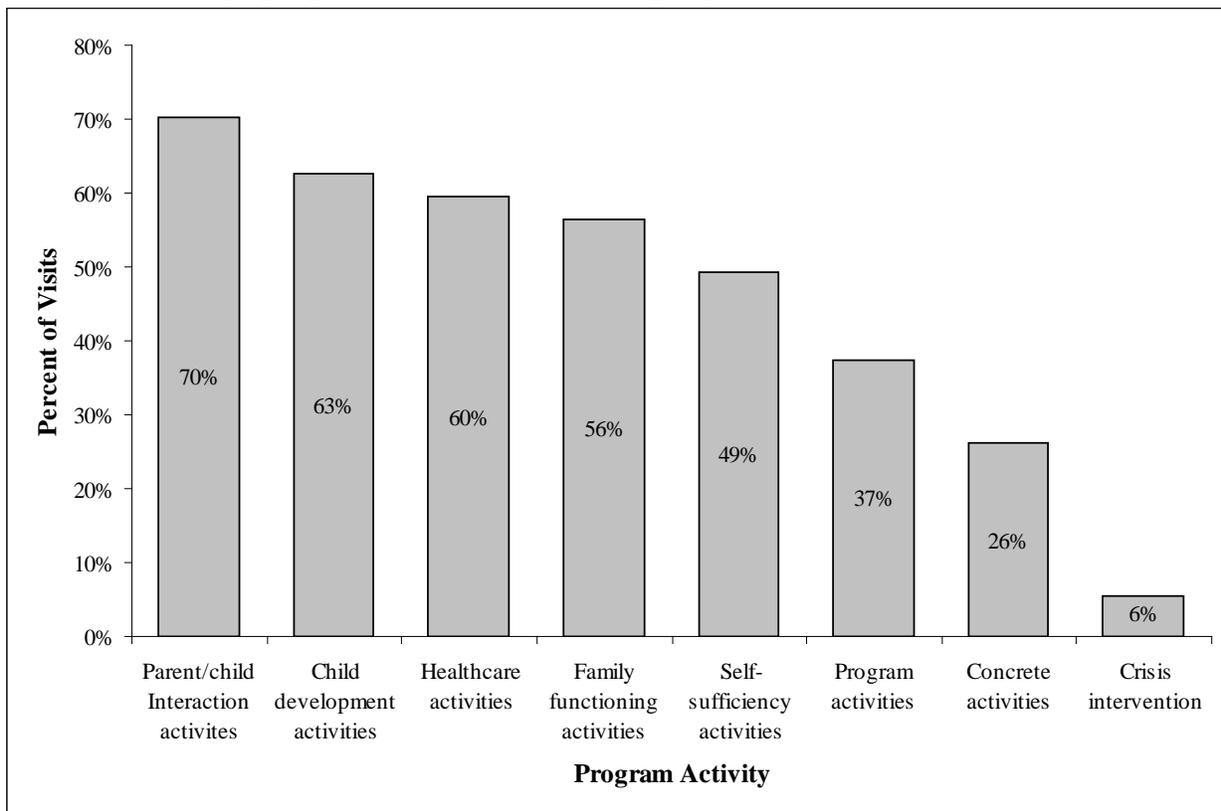
To approximate the percentage of home visitor-participant pairs that were racially/ethnically concordant, we examined data collected in 2002 from 40 home visitors who collectively served 368 of the HFNY RCT study participants. Concordance in terms of race and ethnicity was achieved for about two-thirds of the home visitor-participant pairs. White home visitors matched the race of 68% of the study mothers served, while the race of the home visitor and mother matched at a rate of 62% for both black home visitors and for Hispanic home visitors.

In addition to looking at the matched pairs, we examined self-report data from the Year 1 interview to consider home visitors' sensitivity to the needs of the HFNY mothers participating in the program. Among the 403 program participants interviewed, 82.4% of the respondents felt comfortable talking to their home visitors all or most of the time. Even higher percentages of mothers felt that the home visitors gave good advice (87.3%) and respected participants' ideas about parenting (86.6%). While not a direct assessment of cultural sensitivity, mothers' responses to these questions suggest that there was not only congruence with their home visitor's race/ethnicity, but also a level of comfort, understanding, and appropriateness suggestive of cultural competence.

Supporting parent-child interactions and child development. The FSWs and program participants enrolled in the study engaged in a variety of activities together during home visits. The frequency of engagement in these activities is consistent with the program's goals of supporting positive parent-child interaction, promoting optimal child health and development, enhancing parental self-sufficiency, and preventing child abuse and neglect.

As shown in Exhibit 8, parent-child interaction activities and child development activities occurred most often, followed by healthcare, family functioning, and self-sufficiency activities. Program activities, such as completing forms, group activities, and developing or revising the individualized family service plan, which is updated every six months, were engaged in much less frequently. Based on the MIS records, concrete activities such as providing transportation, serving as a translator, and providing food, clothes, diapers, or household items occurred in about one-quarter of the visits. Crisis intervention activities rarely occurred. Notably, although not presented in the Exhibit, almost all participants (96%) received at least one visit where either parent/child interaction activities and/or child development activities took place.

Exhibit 8. Average proportion of visits in which program activities occurred



Linkages to health and other services. FSWs provide information and referrals as their primary method of facilitating linkages between service providers and families. FSWs may make arrangements for the family, or they can provide the family with the information they need to make the contact themselves. Approximately 80% of families who enrolled in the program had at least one referral for services other than HFNY; the average number of referrals for those families was 10.79 (SD=16.46). On average, families received 5.73 referrals that were arranged (SD=13.72) and 5.07 referrals (SD=6.41) where the worker discussed the service with the family and provided them with the information necessary to contact the provider on their own. Approximately 1.73 (SD=3.03) of these referrals were issued within three months of the Kempe assessment, which was significantly lower ($p<.001$) than the number of referrals issued for the control group (M=2.97, SD=2.72).

Overall, 82% of the families who had a referral received at least one service as a result, with services being received for about 53% of the total number of referrals issued. Exhibit 9 shows the percent of families with at least one referral and the percent of referrals where a service was received, by category of service.

As shown in the exhibit, the most common referral provided was for concrete services, an activity that was not typically performed by the home visitor: 52% of the families had at least one referral for concrete services, and nearly two-thirds (64%) of these referrals resulted in a service being received. About a third of participants were provided referrals to nutrition services; employment, training and education services; or family and social support services. In most service categories, more than one third, and often more than one half of the referrals resulted in receipt of services.

Exhibit 9. Percent of referrals for services provided and received

Referral and Service Type	%
Concrete Services	
Proportion of families with at least one referral	52.5
Proportion of referrals where a service was received	63.8
Nutrition Services	
Proportion of families with at least one referral	35.4
Proportion of referrals where a service was received	59.9
Employment, Training & Education Services	
Proportion of families with at least one referral	33.3
Proportion of referrals where a service was received	39.0
Family & Social Support Services	
Proportion of families with at least one referral	32.6
Proportion of referrals where a service was received	32.6
Health Care Services	
Proportion of families with at least one referral	29.5
Proportion of referrals where a service was received	58.3
Counseling & Support Services	
Proportion of families with at least one referral	19.3
Proportion of referrals where a service was received	26.5
DSS/HRA Services	
Proportion of families with at least one referral	16.7
Proportion of referrals where a service was received	66.4
Other Services	
Proportion of families with at least one referral	31.0
Proportion of referrals where a service was received	43.2

The primary reason for why a service was not received was lack of follow-through on the part of the participant. This occurred more frequently when participants were merely provided with information to make the arrangements themselves rather than having the linkages made for them. This suggests that it is not enough to talk with families about a service and expect that

they will follow through and contact the providers themselves; it is often necessary to arrange the referral if the service is to be obtained.

Caseload management. In accordance with program guidelines, caseloads for each FSW do not exceed 25. However, the maximum caseload size for Level 1 visits is 15, as families at this stage require the most frequent visits (4 per month). HFNY utilizes a weighted caseload management system to determine the caseload size of FSWs serving families at different levels of need. Supervisors monitor case weights with frequent meetings with FSWs.

Staffing, Training & Supervision

Staff recruitment and selection. The HFA model indicates that staff should be recruited based on their skills and their willingness and/or experience in working with diverse populations. These skills include the ability to build trusting relationships and to work effectively with families who may have different beliefs and values. Accordingly, FSWs are most often trained paraprofessionals who live in the communities targeted for services and are representative of the language and culture of the participants being served by the program. More than three quarters of the workers have attended college or received post high school training, and about one third were college graduates.

Basic training. Prior to serving families, all FSWs are required to attend orientation and core trainings designed to develop the skills they need to fulfill their position requirements and meet the goals of the program. The orientation is provided by each individual program. It discusses the program's goals, services, policies and operating procedures as well as the program's relationship with other resources in the community. Additionally, as part of the orientation, FSWs observe at least one home visit. All new HFNY staff members attend a one week core training designed to teach the basic skills needed to perform home visits and

assessments. The training is run by a New York State team of approved HFA trainers from Prevent Child Abuse New York. FSWs receive training in promoting parent-child interaction, child growth and development, strength-based service delivery, communication skills, and providing culturally sensitive and responsive services to families. FSWs practice skills learned in core training with a series of transfer-of-learning exercises, and shadow an experienced home visitor before they are assigned to work with families. FSWs are not permitted to make home visits alone until the training is complete

Intensive training. FSWs also attend centralized formal trainings and local wrap-around trainings to enhance their current knowledge and develop strategies to deal with issues such as family violence, mental health, and child abuse and neglect. During the first year of employment, trainings occur every three months. Ongoing trainings reflect FSWs' individual needs for service provision and interests (e.g., working with fathers, immigrant parents, grandparents, etc.).

Supervision. Each FSW receives a minimum of 1.5 hours of individual supervision per week. The individual sessions are designed to address skills development, professional support, and the quality of the FSW's work. Accordingly, the sessions focus on topics such as the FSW's role in promoting parent-child interaction, strategies of engagement and retention of families, and discussion of techniques and approaches to address specific family situations. If possible, supervisors participate in the first home visit with new participants and may also observe subsequent visits.

Quality assurance. While not a critical element, New York's continuous quality improvement system is noteworthy and typifies the program's consistent and institutionalized efforts to meet the national model's credentialing standards and HFNY's own goals. HFNY

employs a continuous quality improvement system designed to: (1) regularly review program quality; (2) plan for and deliver technical assistance and support to sites; and (3) assess progress toward the stated goals and objectives of the program. The centralized MIS database, which contains comprehensive case-level information from each individual program, is an essential tool in accomplishing these aims. The database provides administrators and program managers with a means of monitoring HFNY programs' compliance with performance targets and service delivery standards and their success in attaining program goals.

In addition, quality assurance visits are conducted by training and staff development specialists biannually to observe assessments, home visits, and supervision practices. Site reviews are conducted annually to monitor compliance with HFA and HFNY standards, with review activities tailored to the individual needs of the programs. Technical assistance and additional support for quality assurance, program compliance, or data management issues are provided upon request or when a specific need is identified. As a result of these strategies, HFNY is recognized as a provider of high quality home visiting services, meeting the nationally established HFA credentialing standards for quality service provision. In 1998, HFNY was the first program in the nation to have all of its sites credentialed by HFA; in 2004 it became one of only four multi-site systems to have completed and received multi-site credentialing.

Discussion

Evaluators and proponents of HFA suggest that the level of success realized by the model is a function of the model itself, a program's implementation of the model, and the family and community contexts in which the programs are implemented (Duggan et al., 2000; DuMont, Mitchell-Herzfeld et al., 2008; Gomby, 2007; Harding et al., 2007). Based on the discussion of the 12 critical elements outlined, HFNY appears to have excelled in identifying families in need

of services and initiating those services in a timely fashion but has struggled in retaining those families throughout the course of the program. While HFNY's retention rate is consistent with other HFA programs (Harding et al., 2004), this does not mean it is acceptable. To significantly affect the lives of at-risk mothers and their infants, it may be necessary for the program to make greater efforts to retain the participants it so effectively recruits.

This fidelity assessment has demonstrated that FSWs are representative of the communities they serve and in touch with the obstacles many of the program participants face in creating and maintaining a stable and healthy environment for themselves and their child/children. Accordingly, the review of program activities indicates that HFNY succeeds in engaging families in activities that support positive behaviors related to parenting, child health and development, family functioning and maternal self-sufficiency. Referrals to outside providers are less frequent for these domains. This may reflect the adaptability of the program to respond to the particular needs of those it serves, the reluctance of participants to accept referrals for family issues of such an intimate nature, or a missed opportunity on the part of home visitors to arrange for additional support.

Overall, the assessment reveals that HFNY is guided by the HFA framework; it adheres to some elements very closely, and allows for growth and change in the adaptability of others.

CHAPTER 5: DOES HFNY PREVENT OR REDUCE CHILD MALTREATMENT?

A critical issue in home visiting research concerns the effectiveness of home visiting in preventing or limiting child maltreatment. As discussed earlier, both HFNY and the HFA model on which it is based strive to promote positive parenting skills and healthy parent-child interactions in order to prevent child abuse and neglect. It is expected that these strategies, along with activities promoting the child's health and development and the families' self-sufficiency will contribute to fewer and less frequent cases of abuse and neglect. The current study examines associations between the HFNY program, levels of positive and negative parenting, and contact with the child welfare system for confirmed instances of abuse and neglect, the initiation of family services to avert a placement, and out-of-home placements.

Analytic Subgroups

As discussed earlier in the report, prior to analyzing data from Year 7, we constructed two policy-relevant analytic subgroups: the Recurrence Reduction Opportunity (RRO) subgroup and the High Prevention Opportunity (HPO) subgroup. The RRO subgroup includes women who were involved in a confirmed report (as a non-victim) within five years prior to random assignment (n=104). The HPO subgroup includes first-time mothers, under the age of 19, who were randomly assigned to the program at a gestational age of 30 weeks or less (n=179).

The limited sample sizes of the two subgroups largely dictated our options for analyses. For the RRO group, assuming 80% power and a .05 level of confidence, effect sizes of .25 were detectable when the rates of maltreatment averaged about 25% or 75%. When confidence standards were relaxed to .10, effects sizes of .20 were detectable with 75% power. While rates of retention were reasonable within the analytic RRO subgroup between the baseline sample and the mom interview (83.7%) and child interviews (65.4%) conducted at Year 7, the reductions in

sample size further restricted our ability to detect meaningful differences with confidence. Consequently, we did not analyze maternal or child reports of parenting for this subgroup.

Our ability to detect differences for the HPO subgroup was slightly better, although effect sizes of .20 were still required to detect differences in rates of abuse and neglect with 95% confidence and 80% power. When confidence standards were relaxed to .10, effects sizes of .15 were detectable with 75% power within the baseline sample, and effect sizes of .20 or higher with the Year 7 follow-up samples. While not ideal, the effect sizes observed for this group at both Year 2 and Year 3 approached these levels. Thus, we decided to conduct analyses for the HPO subgroup for all three samples, with the caveat that our study lacked the power to detect small effects.

Measures of Parenting

Analysis of parenting outcomes rely on data from three sources: CPS reports, maternal self-reports, and child reports of mothers' parenting practices.

Administrative records. As described earlier, information regarding child maltreatment, initiation of family services to avoid a placement, and foster care placement were summarized using OCFS-administered databases. As a result of extraction and coding efforts, information on confirmed reports, tracked services, and foster care placements in NYS were available for all target children in the baseline sample from the time of random assignment up through their seventh birthday. We created a series of variables representing the cumulative rate and total number of reports involving the mother as the confirmed subject and / or the target child as the confirmed maltreated victim. Cumulative rates and totals were created for confirmed reports involving any type of abuse or neglect, confirmed reports involving at least neglect, confirmed reports involving at least physical abuse, and confirmed reports involving at least sexual abuse.

Finally, given HFA's distinct goals regarding parenting, mother-child interactions, and child development, we summarized rates of confirmed reports for mother and child separately, overall, and for each of seven years to investigate possible differences in the program's impact.

Information from the CCRS database was used to create variables representing cumulative rates and numbers of tracked family support services initiated for preventive and protective services or placement and foster care placements from random assignment through the target child's seventh birthday.

Mom Interview. We used the revised parent-child Conflict Tactics Scale (CTS-PC) at Year 7 to measure self-reported parenting practices (Straus, Hamby, Finkelhor, Moore & Runyan, 1998). The scale consists of 27 items that ask about how often each parenting behavior occurred (0–20 times) during the prior year and five items that inquire about negative parenting behavior during the past week. For the current analyses, we used subscales that described parenting behaviors during the last year, including non-violent discipline, psychological aggression, minor physical aggression, serious physical abuse, and neglect. The incidence of very serious physical abuse was too low to analyze. Scores of items grouped within a particular scale were used to indicate the frequency and/or prevalence of a parent's use of specific parenting tactics over the past year. Mothers who did not have custody of the child were ineligible to complete the scale.

At Year 7 we also introduced a new method for collecting sensitive information. The method involved a computer-assisted assessment tool with an audio version of the Conflict Tactics Scale – Parent Child version, which allowed respondents to hear the questions through headphones and to respond privately using a touch screen. Practice items were included so that the interviewer could coach the respondent how to indicate answers, refuse questions, and move

to the question. While this technique helped to substantially reduce levels of missing data, respondents still selected “don’t know” and “refuse” for many of the items, with non-response increasing as the severity of the question increased. In most cases, “don’t know” and “refused” options did not exceed 5%.

To assess the potential for missing data to influence study findings, we constructed and compared four different methods for replacing missing items, including dropping the items, replacing the missing values with 1s or 0s and, where appropriate, the child’s response or a concurrent indicated CPS report. Results included in the exhibits showed a consistent pattern of effects across at least three of the four measures.

Child Interview. Complementing mothers’ reports of parenting practices, the Conflict Tactics Scale-Picture Card Version (Mebert & Straus 2002) was included for children. The CTS-PCV consists of pictures depicting parenting behavior and acts of maltreatment. In the current study, the pictures were accompanied by an audio description of the act being shown and a question asking the child if his/her mother has ever treated him/her the way the mother is treating the child in the picture. Children listened to the questions through headphones and responded privately using a touch screen. While items are typically grouped into 5 subscales: non-violent discipline, psychological aggression, minor physical aggression, serious physical abuse, and very severe physical abuse, due to concerns about exposing children to the more severe pictures, the instrument was restricted to pictures that depict non-violent discipline strategies, psychological aggression, and minor physical aggression. Scores were used to indicate the frequency and/or prevalence of a child’s experience of specific parenting practices.

Analysis Plan: Parenting Outcomes

Prior to analysis of the outcomes, we used Student's *t*-tests and Chi-square tests to assess the comparability of the intervention and control groups for the two analytic subgroups. For tests of treatment effectiveness, we adopted an intention-to-treat approach using all available data, irrespective of a family's participation in the program.

Cumulative summaries of confirmed reports and mother and child reports of parenting practices were analyzed as both dichotomous outcomes and frequency or total count scores. We used generalized models, SAS 9.2, with a binomial distribution and logit function to estimate rates of dichotomized prevalence outcomes. We also used logistic regressions to produce an adjusted odds ratio, which approximate the size of the intervention's effect. In models where the maltreatment or parenting indicators were a count or frequency variable, descriptive statistics were conducted to determine the distributional properties of the dependent variables. All count variables describing administrative indicators, serious physical abuse, and neglect were not normally distributed. More frequently occurring parenting behaviors, such as non-violent discipline, psychological aggression, and minor physical aggression were normally distributed, so a generalized linear model with the identity link function was utilized. Dependent variables involving non-normative distributions were analyzed with a negative binomial distribution and log link function, using generalized linear models, SAS 9.2.

Analyses involving administrative records were run for the sample as a whole and the two subgroups. The remaining analyses were conducted on the whole sample and the HPO subgroup. Covariates were used to maintain equivalence across the two treatment arms and are indicated on each exhibit.

Although not presented in tabular form, we also estimated and plotted the non-cumulative rates for each year between the time between random assignment and the target child's seventh birthday. We used these analyses to assess whether the cumulative rates provided a fair representation of the effects if they had been assessed cross-sectionally.

Results

Description of analytic subgroups. Exhibit 10 shows the baseline characteristics of the RRO and HPO subgroups by treatment arm. While differences within the RRO subgroup exist for women in the treatment and control conditions on receipt of cash assistance, gender, and depressive symptoms, the subsamples are comparable on all other characteristics, including education, presence of partner, age, and count of risk items. The equivalence of the treatment arms within the HPO subgroup is noteworthy, showing similar rates or levels on all baseline characteristics and revealing no significant differences.

It is also notable that despite the many differences between the two subgroups that resulted from the characteristics used to define them, such as age and parity, the women in both subgroups are similar in their counts of moderate to severe risk items and mastery scores.

Exhibit 10: Baseline characteristics of subgroups by subgroup and treatment arm

Characteristic	RRO		HPO			
	Control	HFNY	Control		HFNY	
	(n=52)	(n=52)	(n=88)	(n=91)		
	Percentage		p	Percentage		p
Mother's race/ethnicity						
White, non-Latina	25.0	42.3		23.9	24.2	
African-American, non-Latina	61.5	46.2	.17	52.3	50.5	.40
Latina	13.5	11.5		17.0	23.1	
Mother < 19 years old	1.9	5.8	.31	100	100	--
First-time mother	3.8	7.7	.40	100	100	--
At least high school diploma or equivalent	50.0	55.8	.56	9.1	13.2	.39
Had partner	69.2	64.7	.63	62.5	70.3	.27
Moved in past 12 months	69.2	53.8	.11	59.1	48.4	.15
Cash assistance at random assignment	78.8	61.5	.05	26.1	22.0	.52
Pregnant at random assignment	69.2	64.7	.63	100	100	--
Target child female	51.9	34.6	.08	50.0	46.2	.61
	Mean (sd)		p	Mean (sd)		p
Mean maternal age in years	26.96	26.65	.78	17.20	17.31	.54
	(5.03)	(6.35)		(1.46)	(1.01)	
Total number other biological children	2.67	2.33	.30	0	0	--
	(1.37)	(1.99)				
Depressive symptoms (CESD)	16.54	20.65	.08	16.16	16.20	.98
	(10.29)	(12.90)		(10.31)	(11.07)	
Total mastery score	21.33	20.53	.19	20.47	20.25	.61
	(3.01)	(3.11)		(2.83)	(2.87)	
Total maternal parenting attitudes (AAPI)	134.19	138.21	.15	131.94	129.16	.22
	(14.64)	(13.58)		(15.66)	(14.61)	
Count of risk items (Kempe)	6.04	5.96	.75	5.92	6.09	.41
	(1.27)	(1.22)		(1.39)	(1.32)	
Estimated annual earnings (\$)	1662.92	2290.60	.51	1381.64	1576.72	.31
	(4577.64)	(5013.43)		(3111.07)	(3249.59)	

Involvement with child protective services, preventive services, and foster care. Exhibit 11 displays the rates, counts, and adjusted odds ratios and effect sizes for HNFY as it relates to indicators of child maltreatment, preventive services, and placements outside of the home. Estimates are shown for the sample of women assessed at baseline as a whole and separately for the RRO and HPO subgroups.

Sample as a whole. No program effects were observed for the cumulative prevalence or cumulative number of confirmed reports for the sample as a whole. Five cases of sexual abuse were confirmed between the time of random assignment and the target child's seventh birthday; all five reports involved mothers from the control group as the confirmed subject ($p < .05$).

Although not shown, estimates of the non-cumulative annual rates also revealed no marked differences between the intervention and control groups. Annual rates for women and children in both treatment conditions were relatively similar, and similar to rates documented in incident studies, reaching their highest rates during the target child's first year of life.

To investigate the possibility that a reporting bias influenced the findings, we examined whether mothers in the HFNY group who self-reported serious abuse and neglect during the Year 1 interview, when one would expect the bias to be the greatest, were more likely to have a CPS report, whether unfounded or substantiated, than mothers in the control group who self-reported serious abuse or neglect. For the sample as a whole, 42.9% of HFNY mothers who self-reported serious abuse and neglect had a CPS report as compared to 22.2% of mothers who self-reported serious abuse and neglect in the control group ($p < .05$). These results suggest that mothers assigned to HFNY were more likely to be detected for child maltreatment than mothers assigned to the control group.

Women in the RRO subgroup. HFNY produced unexpected and unprecedented differences in rates of subsequent confirmed reports for HFNY mothers in the RRO subgroup.

As compared to their counterparts in the control group, HFNY mothers had

- lower rates of confirmed CPS reports for any type of abuse or neglect:
 - (41.5% versus 60.4%, $p < .10$);
- lower rates of reports when the study mother was the confirmed subject:
 - (38.2% versus 57.4%, $p < .10$)
- lower rates of confirmed reports involving physical abuse:
 - (3.3% versus 13.4%, $p < .10$)
- a smaller number of total confirmed reports for mothers as the confirmed subject
 - (.8 versus 1.6, $p < .05$)
- lower rates of preventive, protective, and placement services initiating:
 - (38.02 versus 60.02, $p < .05$)

Exhibit 11. Administrative indicators of child maltreatment and foster care from RA to target child’s seventh birthday

Administrative Indicator of Child Maltreatment or Foster Care	Whole Sample ¹ (n=1173)			RRO Subgroup ² (n=104)			HPO Subgroup ³ (n=179)		
	Control (n=594)	HFNY (n=579)	AOR ⁴	Control (n=52)	HFNY (n=52)	AOR ⁴	Control (n=88)	HFNY (n=91)	AOR ⁴
Cumulative Rate	Percentage		AOR ⁴	Percentage		AOR ⁴	Percentage		AOR ⁴
Bio mom OR target child confirmed subject or victim of CPS report	27.10	29.55	1.13	60.36	41.51	.47†	25.03	21.92	.84
Bio mom confirmed subject – N	20.68	22.96	1.14	57.41	38.18	.46†	16.90	19.00	1.15
Target child confirmed victim - N	22.95	24.29	1.08	47.69	36.42	.63	22.64	18.70	.79
Bio mom confirmed – PA	4.24	4.47	1.06	13.44	3.25	.22†	4.47	1.08	.23
Target child confirmed victim – PA	3.05	4.09	1.36	8.81	6.71	.75	3.01	1.05	.34
Bio mom confirmed – SA ⁵	0.7	0.0	.00*	3.8	0.0	.00	0.0	0.0	--
Target child confirmed victim – SA ⁵	0.7	0.2	.26	1.9	1.9	1.00	2.3	0.0	.00
Child welfare services track initiated	18.16	16.21	.87	60.02	38.02	.41*	12.3	8.67	.68
Foster care placement – target child	4.90	4.83	.99	23.62	17.19	.67	1.76	2.88	1.47 ⁵
Cumulative Number	LS Mean		ES ⁶	LS Mean		ES ⁶	LS Mean		ES ⁶
Bio mom OR target child confirmed subject or victim of CPS report	.55	.54	-.01	1.63	.96	-.35†	.49	.31	-.19
Bio mom confirmed subject – CAN	.42	.42	.00	1.59	.79	-.44*	.24	.25	.02
Target child confirmed victim - CAN	.41	.39	-.02	1.00	.66	-.24	.41	.24	-.21

¹ Analyses control for female target child, count of moderate to severe Kempe items, annual earnings at random assignment, and having at least a GED or high school diploma. ² Analyses control for female target child, being white, cash assistance at random assignment, depressive symptoms at baseline, and at least one move in the past year. ³ Analyses control for female target child and at least one move in the past year. ⁴ Adjusted odds ratio. ⁵ Unadjusted percentage; incidence too low to reliable estimate in multivariate model (n=5). ⁶ Effect size. † p< .10 * p< .05

Given the magnitude of the effects and the potential importance of the findings, we conducted post-hoc analyses to examine several factors that might account for the association between the intervention and reduced rates of subsequent reports. After reviewing related literature and considering the program's impact at earlier waves, we considered five possible intervening variables: changes in parenting attitudes from baseline to ages one and two; more appropriate limit-setting at age two; positive and negative parenting behaviors observed at Year 3; and, given its role in research on risks for recurrence and the home visiting literature, the number of subsequent children to whom the respondent gave birth since the baseline interview (c.f., Bae, Solomon & Gelles, 2008).

Using logistic regression analyses, we examined a model identical to the one used to determine the relationship between the program and confirmed reports involving either mother or child. We then evaluated the degree to which each of the candidate mechanisms attenuated the treatment effect when entered in the model individually. When considering parenting attitudes and number of subsequent pregnancies and children, we controlled for baseline values for parenting attitudes and number of other children to capture the change since that time. The odds ratios were remarkably unaffected by the earlier parenting indicators, with only the number of subsequent children and, more specifically, births between baseline and the Year 2 follow-up, reducing the magnitude of the program's effect by 35%. Thus, consistent with mechanisms proposed by Olds and colleagues (1988), fewer subsequent children are associated with lower rates and fewer subsequent reports for the RRO subgroup.

To further validate this relationship, we examined the correlations between home visiting activities and the cumulative rate of CPS reports by Year 7. An interesting pattern emerged: prenatal self-sufficiency activities and prenatal health-related activities, including family

planning, were both inversely related to the cumulative rate of reports, with correlations equal to about -.12. During the prenatal period, program activities such as intensive family service plans were the highest correlates of confirmed reports ($r = -.15$). Given the strong link between service content and the outcome of interest, together with the identification of a partial mediating mechanism, we are confident that the effect observed is robust, and that delayed or averted births played some role in achieving the reductions in subsequent CPS reports observed.

Finally, to investigate the possibility that a reporting bias influenced the findings, we examined whether a program effect due to increased surveillance of mothers in the home visiting program may be reversed for mothers in the RRO subgroup because they are already known to the system. Specifically, we evaluated whether RRO mothers in the HFNY group who self-reported serious abuse and neglect during the Year 1 interview were more likely to have a CPS report, whether unfounded or substantiated, than mothers in the control group who self-reported serious abuse or neglect. Within the RRO subgroup, 71.4% of HFNY mothers who self-reported serious abuse and neglect had a CPS report as compared to 50.0% of mothers who self-reported serious abuse and neglect in the control group. These results suggest that RRO mothers assigned to HFNY were more likely to be detected for child maltreatment than RRO mothers assigned to the control group, minimizing the likelihood that the lower rates observed for the HFNY mothers were due to a reporting bias such as heightened vigilance for mothers in the control group or underreporting of HFNY mothers because they were involved in a known parenting program.

Women in the HPO subgroup. Consistent with analyses conducted at Year 2, no significant differences were detected in the rates or number of cumulative reports for women within the HPO subgroup.

Analyses of non-cumulative annual rates revealed a somewhat different pattern than the analysis of cumulative rates, and as such, are worthy of mention. HFNY Mothers in the HPO subgroup initially experienced higher rates of abuse and neglect during the first year of the program than participants in the control group, although not statistically different. However, rates for these participants were consistently low in subsequent years and followed a fairly consistent pattern across the two treatment arms. In contrast, rates of confirmed reports for HFNY children declined after the first year and remained fairly constant for the duration, while rates for children in the control group rose at two years of age, declined until age four, and then rose steadily as children entered the school age years. This pattern was so pronounced we conducted post-hoc analyses to test the association between treatment group assignment and the cumulative rate of abuse or neglect during the child's fifth to seventh years of life. Similar to results reported by Zielinski, Eckenrode and Olds (2009), we found that rates during this later period were significantly lower for the HFNY participants (9.9%) than for those assigned to the control group (19.3%; AOR .46, $p=.08$). Thus, as children move out of the home and into more structured and public settings such as school, rates of detection may begin to equalize.

Non-violent, abusive and neglectful parenting. Once the analysis of administrative indicators was complete, we evaluated the effectiveness of the program in promoting positive parenting and preventing negative parenting behaviors that may or may not reach the level of severity of confirmed cases of abuse and neglect.

Exhibit 12 shows the effect of HFNY on indicators of non-violent, abusive, and neglectful parenting as reported by study mothers. The top half of the table presents effects on rates or the prevalence of parenting behaviors and the lower half displays differences between the treatment and control group on the frequency of parenting behaviors. In a similar manner,

Exhibit 13 shows the effect of HFNY on a parallel set of indicators of non-violent and coercive parenting, as reported by study children.

For the sample as a whole, reports from both mothers and their children reveal several significant program effects.

- Consistent with findings from the trial at Year 2, HFNY mothers used serious physical abuse less frequently (.03 versus .15, $p < .01$) than mothers in the control group, and used non-violent discipline strategies more frequently (49.27 versus 45.27, $p < .05$).
- Target children also reported lower rates of minor physical aggression for HFNY mothers (70.8% versus 77.2%, $p < .05$); but differences were not found on their reports of mother's non-violent discipline practices.
- No program effects were observed for prevalence (whether an event occurred) of neglect, although the odds ratio and frequencies followed the pattern identified at earlier waves.

Exhibit 12. Year 7 indicators of non-violent, abusive and neglectful parenting by treatment arm: Mother interview

Non-violent, abusive, and neglectful parenting	Whole Sample ¹			HPO Subgroup ²		
	(n=897)		AOR ³	(n=132)		AOR ³
	Control (n=445)	HFNY (n=452)		Control (n=63)	HFNY (n=69)	
Prevalence	Percentage			Percentage		AOR ³
Non-violent discipline ⁴	98.6	100.0	*	100.0	100.0	ns
Psychological aggression	86.49	87.92	1.18	91.19	79.74	.38†
Minor physical aggression	59.17	64.12	1.25	65.58	64.79	.92
Serious physical abuse	3.18	1.76	.55	3.40	3.20	.ns ⁵
Neglect	16.74	15.77	.93	12.53	17.07	1.39
Frequency	LS Mean		ES ⁵	LS Mean		ES ⁵
Non-violent discipline	45.27	49.27	.14*	45.14	43.30	-.06
Psychological aggression	15.21	15.33	.01	12.99	9.93	-.23
Minor physical aggression	4.51	4.36	-.02	5.47	3.10	-.34†
Serious physical abuse	.15	.03	-.20**	--	--	--
Neglect	.64	.53	.05	.28	.27	.01

¹ Analyses control for annual earnings at random assignment and being white. ² Analyses control for female target child, being black, count of moderate to severe Kempe items, and respondent's age at baseline. ³ Adjusted odds ratio. ⁴ Unadjusted percentages; variability too limited to reliably estimate in multivariate model, effect estimated with chi square statistic. ⁵ Effect size.

† p ≤ .10 * p ≤ .05 ** p ≤ .01 *** p ≤ .001

Exhibit 13. Comparison of indicators of non-violent and abusive parenting: Target child survey – Year 7

Parenting	Whole Sample ¹			HPO Subgroup ²		
	(n=793)		AOR ⁴	(n=122)		AOR ⁴
	Control (n=388)	HFNY (n=405)		Control (n=60)	HFNY (n=62)	
Prevalence	Percentage		AOR⁴	Percentage		AOR⁴
Non-violent discipline (CTS)	96.90	97.80	1.33	96.70	95.20	ns ⁵
Psychological aggression (CTS)	85.14	84.47	1.00	87.13	84.93	.87
Minor physical aggression (CTS)	77.23	70.79	.74*	81.57	75.70	.79
Frequency	LS Mean		ES⁶	LS Mean		ES⁶
Non-violent discipline (CTS)	4.02	4.03	.01	3.93	3.60	-.20
Psychological aggression (CTS)	2.68	2.78	.05	2.92	3.00	.03
Minor physical aggression (CTS)	2.35	2.27	-.04	2.59	2.63	.02

¹ Analyses control for being black, annual earnings at random assignment, having at least a GED or high school diploma, and count of respondent’s other biological children at baseline. ² Analyses control for being Hispanic, annual earnings at random assignment, cash assistance at random assignment, respondent’s age at baseline, and at least one move in the past year. ⁴ Adjusted odds ratio. ⁵ Unadjusted or raw percentages; limited variability did not generate a reliable estimate in multivariate model, effect estimated with chi square statistic. ⁶ Effect size.

† p ≤ .10 * p ≤ .05 ** p ≤ .01 *** p ≤ .001

Next, we examined the relationship between treatment group assignment and indicators from the two CTS measures for women in the HPO subgroup (i.e., first-time mothers under age 19 who were randomly assigned at a gestational age of 30 weeks or less).

- HFNY mothers in the HPO subgroup were less likely to engage in psychological aggression (79.7% versus 91.2%, $p < .10$) and less frequently used minor physical aggression tactics (3.7 versus 5.5, $p < .10$) than their counterparts in the control group. These findings are also consistent with results found in years 2 and 3, using two different methods of assessment.
- Also similar to results from earlier waves, no differences were noted for reports of neglect.
- Additionally, no differences were observed between the two treatment arms for the HPO subgroup within the child sample, although the direction of differences was similar for the prevalence of both psychological aggression and minor physical aggression.

Discussion: Parenting Outcomes.

In the current study, data from multiple sources suggest that HFNY may lead to reductions in several types of abusive and neglectful parenting practices during the first seven years of life. Effects on confirmed reports of child abuse and neglect were concentrated in an important subgroup of women with prior CPS involvement. In contrast, maternal-reported strategies, including the increased use of positive parenting strategies and lower levels of serious physical abuse were sustained from earlier waves and detected for the sample overall. The HFNY women within the HPO subgroup also reported lower levels of minor physical aggression and psychological aggression than their counterparts in the control group, providing evidence for another sustained effect on parenting.

Consistent with earlier findings from the current study, and with reports from the randomized trial of Alaska's HFA program (Duggan et al., 2004; Duggan et al., 2007), no significant differences were detected for the sample overall between the control and HFNY groups for indicators of substantiated CPS reports of abuse and neglect by Year 7. In contrast, results observed for maternal reports of physical parenting provide evidence of a sustained effect on serious physical abuse. At Year 1, compared to mothers in the control group, mothers in the HFNY intervention group reported having engaged in significantly fewer acts of very serious physical abuse, minor physical aggression, and psychological aggression in the past year, and harsh parenting in the past week. At Year 2, HFNY parents reported having committed, on average, one-fourth as many acts of serious physical abuse in the past year (e.g., hitting child with fist, kicking child, slapping on face) than the control group. At Year 7, we again find that HFNY mothers report committing fewer acts of serious physical abuse than mothers in the control group.

The inconsistent pattern of results between the lack of an effect on CPS reports for the sample as a whole and the effect on maternal reports of serious physical abuse may be due in part to increased detection of the HFNY group. That is, many of those self-reporting abusive behaviors, also have a CPS report, while those that don't have a report appear to be engaging in serious behaviors but also avoiding detection. Data from the child interview further support this suggestion, as the children surveyed also reported that HFNY mothers engaged in a lower rate of physical aggression.

A second explanation is that the inconsistent findings are due to non-response bias. To investigate this possibility, we compared the prevalence rates on the child welfare indicators for those consenting to the Year 7 interview and those not interviewed. As shown in Exhibit 14, the

pattern of results suggests that the rates of those consenting to participate at Year 7 are not significantly different from the rates of those who were not interviewed, suggesting that the inconsistency across measures is not simply due to non-response bias.

Exhibit 14. Prevalence rates on child welfare indicators for those consenting to the Year 7 interviews and those not interviewed

Administrative Indicator of Child Maltreatment or Foster Care	Consented Y7 (n=942)	Not Interviewed (n=231)	
Cumulative Rate	Percentage		p-value ⁴
Bio mom OR target child confirmed subject or victim of CPS report	29.34	24.89	.20
Foster care placement – target child	4.67	6.13	.37

We also examined the interaction between interview status (consenter versus not interviewed) by treatment arm assignment (HFNY versus control). As reported above, those interviewed at Year 7 have slightly higher rates for confirmed CPS reports than those not interviewed at Year 7; however, this pattern was more pronounced in the control group than the HFNY group, although not statistically significant. Although these differences could potentially contribute to the pattern of results observed for maternal reports, they are offset by the controls used in the analyses to minimize differences between the treatment arms. In addition, regardless of treatment arm assignment, involvement with foster care was lower for those who consented to participate at Year 7 compared to those not interviewed. Thus, the sustained effect on self-reported physical abuse for the whole sample suggests that the program does impact child abuse, but the effect is not strong enough to counteract the surveillance bias in the early years

With respect to the RRO subgroup, we consistently found an unprecedented pattern of statistically significant trends and differences for both rates and levels of confirmed reports for women within the RRO subgroup, such that HFNY mothers in the RRO subgroup had

substantially lower odds and rates of subsequent confirmed reports as compared to a similar group of women in the control group. While we expected that the risks presented by these women would interfere with the effectiveness of the program, it appears that early, intensive, and enduring home visiting services provided the stimulus needed to induce advantages and instigate parenting choices that otherwise would not be realized. Furthermore, we were able to isolate some of the factors accounting for the program's success, including the content of the home visits and the number of subsequent children born to the respondent. These findings are particularly noteworthy given the well-documented difficulty associated with preventing recurrence, and the well-established risk of number of children in the home (Bae et al., 2008). Considering that 55% of the families included in the RRO subgroup had a report within the three months prior to random assignment, strengthening ties between local departments of social services and the HFNY program has promising implications.

The program also resulted in a significantly lower rate of child welfare services being initiated to avoid a possible placement for the RRO group. Lower odds were also noted for tracked services with regards to the sample as a whole and the HPO group, although to a lesser degree. Thus, the low rate raises questions about whether in-home programs help families circumvent the need for additional preventive or protective services. Conversely, families in the control group may only seek services or be detected when they are faced with an emerging and unsupported crisis that may instigate the initiation of services to avert a potential foster care placement. These are important areas for future investigation.

It is also noteworthy that the program demonstrated gains for women in the HPO subgroup, with significant differences emerging between rates for the HFNY and control groups between the ages of five and seven years. Although involvement in the child welfare and

protective services systems provides a lens for viewing some of the most severe cases of parenting and childhood experiences, it is also important to consider negative parenting behaviors that may not rise to the level of a CPS report, but can never the less have a deleterious impact on the child. Within the HPO subgroup, results from the current study suggest a sustained program impact on coercive parenting, including harsh, psychological and minor physical aggression, which was first detected at the target child's second birthday (DuMont et al., 2008) and observed again at age three (Rodriguez et al., 2010). At Year 7, both mothers' reports of any use of psychological aggression and frequency of use of minor physical aggression followed the pattern of effects observed at earlier waves. Given the magnitude of the earlier effects and its persistence, we expect that the program will also impact outcomes for the target children; although the small sample size and restricted power make effects of a smaller magnitude difficult to assess.

In addition to considering effects on negative parenting, it is also instructive to consider the parenting behaviors that the home visitor has targeted to nurture, instruct, model, and support. The study's effect on non-violent discipline for the sample as a whole is consistent with early results from the first three years of the HFNY randomized trial (Mitchell-Herzfeld et al., 2005; DuMont et al., 2005; Rodriguez et al., 2010). Of particular interest are results observed at Year 3, when we introduced a protocol to obtain videotaped observations of parent-child interactions in four situations that impose a different set of demands on the mother and the child (Rodriguez et al., 2010). In all situations, mothers assigned to the HFNY intervention group were more likely to display at least one positive parenting strategy than mothers in the control group. Also similar to Year 3, the rates documented for positive parenting at Year 3 (Rodriguez et al., 2010) reveal that the majority of mothers and children surveyed at Year 7 reported use of

at least one non-violent parenting strategy when faced with a stressor. Despite the presence of rates in excess of 96% in both groups, mothers in the intervention group were significantly more likely to use this strategy and to use it more frequently. Using alternative strategies to deal with challenging situations may increase the likelihood of a smooth transition to formal schooling (Fagot & Gauvain, 1997; Smith, Landry & Swank, 2000), which, in turn, may decrease the child's risk for severe, negative long-term outcomes such as delinquency, school dropout, and illiteracy.

CHAPTER 6: DOES HFNY LIMIT PRECURSORS TO DELINQUENCY?

Research suggests that the earlier the onset and the greater the severity and persistence of problem behaviors during childhood, the higher the probability that an individual will exhibit deviant behaviors in the future (Farrington, 2005; Moffitt, 1993; Loeber, Green & Lahey, 2003; Tremblay, Japel, Perusse, Boivin, Zoccolillo, Montplaisir & McDuff, 1999). This chapter presents the intervention's association with variables or indices representing risks for poor school outcomes, problem behaviors, socio-emotional difficulties, and self-regulation.

Analytic subgroup

As discussed earlier in the section on abuse and neglect, due to concerns about sample size and group representativeness, analysis of data from the Year 7 interviews was only appropriate for the HPO subgroup. While the effect of the program on parenting within the HPO subgroup was known to approximate a medium effect size, its impact on child outcomes is not known. Therefore, while we decided to continue our study of relationships between the program and precursors to delinquency within the HPO subgroup, we did this cautiously.

Primary Measures: Precursors to Delinquency

Measures related to later delinquency were selected from both the interview with the study mother and the child interview.

Mom Interview. Mothers were asked a series of questions regarding the target child's school experiences, including an assessment of their involvement in gifted and talented programs, special education services, and remedial services for math and writing, repeating a grade, and skipping school or playing hooky more than once. Each experience was evaluated as present (1) or not (0).

Mothers also completed the Child Behavior Checklist for Ages 6-18 (CBCL/6-18: Achenbach & Rescorla, 2001) to assess children's emotional and behavioral problems. The scale consists of 112 structured items and a single open-ended item rated according to how frequently each problem occurs. As in other studies of home visiting, we present results for several of the syndrome subscales (Achenbach & Rescorla, 2001). We selected the five subscales that aligned most strongly with the constructs of interest, including attention problems as a risk for poor school outcomes, rule breaking and aggressive behaviors as an assessment of problem behaviors, and social problems, and the anxious-depressed and withdrawn-depressed syndrome scales as indicators of socio-emotional difficulties. We used raw scores in all analyses.

Child Interview. To provide a measure of receptive language skills, children were administered the Peabody Picture Vocabulary Test 4th Edition (PPVT-IV: Dunn & Dunn 2007). Raw scores are converted to standard scores. To better describe children who may be at increased risk for delinquency due to language or cognitive difficulties, we also report the percent of children who fall below the standardized average of 100; similar approaches have been used by other evaluators (c.f., Love et al., 2005).

The *Loneliness and Social Dissatisfaction Questionnaire* (Cassidy & Asher 1992) was used to provide an assessment of children's self-reported feelings of loneliness and social dissatisfaction, particularly with peers at school. Questions are answered on a 3-point scale where no=1, sometimes=2, and yes=3. In the HFNY evaluation, responses were scored on a scale from 0 to 2. Responses are summed for a total score, with higher scores reflecting a greater degree of loneliness and dissatisfaction.

To provide an assessment of children's self-reported anti-social tendencies, target children were asked select questions that were adapted from the Seattle Social Development Project (Hawkins, 2003) and the Dominic-R (Valla, Bergeron & Smolla, 2000; Valla et al., 1994). The questions tap antisocial behaviors such as stealing, cheating, and fighting, and relational aggression such as social exclusion and retaliation. Based on factor analysis, two scales were created, one representing *bullying activities* and the other *deviant activities*. One item, child's report of "often skipping school" did not load well on any factor and was therefore analyzed independently. Consistent with analysis of parenting outcomes, dichotomous and count variables were created for the two subscales.

An automated *Delay of Gratification Task* was developed for the project using the laptop to obtain a direct behavioral assessment of self-regulatory processes, including three related components: inhibition, attentional strategies, and delay of gratification (Rodriguez, project documentation). Delay-of-gratification tasks consistently show associations with social and academic outcomes (Mischel, Shoda & Rodriguez, 1989). Children were presented with a series of choices and snacks and then asked to wait to eat the snack until the interviewer finished with her paper work. The snack was placed within reach of the child. The child was allowed to distract him or herself using one of two predesigned computer screens. While the child waited, the computer recorded the elapsed time as well as a count of the number of times the child used an electronic pencil to touch either an innocuous picture on the screen's tablet, a fish, or a picture that featured the snack (temptation). The task was automatically terminated if the child visited the temptation screen two times in a row. Variables for analysis included a summary of the number of visits to the distractor screen, a dichotomous variable representing resistance of the temptation screen (i.e., zero clicks), and a dummy variable indicating whether or not the child

waited the full task time (i.e., eight minutes). On average, children pushed the fish picture 33.9 times (SD 24.0), with 18.8% completely resisting touching the temptation picture, and 56.3% of the children waiting the full eight minutes. Analyses of the assessment's relationship with other variables suggest that the task operates as expected, correlating positively with scores from the Peabody Picture Vocabulary Test and several other child outcomes.

Analysis Plan: Precursors to Delinquency

We first ran descriptive statistics to document the experiences and challenges children in the sample were facing. Next, we utilized all available interview data, irrespective of the families' participation in the program to analyze maternal and child assessments of child outcomes. Generalized linear models were used to estimate the effects of the program on a variety of self-report, perceived, and direct assessments of children. Dichotomous outcomes such as the percent participating in special education or gifted programs, self-reported deviant activities, or scoring below the standardized mean on the Peabody were estimated with a binomial distribution and logit function. Similar to the approach with parenting outcomes, we also used logistic regressions to estimate adjusted odds ratios for these variables. All CBCL syndrome scales involved non-normative distributions and were analyzed with a negative binomial distribution and log link function using generalized linear models, SAS 9.2. We used a generalized linear model with the identity link function for the analysis of the distractor count variable, which was appropriate for normally distributed data. The above analyses were run for the sample as a whole and again for the HPO subgroup. Analyses also included sample and subgroup specific covariates to help maintain equivalence across the two treatment arms. Given differences in the distributions on several of the outcomes for boys and girls, gender was also

included as a covariate in all models. Covariates specific to each analysis are indicated on the corresponding exhibits.

Findings: Precursors to Delinquency

Descriptive statistics. On average, children were 7.45 (SD = .39) years of age when the child interview was conducted; 80% of the children were age seven, 10.4% were age six, and 8.5% were eight years or older. Slightly more than half (52.8%) of the children were attending 2nd grade and another 39.9% were in 1st grade. The children's ages were similar across the two treatment arms.

Reports from mothers and children, as well as results from the direct behavioral assessment revealed that many of the children had experiences or behaviors associated with delinquency. Based on mothers' reports, 15.4% of the children were receiving special education services, 31.3% received remedial support for math or reading, and 12.3% had already repeated a grade. Almost half (44%) of the children scored below the standardized mean of 100 on the PPVT. According to the children, nearly 5% of the children skipped school often; 31.5% reported engaging in deviant activities such as cheating often, stealing out of someone's hand, and fighting; and 37.5% reported at least one bullying activity. For comparison, Bauer and colleagues (2006) found that 34.6% of six to nine year old children from a community sample engaged in some bullying during the past year.

HFNY and precursors to delinquency. Exhibits 15 and 16 display results for the whole sample and HPO subgroup using generalized linear and logistic regression analyses to examine HFNY's impact on the precursors to delinquency. Effects on precursors summarized from interviews with mothers are presented in Exhibit 15, and Exhibit 16 presents impacts on indicators derived from the child interview.

With respect to risks for poor school outcomes, a significantly higher percentage of children from the HFNY were reported to participate in gifted programs as compared to children in the control group (AOR: 2.80, $p < .01$). Conversely, compared to children in the control group, fewer children in the HFNY group were receiving special education services (AOR: .70, $p < .10$) or self-reported skipping school (AOR: .35, $p < .01$); however, this latter finding was not supported by maternal reports (AOR: 1.08, ns). As hypothesized, results suggest that HFNY limited the emergence of several outcomes documented to contribute to poor academic performance, school drop-out, and truancy (Pagani, Tremblay, Vitaro, Boulerice, & McDuff, 2001).

With respect to problem behaviors, socio-emotional difficulties, and self-regulation, no significant differences were detected, suggesting that the program's impacts for children were limited to the verbal abilities and experiences at school.

Next, we examined whether the program was more effective for women classified as part of the HPO subgroup. Similar to the pattern observed for the sample as a whole, the program prevented risks associated with poor school outcomes. Notably, HFNY children in the HPO subgroup were less likely to score below average on the PPVT-IV (AOR: .43, $p < .05$); less likely to repeat a grade (AOR: .45, $p < .10$), and more like to participate in a gifted program (5.8% versus 0%, $p < .10$). Similar to results for the sample overall, no differences were present for the other domains of child functioning, although the direction of effects were in the desired direction. Unfortunately, the small sample size of the subgroup limited the study's power to detect differences for small effect sizes.

Exhibit 15. Comparison of indicators of target school engagement, behavior, and mental health at age seven – maternal survey

	Whole Sample ¹			HPO Subgroup ²		
	(n=897)		AOR ³ or ES ⁴	(n=132)		AOR ³ or ES ⁴
	Control (n=445)	HFNY (n=452)		Control (n=63)	HFNY (n=69)	
Risks for poor school outcomes	Percentage/ LS Mean		AOR³ or ES⁴	Percentage/ LS Mean		AOR³ or ES⁴
% participating in a gifted program	1.99%	5.38%	2.80**	0	5.8%	† ⁵
% receiving remedial services	33.31%	32.83%	.98	32.25 %	27.31 %	.73
% receiving special education	16.74%	12.33%	.70†	14.95 %	11.41 %	.78
% repeating a grade	12.60%	12.33%	.98	23.94 %	12.40 %	.45†
% skipping school more than once	2.20%	2.40%	1.08	2.20 %	2.40 %	.ns
Attention problems (CBCL)	4.75	4.77	.01	5.31	4.33	-.24
Problem behaviors	Percentage/ LS Mean		AOR³ or ES⁴	Percentage/ LS Mean		AOR³ or ES⁴
Rule breaking behaviors (CBCL)	2.66	2.74	.03	2.90	2.38	-.23
Aggressive behaviors (CBCL)	6.72	6.99	.04	6.76	6.06	-.12
Socio-emotional difficulties	Percentage/ LS Mean		AOR³ or ES⁴	Percentage/ LS Mean		AOR³ or ES⁴
Anxious depressed (CBCL)	2.97	2.89	-.03	2.80	2.64	-.05
Withdrawn depressed (CBCL)	1.54	1.47	-.04	1.35	1.16	-.13
Social problems (CBCL)	1.15	1.31	-.04	1.25	.93	-.23

¹ Analyses control for being white, annual earnings, and target child is female. ² Analyses control for being black, respondent's age at baseline, and target child is female. ³ Adjusted odds ratio; ⁴ Effect size. ⁵ Unadjusted or raw percentages; incidence too low to reliably estimate in multivariate model, effect estimated with chi square statistic.

† p ≤ .10 * p ≤ .05 ** p ≤ .01 *** p ≤ .001

Exhibit 16. Comparison of indicators of potential for school success, problem behaviors, socio-emotional difficulties, self-regulation

	Whole Sample ¹ (n=793)			HPO Subgroup ² (n=122)		
	Control (n=405)	HFNY (n=388)	AOR ³ or ES ⁴	Control (n=60)	HFNY (n=62)	AOR ³ or ES ⁴
Risks for poor school outcomes	Percentage/ LS Mean		AOR³ or ES⁴	Percentage/ LS Mean		AOR³ or ES⁴
Receptive vocabulary (PPVT-IV):						
Standard score	95.81	95.71	-.16	94.41	95.49	.62
% below average (100)	67.97	64.98	.87	77.61	59.40	.43*
Skip school often (%)	6.47	2.35	.35**	4.53%	1.85%	.35
Problem behaviors	Percentage/ LS Mean		AOR³ or ES⁴	Percentage/ LS Mean		AOR³ or ES⁴
Bullying activities:						
Ever (%)	36.43	38.14	1.08	40.68	36.68	.85
Total (#)	.64	.63	-.01	.66	.74	.07
Deviant activities:						
Ever (%)	32.43	30.24	.90	30.6	25.76	.79
Total (#)	.48	.48	-.01	.58	.49	-.08
Socio-emotional difficulties	Percentage/ LS Mean		AOR³ or ES⁴	Percentage/ LS Mean		AOR³ or ES⁴
Social isolation	5.48	5.30	-.03	5.49	5.35	-.02
Ever bullied by others (%)	53.04	53.68	1.03	58.26	52.56	.79
Self-regulation	Percentage/ LS Mean		AOR³ or ES⁴	Percentage/ LS Mean		AOR³ or ES⁴
Delayed / Waited	56.68 %	56.00%	.99	54.08	50.13	.85
Resisted temptation	18.34 %	18.32%	1.00	15.95	16.91	1.07
Level of distracters generated	2.35	2.45	.08	2.25	2.20	-.04

¹ Analyses control for being black, annual earnings, number of other biological children, and target child is female. ² Analyses control for being Hispanic, cash assistance at random assignment, annual earnings, and target child is female. ³ Adjusted odds ratio; ⁴ Effect size.

† p ≤ .10 * p ≤ .05 ** p ≤ .01

Discussion

The current study evaluated the ability of HFNY to limit social, behavioral, and cognitive/academic problems among children served by the program. Given the paucity of research on the effects of home visiting immediately following the child's entry into school, the current study helps to fill an important gap. Results from HFA-based programs primarily have focused on the first three years of the child's life and reveal an early impact on the mental development of the infant (Caldera et al., 2007; Landsverk, 2002). Unfortunately, there is little research on HFA-based programs available past age three to comment on the persistence of this effect. However, the Nurse Family Partnership (NFP), another home visiting model based on a strong conceptual model, has had the opportunity to examine effects at age six, finding a program effect on receptive vocabulary and mental processing skills (Olds et al., 2004). While the current study did not replicate this finding, the results are consistent with the early gains in mental development reported by Healthy Families San Diego's RCT (Landsverk et al., 2002) and Alaska's RCT (Caldera et al., 2007). The pattern of findings observed within the HFNY RCT also further the idea that home visiting effects may be concentrated on outcomes that position home visited children for greater school success over the long-term.

In practice, the current study's findings are especially important because early grade retention can contribute to poor academic performance, which in turn can lead to early drop-out (Pagani et al., 2001). Similarly, students who are often absent from school are at risk for developing a sustained pattern of such behavior, which increases the risk for school drop-out (Sewell, Palmo & Manni, 1981). The impacts noted on these domains, for both the sample as a whole and the HPO subgroup, suggest that home visitors are capitalizing on the opportunity to help parents become involved in their child's learning experiences early on, and perhaps to also

help establish stable living arrangements that contribute to fewer school disruptions. Indeed, researchers have shown that poverty, parent involvement, and residential stability can affect children's early academic performance (Dubow & Ippolito, 1994). In future work, we will investigate these paths as possible intervening variables to help explain the effects observed.

For outcomes regarding children's self-regulatory ability and socio-emotional and behavioral functioning, the program did not meet expectations. Rather, children in the intervention group had similar scores on all indicators of problem behavior and socio-emotional difficulties. These findings, while inconsistent with the sustained effects on positive and negative parenting noted in the previous chapter, are consistent with the body of evidence on home visiting more generally.

With the exception of findings from the Healthy Families Alaska program, few RCTs have reported impacts on behavioral or socio-emotional functioning prior to age six. Again, when we look to the larger body of research, results from NFP suggest that the CBCL's continuous syndrome scale may not effectively discriminate differences between treatment arms, but when clinical classifications are used, the cut points are predictive. As seen with school-related outcomes, classifications such as requiring special education services, repeating a grade, and scoring below average on the receptive vocabulary test point to program impacts that keep some of the neediest children from faltering. In the current study, we limited our analyses of the CBCL to continuous scores, but as with the majority of studies that employed this method, tests of program effectiveness fell short. As we move forward, we plan to investigate if clinical classifications on the CBCL yield a different result. In the interim, the consistency of effects across both mother and child-derived indicators of behavioral and socio-emotional problems suggest that program impacts on this domain are limited or small. What is unclear is whether, as

children age and move into adolescence, the school experiences noted above will produce differences in the behavior across the two groups. Thus, as with NFP, it is possible that effects on the children's behavior don't fully emerge until adolescence.

Another possibility is that the levels of disadvantage presented in both the home and the communities targeted to receive the intervention may overwhelm the program's ability to effect changes in these areas. Additional resources or connections may be needed to capitalize on the earlier gains and to more fully support families as their children move into an ever widening and challenging world. It is also possible that the study's requirement that mothers must have custody of the child in order for the mother to report on the child's functioning and for the child to complete an interview may have minimized otherwise detectable effects by removing some of the most symptomatic and traumatized children from the sample (e.g., those in foster care).

HFNY's Year 7 evaluation extends the discussion regarding the effectiveness of home visiting in preventing early precursors of juvenile delinquency. Given the substantial consequences and costs associated with these risks, additional research is needed to better understand the mechanisms promoting school success and the factors limiting the program's influence on children's emotional and behavioral adjustment. Increased knowledge in these areas will help policy and program administrators develop better strategies to avoid missing opportunities to serve these at-risk children.

CHAPTER 7: DO THE BENEFITS OF HFNY OUTWEIGH ITS COSTS?

Introduction

Prior research has shown that the benefits of early childhood interventions far exceed the costs of such programs (Karoly et al., 2005). Home visitation programs in particular have been widely promoted as an efficient use of resources. Unfortunately, few evaluations of home visiting programs have included an economic component to support this level of confidence. In the current study, we generate a cost benefit analysis of the HFNY program from the perspective of the government to answer the following questions:

- What are the costs associated with the program?
- Does HFNY reduce spending for government supported programs?
- Does HFNY increase tax revenues?
- Do the benefits of HFNY exceed the costs?
- Do the specific characteristics and/or experiences of HFNY participants influence the costs and benefits related to the program?

Cost Benefit Analysis Variables

Healthy Families New York Program Costs. Our cost calculations reflect the two distinct service components that are provided by the HFNY program: outreach and assessment and home visitation. In general, all individuals referred to the HFNY program benefit from outreach efforts as well as the in-depth assessment that is offered to determine their level of risk to maltreat and their need for other services not offered by the program. Once the assessment is complete, referrals are issued for home visitation, the second service component of the program, and/or other services in the community. In the case of the random assignment study, respondents in both the intervention and control groups benefited from the outreach and assessment efforts, but

only those assigned to the intervention group were offered home visitation. Thus, separate estimates were derived for the two different service components. In addition, we calculated discrete estimates for each of the three study site programs to investigate a range of costs associated with operating the program.

Exhibit 17 summarizes the number of families who received an assessment, the number of families served, and the program cost estimates for outreach and assessment, home visitation, and an overall total. The number of families assessed was used as the denominator in calculating estimates for outreach and assessment. As shown in Exhibit 17, this number was unusually high during the evaluation's enrollment period (2000-2002), when the outreach and assessment service component was expanded in order to facilitate random assignment. To a lesser degree, it also affected the count of the number of families served, which includes the number of families assessed per year plus the number of families with an open home visitation case. The number of families served was used as the denominator to calculate costs for the home visitation and total per family per year estimates.

Cost estimates, the numerator for the per family per year estimates, were derived from the contracts with the three study sites, which included expenditures for personal services, contractual services, travel, equipment, and supplies. Program costs also included other expenditures related to training, quality assurance, and the data management information system, which were documented in other contracts. The contracts for these latter costs support infrastructure for operating the program statewide, and were therefore prorated according to the number of programs operating in each specific year.

Assessment costs included the following components: FAW salary and fringe, proportion of supervisor salary and fringe for supervising FAWs, proportion of salary and fringe of support

staff supporting FAW activities, proportion of costs for FAW as a function of total staff, local and out-of-town travel, wrap-around/professional development training, office supplies, and data management. Outreach costs are specified within the budgets to cover advertisements in newspapers, TV, radio, community associations, etc.

Home visiting costs included the following components: FSW salary and fringe, proportion of supervisor salary and fringe for supervising FSWs, proportion of salary and fringe of support staff supporting FSW activities, proportion of costs for FSW as a function of total staff (e.g., office space, phone/communication, equipment use/maintenance, fiscal, insurance, etc.), travel (local and out-of-town for trainings), wrap-around/professional development training, office supplies (general, postage, printing, etc.), and data management.

Expenses for the randomized controlled trial were not included in any of the cost calculations; rather, the following components were extracted from the program cost: interviewer salary, research evaluator salary, fringe (if applicable, some interviewers were hired as consultants and thus had no fringe), space, phones, travel (local and out-of-town for trainings), office supplies, incentives/money orders, legal fees, and equipment such as laptops for conducting interviews. In many cases, each of these costs was directly specified within the budget justifications. There were several costs, however, that had to be calculated as a proportion of the total number of staff accruing a cost (e.g., cellular phone units, office space, supplies, etc.). Whether or not specific costs were directly specified or had to be calculated varied as a function of the time period of the funding as well as by site. Only costs that were directly attributable to the randomized controlled trial were excluded using these methods.

Exhibit 17 also displays estimates for funds contributed by the state (90%) and those contributed as part of a local share (10%). For the years presented, the program sites and the

training contractor were required to provide a local share consisting of in-kind or cash contributions of at least 10% in order to receive state funding.

Combining the expenditures for outreach and assessment and home visitation for the families served, as supported by contributions from both the state and the local share, the average total cost of the HFNY program (across the three study sites) was approximately \$3,074 per family per year in undiscounted 2000 dollars. Total costs ranged from \$2,645 at Site B to \$3,836 at Site C. We attribute Site C's elevated costs to the higher cost of living in this area, and to the lower numbers of families assessed and served per year with a constant number of staff persons being maintained, particularly in the outreach and assessment positions.

In order to calculate a per family per year cost, we adjusted the yearly total program cost into a cost per day. This resulted in a daily cost of \$6.85 for Site A, \$6.53 for Site B, and \$9.32 for Site C, in addition to the assessment cost, which is a standard onetime cost per family (\$384, \$486, and \$857 for Sites A, B, and C respectively). We then multiplied the daily cost of the program by the number of days the family was enrolled in the program in each year to obtain a total cost per year, using the rate for the site in which each family was enrolled.

Exhibit 17. Healthy Families New York home visiting program costs (in undiscounted 2000 dollars)

Site	Outreach & Assessment Costs			Home Visiting Services Costs			Total Program Costs	
	Number of Families Assessed/Year	Outreach & Assessment Per Family Cost/Year	Outreach & Assessment Per Family Cost with 10% LS/Year	Number of Families Served/Year	Home Visiting Per Family Cost/Year	Home Visiting Per Family Cost with 10% LS/Year	Total Program Cost Per Family /Year	Total Program Cost with 10% LS Per Family /Year
Site A								
FY 00-01	515	\$385.52	\$424.01	566	\$1,516.42	\$1,666.93	\$1,867.20	\$2,052.73
FY 01-02	362	372.16	409.26	625	1,858.54	2,043.14	2,074.09	2,280.19
FY 02-03	328	297.68	327.32	552	2,114.52	2,324.58	2,291.40	2,519.08
FY 03-04	253	369.02	405.80	484	2,266.92	2,492.44	2,459.82	2,704.56
FY 04-05	235	366.67	403.21	412	2,684.91	2,951.93	2,894.05	3,181.91
FY 05-06	259	321.06	353.06	443	2,441.02	2,683.99	2,628.73	2,890.41
FY 06-07	240	332.49	365.63	420	3,053.15	3,357.38	3,243.14	3,566.31
	Average Cost	\$349.23	\$384.04		\$2,276.50	\$2,502.91	\$2,494.06	\$2,742.17
Site B								
FY 00-01	275	\$297.96	\$327.64	284	\$2,043.87	\$2,246.00	\$2,332.39	\$2,563.26
FY 01-02	243	346.43	380.90	357	1,967.82	2,162.41	2,203.62	2,421.68
FY 02-03	191	367.52	404.06	392	1,867.17	2,051.94	2,046.25	2,248.82
FY 03-04	179	414.22	455.47	387	1,836.40	2,018.56	2,027.98	2,229.23
FY 04-05	130	558.39	613.99	338	1,899.60	2,087.78	2,114.37	2,323.93
FY 05-06	135	518.85	570.54	271	2,469.50	2,714.60	2,727.97	2,998.82
FY 06-07	116	589.39	648.12	237	3,100.96	3,409.13	3,389.44	3,726.35
	Average Cost	\$441.82	\$485.82		\$2,169.33	\$2,384.35	\$2,406.00	\$2,644.58
Site C								
FY 00-01	230	\$500.48	\$550.38	233	\$3,371.60	\$3,706.02	\$3,865.64	\$4,249.31
FY 01-02	154	730.03	802.77	253	3,123.61	3,432.89	3,567.98	3,921.53
FY 02-03	115	890.95	979.70	266	3,202.62	3,520.00	3,587.81	3,943.56
FY 03-04	143	682.97	751.06	282	2,997.94	3,295.71	3,344.27	3,676.56
FY 04-05	106	858.83	944.41	271	2,671.75	2,936.70	3,007.68	3,306.10
FY 05-06	91	1,037.72	1,141.20	242	2,708.44	2,977.22	3,098.66	3,406.34
FY 06-07	97	751.48	826.38	202	3,595.82	3,953.14	3,956.68	4,349.96
	Average Cost	\$778.92	\$856.56		\$3,095.97	\$3,403.10	\$3,489.81	\$3,836.20
Average Cost (All Sites)		\$523.32	\$575.47		\$2,513.93	\$2,763.45	\$2,796.63	\$3,074.32

Food Stamps. We obtained food stamps eligibility and payment information from OTDA. Food stamps payments are based on the number of people who live together and buy food and cook meals together. As a result, the actual payment amount we received for each payment was for more people than just the respondent and her biological children and is likely to overestimate actual costs attributable to the study respondent and her target child. Payment data were converted into 2000 dollars prior to being allocated to the appropriate time period.

Public Assistance. OTDA also provided us with public assistance eligibility and payment data for the respondent and/or the target child. Payment data were converted into 2000 dollars prior to being allocated to the appropriate time period.

Foster Care. Information describing the actual foster care experiences of the target children was extracted from the CCRS. We were able to describe the total number of foster care placements, the stop and start dates for each placement, and the total length of time spent in each placement. We identified the number of days within each time period that the target child was in a foster care placement. We then applied the NYS age-adjusted foster care per diem rate for the 2000-2001 FY to each day spent in care. The rates by age are as follows: \$47.08 per day for children from birth to age 3, \$45.53 per day for children from ages 4 to 5, and \$48.47 per day for children age 6 to 11. These rates include administrative costs in addition to the costs of housing, clothing and diapering the children who were placed in care. We did not obtain data on or value foster care costs for siblings of the target children.

Preventive Services. We obtained information from the CCRS regarding whether or not the target child and his/her family received prevention and support services in any given year. Because we were unable to determine the specific type of preventive service provided, or the length of time the preventive service was received, we chose to apply the average yearly cost per

individual of receiving prevention and support services only to those children who were not also placed in foster care during that time period to avoid over counting services. We applied the 2008-2009 FY average expenditure per individual served for Prevention and Support Services adjusted to 2000 dollars (\$3,865) to each year in which prevention services without foster care placement were provided (OCFS, 2008).

CPS Investigation. Child protective services reports were extracted from CONNECTIONS from random assignment up through the target child's seventh birthday for the 1173 study mothers who completed the baseline interview and their target children. We created a set of variables to describe the number of unique confirmed CPS reports where the respondent was a confirmed subject and/or the target child was a confirmed abused/maltreated child during each time period. We applied the 2008-2009 FY average expenditures per individual served for a NYS CPS investigation adjusted to 2000 dollars (\$1,762) to each investigation to obtain a total cost for each time period (OCFS, 2008). Although we were unable to specifically value costs for CPS investigations for the respondents' other biological children, by virtue of including all reports where the mother was a confirmed subject of a report regardless of whether or not the target child was involved, we have obtained a partial estimate of the additional costs expended to provide child protective services to the families in this study.

Medicaid Delivery and Hospitalizations. We obtained information regarding low birth weight from the New York State Department of Health for women who were randomly assigned prenatally to obtain the birth weight of the target child. This information was used to corroborate women's self reports of birth outcomes. We used estimates from Schmitt and colleagues (2006) to value the costs attributable to maternal and infant delivery and hospitalization prior to initial discharge home (\$38,558 in 2000 dollars for low birth weight (LBW) and \$4,636 in 2000 dollars

for those who were not LBW) for a subset of the sample who had a single birth, were randomly assigned to the study at a gestational age of 30 weeks or less, and reported receiving medical services funded by Medicaid or were uninsured at the time of the child's birth. For the few cases where the birth interview was not completed, we used the response from the initial interview to determine whether or not mothers' were uninsured or had their medical services funded by Medicaid to estimate costs.

Because we expected to be able to obtain administrative data from the NYS DOH regarding Medicaid use, we did not request detailed information about medical insurance receipt in the Year 7 interview. Unfortunately, HIPAA requirements prevented us from obtaining this information. Because of this and the lack of detail in the Year 7 interview, we were unable to accurately value families' use of Medicaid across the time span of the study.

Tax Revenues from Earned Income. Respondents were asked about their employment experiences at each time point so that we could describe their employment from random assignment to the target child's 7th birthday. Respondents provided information on the number of jobs employed at in each time period, as well as the start and stop dates for each job and the total number of hours per week worked at each. Using this information, we were able to create monthly cost streams for each respondent's income from random assignment to the target child's 7th birthday.

Federal tax revenues were calculated from the respondent's annual earned income for each calendar year covered in the analysis. We used the federal tax tables from 2000 to 2009 for single head of household for deductions and taxable incomes. Annual taxable income was computed for each respondent by subtracting from the annual earned income the standard deduction for head of household filers and the total value of exemptions based on the number of

biological children under age 19 years and the respondent herself. The standard deduction value, the value per exemption, and the total number of exemptions are specific to the calendar year of the annual earned income. Annual tax revenue was calculated from annual taxable income using each year's respective tax rates. The annual tax revenue was prorated to each month in the analysis in which the mother worked according to that month's proportion of annual earned income as described above.

New York State tax revenue was computed in a substantially similar method as that used for Federal tax revenue. There were two primary differences in the calculation of New York State tax revenue. First, the standard deduction for head of household filers and the exemption values are different for New York State. In addition, the total value of exemptions in New York State is based only on the number of biological children under age 19 years. Because of this, the New York State taxable income amount is different from that used for Federal tax revenue. A second difference concerns the New York State Household Credit. This credit reduces New York State income tax for low income earners with dependents. The credit is calculated based on the number of New York State exemptions claimed for that year and the mother's New York State taxable income. The Household Credit is subtracted from the annual New York State tax revenue to yield the annual net New York State tax revenue. This annual tax revenue is then prorated to each worked month in the analysis in the same manner as described for Federal tax revenue above.

Medicare and Social Security tax revenues were assessed directly as percentages of monthly earned income. Employer and employee contributions to Medicare together equaled 2.9% of monthly earned income and employer and employee contributions to Social Security together equaled 12.4% of monthly earned income.

The Earned Income Tax Credit (EITC) is a refundable federal income tax credit to offset the burden of social security taxes and to provide an incentive to work. However, to qualify, taxpayers must file a tax return, even if they do not have a filing requirement. It is estimated that 15 million households and individuals do not file a return at all. The IRS does not require individuals or families under specific income levels to file federal tax returns (\$12,000 in 2009 for the single head of household). When individuals do file a return, they often fail to claim the credit (Berube & Gale, 2005), especially those unaware that the credit exists, facing language or cultural barriers, and for fear that by claiming the credit they will sacrifice their eligibility for other important income-support programs. Research suggests that eligible families with very low incomes, children, income from self-employment, limited education, or language barriers are more likely to miss out on this credit (Berube & Gale, 2005).

Given the complexity of filing tax returns and low earned incomes, as well as the demographic characteristics of our sample, we believe that many of the respondents did not receive EITC even though they may have been entitled. Therefore, we elected to exclude these revenues in our calculations.

While information regarding revenues was available for the majority of respondents, revenue data was missing for people who did not complete an interview in a given year. Revenues were available for 98% of respondents between baseline and target child age two, for 84% of respondents between target child age two to age three, and for 81% of respondents between target child age three to seven. To assess the pattern of our missing data, we created a dummy variable reflecting the presence or absence of missing data. Examination of the Pearson product-moment correlations (not presented) between the missing data dummy variable and selected baseline demographic characteristics suggested that the data were not randomly

distributed among the cases, but were related to several other variables (e.g., site, ethnicity, country of origin, and Kempe score). Rather than simply exclude those who had incomplete data from the analysis, since the percentage missing becomes rather substantial by target child age three, we used SAS 9.2 to multiply impute missing values. We used monotone missing techniques with a predictive mean matching method and 10 imputations to estimate the missing data. We then applied the average of the 10 imputations to the respective monthly time period and case where the missing data was estimated.

Cost Streams. We felt it was important to include costs from random assignment to the child's birth for those who were enrolled prenatally since involvement in home visiting might be expected to have an impact on connection to and receipt of services. Because our sample included women who were assigned prenatally or up to three months postnatally, and therefore had different lengths of time from random assignment to birth or from birth to random assignment, we were constrained in our ability to create monthly cost streams from the period from random assignment to six months post birth. We therefore collapsed costs for these time periods into one block of time (random assignment to six months post-birth).

Cost streams for government supported services such as food stamps, public assistance, and foster care were created from random assignment to six months post-birth, and then monthly from six months post-birth to 84 months post-birth. Similar cost streams were generated for federal and state taxes, as well as for Social Security and Medicare taxes.

Once these costs were allocated to the appropriate time period, we added the monthly costs for each service category together to create a total category cost for seven distinct time points: random assignment to target child age 1, age 1 to age 2, age 2 to age 3, age 3 to age 4, age 4 to age 5, age 5 to age 6, and target child age 6 to age 7.

Preventive services and child protective services investigation costs were allocated on a per event basis within these seven time points, while Medicaid costs for maternal and child delivery and hospitalizations, including those related to low birth weight, were allocated to the period from Random Assignment to target child age one.

Inflation. The value of a dollar from one year to another is not generally comparable. For example, the value of a dollar in 2000, when our study began, is worth more (i.e., has greater purchasing power) than a dollar in 2009, the final year in which some of our costs and benefits occurred. This difference in value is due to the effects of inflation. Because our study spanned up to nine years in time due to different start dates in each site, and because our costs and benefits occurred at various points within this time frame, we needed to make adjustments to the dollar value of those costs and benefits so that we could directly compare their relative value. To equalize the purchasing power of dollars from various years, we used the Consumer Price Index for all items for urban consumers for the New York-Northern New Jersey-Long Island area to convert the costs and benefits into real 2000 dollars, the year the program started (Kokoski, 2010).

Discounting. Not only do we have to account for the effects of inflation on the value of real dollars, we also have to account for the time value of money to reach its present value. What this really means is that money received in the present is more valuable than money received in the future. Discounting is the method used to apply this to real dollars, reducing the relative weight of dollars received in the future and generating a present value. We applied a discount rate of 3% to our estimates using the discount factor appropriate for each year (Drummond & McGuire, 2004). We discounted our estimates beginning in the second year of

the target child's life. Sensitivity analyses were also conducted to examine how the application of different discount rates changed the estimates.

Aggregating & Estimating Net Cost to Government. After adjusting for inflation and discounting the costs and benefits in each year, we aggregated the yearly present values for the costs and benefits to obtain total values for each. We then used these estimates to obtain separate total HFNY program costs, total government supported program costs, and total revenues for the control and HFNY groups. The net cost to government was obtained by subtracting the tax revenues from the government supported programs.

Net cost to government = total government supported program costs – total revenues

Subtracting the HFNY program cost from the net cost to government will yield a total cost savings attributable to averted expenditures and generated revenues.

Costs savings = net cost to government – program costs

We can also create a benefit cost ratio by dividing the net cost to government by the total HFNY program cost.

Benefit cost ratio = net cost to government / total program cost

This value represents the return in dollars for every dollar invested.

Alternatively, we can produce a percentage that reflects the total recovery of the government's investment in the program generated by savings in government supported programs or increased revenues by multiplying the benefit cost ratio by 100.

Percent recovery = (benefit cost ratio)*100

Data analysis

Data analyses included all of the original 1173 respondents who were randomly assigned to the study and completed a baseline interview. We used descriptive statistics to examine the distributional properties of our dependent variables prior to conducting multivariate analyses. Examination of these distributions revealed that the dependent variables were not normally distributed (e.g., food stamps costs, public assistance costs, child protective services costs, preventive services costs, foster care costs, Medicaid delivery and hospitalization costs, total government supported programs, total revenues, and total HFNY program costs), with the exception of the net cost to government, which was normally distributed. The dependent variables were then analyzed using generalized linear models in SAS 9.2, applying a negative binomial distribution with a log link function for data that were not normally distributed and a normal distribution with an identity link function for normally distributed data. While we did not calculate cost estimates for foster care alone because the incidence of children in foster care was too small to estimate costs reliably, we did include these costs in the estimates for total government supported programs. This allowed the actual costs to be captured in the cost benefit ratio and reduced the likelihood that we would report unreliable foster care cost estimates.

Analyses involving the total sample controlled for the following covariates: earnings at random assignment, receipt of public assistance at random assignment, the count of Kempe items endorsed as moderate or severe, female target child, and having at least a high school diploma or GED. We also conducted separate analyses for two different subsamples of women: the recurrence reduction opportunity (RRO) subgroup and the high prevention opportunity subgroup (HPO). Covariates for the RRO subgroup analyses included earnings at random assignment, receipt of public assistance at random assignment, female target child, total depressive

symptoms, being African American, general health status, and any moves in the past year. We used earnings at random assignment, receipt of public assistance at random assignment, female target child, and any moves in the past year as covariates for analyses involving the HPO subgroup. The use of different statistical tests resulted in estimates for total government supported programs, tax revenues, and net cost to government that cannot be obtained by summing the individual components of which they are comprised.

Results

Exhibit 18 presents the government supported program costs accrued by families from random assignment until the target child's 7th birthday for the control and treatment groups. All cost estimates are presented in 2000 dollars discounted at 3%. For the sample as a whole, there were no significant differences in costs for any of the government programs, although Medicaid delivery and hospitalization costs were comparatively lower in the home visited group, but not statistically so. These savings are likely a result of the significantly lower rate of low birth weight found for a subgroup of women who were randomly assigned to the home visited group prior to 30 weeks gestation (Lee et al., 2009).

A much different pattern emerged when we examined the use of government supported programs among women in the RRO subgroup (Exhibit 18). On the whole, women in this subgroup who received home visiting services had lower government supported program costs than women in the control group ($p < .12$). Involvement in HFNY had a substantial, although not statistically significant, effect on reducing costs of food stamps and public assistance for home visited women, compared to women who did not receive home visiting services. Use of child protective and preventive services was also lower for the HFNY group, again not significantly

so. The only area in which women in the home visited RRO subgroup had higher costs was for Medicaid deliveries and hospitalizations. This difference was also non-significant.

There were no significant differences in the costs between home visited women and control group women in the HPO subgroup, although some savings were realized in three of the five government supported program cost categories presented (Exhibit 18). On average, women in the home visited group had slightly higher food stamps and public assistance costs, but slightly lower child protective and preventive services costs. Medicaid delivery and hospitalization costs were also slightly lower for home visited women in this subgroup.

The average per family costs to government for the whole sample and the two subgroups are presented in Exhibit 19. As a reminder, due to the use of different statistical estimation methods that were specific to the unique distributions of the various cost categories, the figures shown within each table cannot be summed. Rather the net cost to government reflects a separate estimate that builds on the analyses used to estimate government supported programs and tax revenues.

As shown in Exhibit 19, there were no significant differences for the whole sample between the control and HFNY groups in government supported program costs or tax revenues in the time from random assignment to the target child's 7th birthday. While home visited women had slightly lower government supported program costs, control women earned slightly higher incomes, offsetting most of the difference. Overall, the group of women who received HFNY had an average savings of \$628 (SE=\$1,613) in the net cost to government over the women in the control group. Taking into account the net program cost (\$4,101) this resulted in a recovery of 15% of the cost to provide HFNY services. Stated differently, for every dollar invested, the program returned \$0.15.

For women in the RRO subgroup, those who were assigned to the home visiting treatment arm had lower government supported program costs than their counterparts in the control group ($p=.12$). Although the control group reported slightly higher revenues, this difference did little to offset the substantial savings in government supported program costs generated by home visited women. In fact, investment in HFNY produced a savings in the net cost to government of \$12,395 ($SE=\$7,247$) per family ($p=.09$), a net cost savings of \$8,475 after taking into account the cost of the program (\$3,920), and a return of \$3.16 for every dollar invested by the time the target child was seven years old. This amounted to a 316% recovery of the net HFNY cost invested.

There were no significant differences between the groups on government supported program costs and tax revenues for women in the HPO subgroup. Overall, the group of HPO women who received HFNY generated a savings in the net cost to government of \$1,020 ($SE=\$3,731$) over women in the control group, a recovery of 25% of the investment in the program by the target child's 7th birthday. This translated into a return of \$0.25 for every dollar invested.

Exhibit 18. Government supported program costs per family from random assignment to target child's 7th birthday, 2000 dollars discounted at 3%

Government Supported Program^a	Control	HFNY	Difference (control-HFNY)	p value
Whole Sample^b	(N=594)	(N=579)		
Food Stamps	\$10,950.13	\$11,091.14	\$-141.01	0.89
Public Assistance	10,971.40	10,474.70	496.70	0.74
CPS Investigations	846.86	859.88	-13.02	0.96
Preventive Services	1,136.25	965.97	170.28	0.73
Medicaid Delivery and Hospitalizations	3,374.90	2,276.92	1,097.98	0.13
RRO Subgroup^c	(N=52)	(N=52)		
Food Stamps	\$17,763.46	\$15,818.50	\$1,944.96	0.54
Public Assistance	22,179.04	16,663.76	5,515.28	0.41
CPS Investigations	2,666.84	1,485.66	1,181.18	0.46
Preventive Services	3,187.58	2,653.66	533.92	0.84
Medicaid Delivery and Hospitalizations	1,120.22	2,416.54	-1,296.33	0.50
HPO Subgroup^d	(N=88)	(N=91)		
Food Stamps	\$11,045.60	\$12,217.53	\$-1,171.93	0.59
Public Assistance	11,327.26	12,901.86	-1,574.60	0.70
CPS Investigations	581.05	387.78	193.26	0.67
Preventive Services	1,177.65	99.24	1,078.42	0.36
Medicaid Delivery and Hospitalizations	6,711.25	5,649.52	1,061.73	0.58

^a Individual items do not sum to total government supported program costs due to differences in the distribution of data and statistical tests used to obtain the estimates.

^b Covariates for these analyses include earnings at random assignment (RA), receipt of public assistance at RA, the count of Kempe items endorsed as moderate or severe, female target child, and having at least a high school diploma or GED

^c Covariates for these analyses include earnings at RA, receipt of public assistance at RA, female target child, total depressive symptoms, being African American, general health status, and moved in the past year.

^d Covariates for these analyses included earnings at RA, receipt of public assistance at RA, female target child, and moved in the past year.

Exhibit 19. Average per family cost to government from random assignment to target child's 7th birthday, 2000 dollars discounted at 3%

	Control (N=594)	HFNY (N=579)	Difference (control-HFNY)	p value
Whole Sample^a				
Government Supported Programs	\$28,763.41	\$27,357.50	\$1,405.91	0.53
Tax Revenues	4,389.75	4,194.83	194.92	0.69
Net Cost to Government ^b	26,858.71	26,231.06	627.65	0.70
HFNY Program Cost	517.93	4,618.92	-4,100.99	
Cost Savings			\$-3,473.34	
Benefit Cost Ratio			0.15	
RRO Subgroup^c				
Government Supported Programs	\$56,952.09	\$43,815.63	\$13,136.46	0.12
Tax Revenues	3,181.96	1,704.06	1,477.90	0.34
Net Cost to Government ^b	56,955.33	44,560.48	12,394.85	0.09
HFNY Program Cost	484.16	4,404.04	-3,919.87	
Cost Savings			\$8,474.98	
Benefit Cost Ratio			3.16	
HPO Subgroup^d				
Government Supported Programs	\$31,391.08	\$33,107.23	\$-1,716.15	0.66
Tax Revenues	3,753.20	3,705.54	47.66	0.96
Net Cost to Government ^b	30,687.78	29,667.38	1,020.40	0.78
HFNY Program Cost	509.14	4,634.72	-4,125.58	
Cost Savings			\$-3,105.18	
Benefit Cost Ratio			0.25	

^a Covariates for these analyses include earnings at random assignment (RA), receipt of public assistance at RA, the count of Kempe items endorsed as moderate or severe, female target child, and having at least a high school diploma or GED

^b Individual items do not sum to total values presented due to differences in the distribution of data and statistical tests used to obtain the estimates.

^c Covariates for these analyses include earnings at RA, receipt of public assistance at RA, female target child, total depressive symptoms, being African American, general health status, and any moves in the past year.

^d Covariates for these analyses included earnings at RA, receipt of public assistance at RA, female target child, and any moves in the past year.

Sensitivity Analysis

We conducted a series of post hoc sensitivity analyses to examine the effect of different assumptions on our cost estimates. Specifically, we examined how the choice of discount rate used and how the application of an alternate method of allocating costs to Medicaid and

uninsured deliveries and hospitalizations influenced our estimates. We also present two additional considerations and describe their influence on the assumptions we made about costs.

Discount Rates. We re-estimated our models to examine the influence of discount rates ranging from 0% to 7% on the benefit cost ratio for the whole sample and the two subgroups. The estimated benefit cost ratio is presented for each sample in Exhibit 20. There was very little change in the ratio as the discount rate increased for the whole sample and the HPO subgroup. Only for the RRO subgroup, where we saw a positive net cost savings, did the benefit cost ratio decrease as the discount rate increased. Even at the highest discount rate estimated (7%), the cost savings generated by the RRO group remained positive.

Exhibit 20. Benefit cost ratio by discount rate and group

	0%	1%	3%	5%	7%
Whole Sample	0.14	0.14	0.15	0.16	0.17
RRO Subgroup	3.37	3.29	3.16	3.05	2.95
HPO Subgroup	0.25	0.25	0.25	0.25	0.25

Medicaid and Uninsured Delivery & Hospitalizations. We used the average delivery and hospitalizations costs for normal weight and low birth weight babies (Schmitt et al., 2006) for our initial cost calculations. We felt that the use of the average cost was most appropriate because there were so few observations within an individual birth weight category (e.g., 2000-2499, 1750-1999, etc.) that the resulting distribution of costs varied considerably. Additionally, use of average costs in this instance takes into account the potential range of low birth weight costs that will be important in budget decisions, which are not necessarily represented in the current sample because low birth weight is of low incidence and requires a larger sample to capture the appropriate distribution. Given these concerns, we selected stability over individual cost assignments. However, we conducted additional analyses using the average cost of delivery and hospitalizations according to birth weight category in grams according to the estimates generated by Schmitt and colleagues (2006) to see how our original estimates might change with

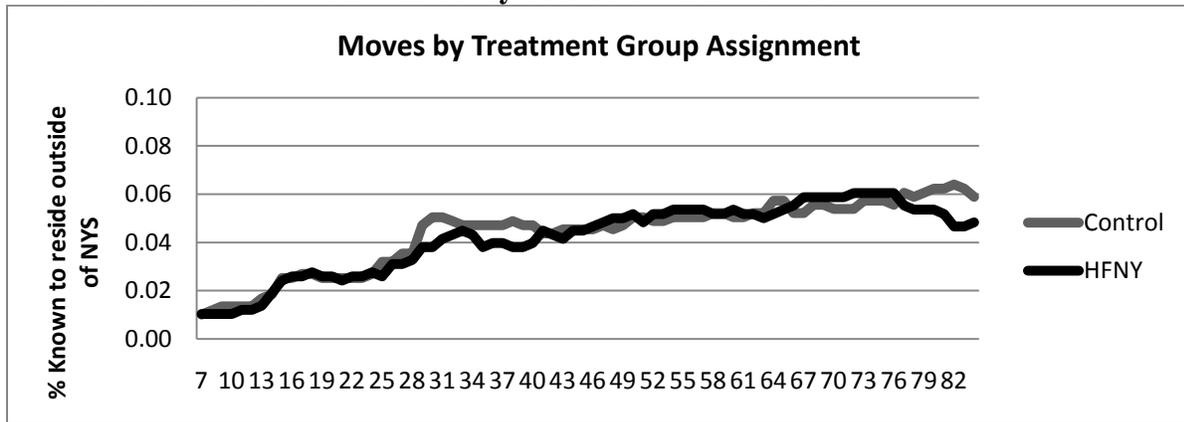
different assumptions. Exhibit 21 presents the results of these analyses for the whole sample and the two subgroups. The changes to the benefit cost ratios for the whole sample and the HPO subgroups were as expected after applying the specific costs related to the more narrowly defined birth weight categories. In both groups, the benefit cost ratio decreased. The benefit cost ratio for the RRO subgroup was unchanged. This is likely due to the small number of low birth weight babies born in this subgroup.

Exhibit 21. Medicaid and uninsured delivery and hospitalization costs by birth weight category and corresponding updated benefit cost ratio

	Control	HFNY	Difference	p value	Revised Benefit Cost Ratio
Whole Sample	\$2,834.24	\$2,135.74	\$698.50	0.27	0.06
RRO Subgroup	1,031.42	2,102.62	-1,071.20	0.52	3.16
HPO Subgroup	5,666.91	4,956.25	710.66	0.66	0.03

Use of administrative data. The current study used administrative data from NYS alone to estimate involvement with the child welfare system and many of the costs associated with government supported programs. This decision may have introduced bias into the results if those who left the state are different with respect to their use of social services by treatment group. To examine the potential for bias, we graphed the pattern of out of state residence for mothers in the HFNY and control groups with known moves across the 84 months of follow-up. As shown in Exhibit 22, the rate for residence outside of NYS was below 7% throughout the study, and women in the two treatment arms showed similar patterns of residency. Thus, it is unlikely that the use of administrative records in the current study substantially biased the outcomes reported.

Exhibit 22. Rates of known residency outside of New York State



Summary of possible under or over estimates in cost benefit analysis. Estimates derived in any cost benefit analyses are influenced by the data available and decisions regarding how best to utilize data. Exhibit 23 presents a summary of possible over or under estimates for each cost outcome.

Exhibit 23. Over / underestimates of cost outcomes

Cost Outcome	Direction of Estimate
Food Stamps	Overestimate: Use of actual food stamps payments are based on the number of people who live together, buy food, and cook meals together which overestimates costs attributable to mother and target child. In addition, other cost-benefit analyses typically assign an approximated or average cost. Underestimate: Food stamps payments for individuals living out-of-state are not included.
Public Assistance	Overestimate: Use of Public Assistance payments are based on family groups which overestimates costs attributable to mother and target child. In addition, other cost benefit analyses typically assign an approximated or average cost. Underestimate: Public assistance payments for individuals living out-of-state are not included.
CPS Investigations	Underestimate: CPS Investigation costs for individuals living out-of-state are not included.
Preventive Services	Overestimate: We were unable to determine the specific type of preventive service provided, or the length of time the preventive service was received so we chose to apply the average yearly cost per individual of receiving prevention and support services. It is unlikely that all families received services for a full year. Underestimate: Preventive Services costs for individuals living out-of-state are not included.
Foster Care Placements	Underestimate: Foster Care placement costs for individuals living out-of-state are not included.
Medicaid Delivery & Hospitalizations	Underestimate: Average cost does not take into account specific costs related to the child's experience.
Net Program Cost	Overestimate: We did not include program costs for those control group families who were erroneously enrolled in HFNY post-random assignment.

Discussion

With this cost benefit analysis, we set out to answer a series of questions to help us better understand the reality of the monetizable costs and benefits generated by involvement in a paraprofessional home visitation program. Only for the RRO subgroup did the benefits obtained exceed the costs of implementing the program; however both the whole sample and the HPO subgroup were moving toward a recovery of the initial investment.

The majority of savings were due to decreases in government supported program costs as opposed to increases in tax revenues. The most substantial source of savings was in the reduction of public assistance and food stamps benefits within the RRO group. This is likely a result of the reduction in subsequent births found for the RRO group in the time period between Baseline and Year 2. A similar finding was reported by Olds and colleagues (1993) who found that reductions in subsequent pregnancies in their low-income subgroup accounted for 32% of their total savings. As of this report, we have not conducted an in-depth analysis of the effect of employment on use of public assistance and food stamps programs. We anticipate that these future analyses may be able to provide us with information on how government supported programs are utilized by our sample when employed, as it is likely many of the employed women worked in low-wage jobs that did not prevent them from continuing to receive at least some government supported services.

The recovery of the government's investment (25%) for the HPO subgroup, a group of young, first-time mothers, created to replicate the sample served by NFP, is slightly lower than those obtained by Olds and colleagues (1993, Glazner, Bondy, Luckey & Olds, 2004) in the Elmira, Memphis, and Denver studies when children were approximately four years old (51%, 29%, and 26%, respectively). Although we only present the combined savings by the time of our

target child's 7th birthday in this report (25%), an examination of the savings generated by our target child's 4th birthday revealed a recovery of 20% of the initial program costs for this subgroup. While we cannot directly compare the results of the NFP studies to the current cost benefit analysis because of differences in the outcomes that were monetized and the methodological decisions that were made, the findings do suggest that HFNY is on track to generate long-term government savings.

The results of this cost benefit analysis should be interpreted with caution for several reasons. First, we limited our analysis to outcomes for which we could obtain NYS administrative data and their associated costs. We were interested in the cost of the program from the perspective of the government in general, and NYS in particular. Therefore, we started with the most direct administrative costs, which for OCFS included child protective, preventive and foster care services, for DOH included birth weight data, and for OTDA included accessible indicators of welfare and food stamp usage. While we could have included costs associated with receipt of special education services, which would likely, based on the size of the program effect in this area, advantage the HFNY group, we opted to restrict our analysis to administrative data sources for two reasons: (1) these data were available for the majority of families over most of the study duration; and (2) imputation for these school outcomes would be based on a single data point (as opposed to the many additional data points available and used to estimate the missing tax revenue information) for many of the children known to have foster care experiences and missing from the sample.

Secondly, the Kempe assessment provided to all individuals in the HFNY and control groups is in fact a service. Families who screen positive on this measure were eligible for the study. Those in the HFNY group received home visiting services while those in the control

group received referrals to services other than home visiting when a need was reported on the Kempe. It is possible that this service had the effect of connecting control group families to services that increased their functioning beyond where it would have been had they not received the referrals. This would have the effect of reducing the potential savings difference between the two groups.

An additional concern related to service referrals is the possibility that those who received home visiting services would be connected to more services and thus incur higher costs than the control group. While some of these connections, for both the home visited group and the control group, were captured within the administrative cost data we obtained, we were unable to accurately measure or place a cost on other services resulting from referrals received by families. However, the data presented earlier in the report on the average number of referrals immediately following random assignment was significantly higher ($p < .001$) for the control group (2.97 referrals) than the treatment group (1.73 referrals). Although we have no indication of whether or not a service was actually received as a result of the control group referrals, this is an indication that they had a significant opportunity to also accrue costs related to services, thus minimizing the potential cost differences between home visited and control families.

Fourth, we had to impute tax revenue data for up to 19% percent of missing cases between the child's third and seventh birthdays. While we used sophisticated techniques to multiply impute the missing data in the least biased way possible, it is likely that the imputed values are only a close approximation of the true observations. There is still some measure of error that must be taken into account.

Finally, this economic analysis adopts one perspective, the standard of government. While highlighting the financial costs and savings, it does not represent, in the current format,

the perspective of the families served. In terms of their quality of life, the outcomes we assess are measurable and suggest improvements in a variety of domains. This perspective is equally important although not monetarily quantifiable.

Generally, cost benefit studies identify their greatest savings over the long-term. Most of the savings accrue from reductions in criminal activity, positive educational outcomes, reductions in welfare use, and increased tax revenues due to higher education. Given the savings currently generated and the potential cost savings from outcomes that have been observed but not valued, we expect savings to continue to accrue in each of the groups.

CHAPTER 8: CONCLUSIONS AND IMPLICATIONS

The recently funded Maternal, Infant, and Early Childhood Visitation Program invests in home visiting, featuring it as important strategy for preventing child maltreatment. The new law outlines ambitious objectives for funded home visiting programs, including high expectations regarding performance targets and rigorous criteria for selecting effective programs. Within the context of this new legislation, the current study presents results from a longitudinal, randomized controlled trial that are both timely and important for thinking about how to optimize resources while also improving the performance of programs. Throughout the report, we have presented two perspectives: one that captures the family's experience and one that captures the government's experience. In considering the implications of the findings, we weighed the influence of each perspective to propose recommendations that appropriately reflect each vantage point. Taken together, the findings presented in the fidelity, outcome, and cost benefit analyses lend themselves to the following conclusions and recommendations.

HFA-based programs can produce sustained effects with a diverse population.

HFNY's persistent effects on serious physical abuse and positive parenting, which were first observed during the toddler and/or preschool years and then again at Year 7, are relevant to home visiting policy in at least two important ways. First, the sustained effects were detected when analyses were conducted with the whole sample. Thus, on average, the program produced significant differences among a very diverse group of families on both cross-sectional indicators of school engagement and longitudinal indicators of parenting. Furthermore, we believe HFNY to be a sound investment over the long term for families who have similar histories and characteristics as those in the whole study sample. Given the concentration of significant non-monetized findings for parenting and the not yet assigned benefits for outcomes related to

school, we consider the results of the cost benefit analysis to be an underestimate of the savings that might be expected to accrue following the initial investment in the program. Accordingly, we strongly recommend that the program continue to target a diverse group of mothers who are at considerable risk to maltreat, live in challenging communities, and depend on their home visitor to help their child chart a life course that averts risks for delinquency and promotes experiences associated with long-term school success.

A second important implication of the pattern of findings observed pertains to discussions regarding the effectiveness of different nationally-based home visiting models. While weak program effects are sometimes attributed to the use of paraprofessional staff or the HFA model, findings from the Year 7 follow-up suggest that HFA-based programs delivered by paraprofessionals can produce sustained effects on parenting that extend past the intended period of service. This pattern of effects helps to fill an important gap in the research on HFA-based home visiting programs. For example, data from several existing RCTs of HFA-based programs show a number of effects on less severe forms of parenting behaviors during the early years of a child's life (Duggan et al., 2004; Duggan et al., 2007; Landsverk et al., 2002; Mitchell-Herzfeld et al., 2005; DuMont, Mitchell-Herzfeld et al., 2008), but until now there has been little longitudinal data available to comment on the long-term effectiveness of these programs.

In sharp contrast to the number of evaluations in the field with longitudinal data on HFA-based home visiting programs, most appraisals of effective practices include criteria that require evidence of "significant, sustained, positive outcomes." The recently passed Maternal, Infant, and Early Childhood Home Visiting Program law also houses such a statement, further illustrating the prominent use of this criterion. Since language such as this is often a factor in determining a program's or model's designation as an evidence-based or effective practice,

having this information available is critical to assigning an effectiveness rating to the HFNY program or the HFA model. In turn, the field of home visiting more generally may also benefit. Given the large number and diversity of low-income families in need of assistance, the availability of multiple evidence-based home visiting programs that complement one another will ultimately maximize the delivery of effective, culturally responsive services to meet the complex needs of a broad population of low-income families.

Who is offered home visiting services matters.

While our overall findings demonstrate the benefits of providing HFNY services to all women in the sample, we also observed significant differences in program effects depending on who was offered HFNY services. For the two subgroups analyzed, the evaluation revealed several effects of clinical significance.

With respect to practice, the subgroup findings suggest ways in which HFNY resources may be optimized.

Establish strong links between local department of social services and HFNY. Program effects on confirmed reports of child abuse and neglect were most robust among the group of women with prior confirmed CPS reports who were assigned to HFNY. These findings are particularly significant given the lack of evidence that other interventions can successfully lower rates of maltreatment recurrence. Rather, home visiting programs working with these women typically report similar rates of maltreatment across the two treatment arms (c.f., MacMillian et al., 2005). While we cannot comment on how the children of these women fare, the consistent pattern of lower rates and levels of confirmed reports for both neglect and physical abuse, along with an identifiable intervening mechanism, are promising. Administrative cost data also suggest that these families are less resource dependent. Thus, we strongly urge that the effects

produced by HFNY's home visiting program for this subgroup be viewed as an opportunity to create meaningful change in the lives of other families with prior histories of confirmed reports. We recommend accomplishing this change by encouraging local child protective services agencies to refer recently or actively indicated CPS cases to HFNY when the mother is expecting or has recently delivered a child.

Prioritize services for those entering during the prenatal period, especially women fitting the descriptions of the two subgroups. With regards to less severe indicators of harsh and punitive parenting and child outcomes, effects were most pronounced among young mothers enrolled prior to the birth of their first child. HFNY parents in the HPO subgroup have consistently shown lower rates or levels of harsh and coercive parenting and minor physical aggression than their counterparts in the control group since age two and across methods and informants. In addition, the target children of these women are less likely to repeat a grade immediately following school entry and to score below average on the receptive vocabulary assessment. This subgroup also holds the potential for considerable long-term savings. Although the cost benefit analysis showed a relatively high initial investment, the benefits estimated in the current study are likely underestimates of the percent recovery because many of the benefits realized for the HPO sample are ones that are not readily monetizable. Specifically, there is the promise of greater returns resulting from the sustained impacts on rates of harsh parenting seen for this group of mothers at Years 2, 3, and 7, and the marked improvements in children's receptive language skills by age seven.

In light of these findings, we recommend that HFNY focus screening efforts on all pregnant women in a community, rather than adhering to the universal focus on all *new mothers*. Prioritizing the initiation of services during the prenatal period would also capitalize on the

program's effectiveness in helping mothers attain better birth outcomes (Lee et al., 2009). Furthermore, when young, first-time mothers or those with a prior substantiated report are referred during pregnancy, we recommend giving priority to these individuals. We do not recommend limiting home visiting services to these groups, rather, home visiting services for women who had recently given birth would be offered whenever a slot is available. This recommendation retains the opportunity for all eligible women to access services, takes full advantage of the opportunity for individual women to benefit from prenatal services (i.e., delivering a healthy weight baby), and maximizes the opportunity for the program to effect the greatest degree of change possible.

With regards to research, we recommend that future evaluations of home visiting programs corroborate the current findings with larger samples to allow for statistical tests that are adequately powered to detect small to medium effects. In addition, study designs that are stratified from the outset would help to minimize potential differences across the treatment groups. In the current study, statistical controls were used to compensate for differences in the groups' baseline characteristics, particularly for the RRO group, and the equivalent or nearly equivalent cell sizes (e.g., the number of respondents in the HFNY and control groups within each subgroup) helped to limit heterogeneity in variances across the two groups, thereby promoting valid and comparable estimates of the groups' standard errors.

Examining patterns of effects on neglect may inform program practice.

Consistent with earlier findings from the trial, the HFNY home visiting program presented with both strengths and weaknesses. While significant results can inform our recommendations for policy and practice, so too can changes in the pattern of results over time.

In the current study, administrative indicators and maternal reports of neglect provide a rich context for informing program practice.

In earlier waves of the RCT, there was some evidence to suggest that HFNY might prevent program participants from neglecting their target child (DuMont, Mitchell-Herzfeld et al., 2008). At Year 1, a trend was noted in the rate of neglectful behaviors as self-reported by mothers, whereby mothers in the intervention engaged in neglectful acts less often than mothers in the control group. At Year 2, a similar trend was noted in the frequency of neglectful acts. With regards to CPS data, rates were slightly but not significantly higher for families in the intervention arm, which was in part attributed to increased surveillance (DuMont, Mitchell-Herzfeld et al., 2008). By Year 7, both types of indicators of neglect suggest that the early evidence of a program effect in this area has attenuated. In contrast to the pattern of results for the sample as a whole, results for the newly tested RRO subgroup suggest that the program can effectively prevent neglect for at risk families, even at the level of administrative reports. This disparity creates an opportunity to examine how the program can be effective for one group, but not generalize to the entire sample.

In a recent study investigating the risk and protective factors associated with neglect during early childhood, three separate longitudinal studies reported consistencies related to indicators of economic hardship and parent well-being when predicting official reports of investigated neglect allegations (Shook-Slack et al., submitted). The potential for home visitors to promote parent well-being and lower the level of economic hardship for program participants may be of critical importance to the development of more effective service delivery strategies for neglect. This may require intensifying or altering efforts to promote mothers' mental health and/or self-sufficiency activities. The partial mediating role of subsequent pregnancies for the

RRO subgroup highlights the importance of strengthening families' chances for economic self-sufficiency via delayed or averted pregnancies. Thoughtful attention to this important area may benefit HFNY's participants more broadly.

Conclusion

There is evidence to suggest that involving families in home visiting services early on promotes positive experiences within the home during the initial years of life for both the mother and the child. These benefits range from healthier birth outcomes (Lee et al., 2009) to healthy parenting (DuMont, Mitchell-Herzfeld et al., 2008; Rodriguez et al., 2010) to positive school experiences. Home visiting presents a unique opportunity for trained workers to forge enduring relationships with families at a time when parents are vulnerable and the developmental path of the newborn is particularly malleable. Home visits provide a forum for encouraging healthy prenatal behaviors and parenting attitudes, engaging infants in play, modeling a positive adult-child bond, promoting self-sufficiency skills, and facilitating linkage to supportive services. Despite these services and the range of benefits reported, the evaluation also identified patterns of results that were inconsistent with the program's expectations, particularly regarding children's emotional and behavioral adjustment. Additional research is needed to reconcile this juxtaposition and to provide information on factors that may hinder the program's success, such as risks extending from the disadvantaged neighborhoods in which the programs are housed. The current study makes a timely contribution to our understanding of the relationship between home visiting and child maltreatment and youth outcomes that protect against or pose risks for later delinquency.

References

- Achenbach, T.M. & Rescorla, L.A. (2001). Manual for the ASEBA School-Age Forms & Profiles, University of Vermont, Research Center for Children, Youth & Families: Burlington, VT.
- Aos, S., Lieb, R., Mayfield, J., Miller, M. & Pennucci, A. (2004). *Benefits and costs of prevention and early intervention programs for youth, technical appendix*. Olympia, WA: Washington State: Institute for Public Policy.
- Bae, H., Solomon, P.L. & Gelles, J. (2008). Multiple child maltreatment recurrence relative to single recurrence and no recurrence. *Children and Youth Services Review*, *31*, 617-624.
- Base SAS 9.1.3 (2006). Procedures Guide, Second Edition, SAS Publishing, Vol. 1-4.
- Bauer, N.S., Herrenkohl, T.I., Lozano, P., Rivara, F.P., Hill, K.G. & Hawkins, J.D. (2006). Childhood bullying involvement and exposure to intimate partner violence. *Pediatrics*, *118*, 235-242.
- Bavolek, S.J. & Keene, R.G. (1999). *Adult-Adolescent Parenting Inventory (AAPI-2): Administration and development handbook*. Park City, UT: Family Development Resources Inc.
- Berube, A. & Gale, W.G. (2005). *Tax policies to help working families in cities*. Retrieved from http://www.brookings.edu/reports/2005/06childrenfamilies_berube.aspx.
- Broidy, L.M., Nagin, D.S., Tremblay, R.E., Bates, J.E., Brame, B., Dodge, K.A., Fergusson, D., Horwood, J.L., Loeber, R., Laird, R., Lynam, D.R., Moffitt, T.E., Pettit, G.S. & Vitaro, F. (2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross national study. *Developmental Psychology*, *39*, 222-245.

- Cairns, J. (2004). Discounting in economic evaluation. In M. Drummond & A. McGuire (Eds), *Economic evaluation in health care: Merging theory with practice*, (pp.236-255). New York: Oxford University Press.
- Caldera, D., Burrell, L., Rodriguez, K., Shea Crowne, S., Rohde, C. & Duggan, A. (2007). Impact of a statewide home visiting program on parenting and on child health and development. *Child Abuse & Neglect*, 31, 801-827.
- Campbell, S.B. (1995). Behavior problems in preschool children: A review of recent research. *Journal of Child Psychology and Psychiatry*, 36, 113-149.
- Cassidy, J. & Asher, S. (1992). Loneliness and peer relations in young children. *Child Development*, 63, 350-365.
- Children's Safety Network Economics & Data Analysis Resource Center, 2000.
- Daro, D. & McCurdy, K. (2007). Interventions to prevent child maltreatment. In L. Doll, J. Mercy, R. Hammond, D. Slett & S. Bonzo (Eds.), *Handbook on injury and violence prevention intervention*, New York: Kluwer Academic/Plenum Publishers.
- DePanfilis, D. & Zuravin, S.J. (2002). The effect of services on the recurrence of child maltreatment. *Child Abuse & Neglect*, 26, 187-2005.
- Díaz, J., Oshana, D. & Harding, K. (2004). *Healthy Families America: 2003 profile of program sites*. Chicago: Prevent Child Abuse America.
- Dodge, K.A. (1980). Social cognition and children's aggressive behavior. *Child Development*, 51, 162-170.
- Dodge, K.A., Bates, J. & Pettit, G.S. (1990). Mechanisms in the cycle of violence. *Science*, 250, 1678-1683.

- Dubow, E.F. & Ippolito, M.F. (1994). Effects of poverty and quality of home environment on changes in the academic and behavioral adjustment of elementary school children. *Journal of Clinical Child Psychology, 24*, 401-12.
- Duggan, A., Caldera, D.L., Rodriguez, K., Burrell, L.D., Rohde, C. & Crowne, S.S. (2007). Impact of a statewide home visiting program to prevent child abuse. *Child Abuse & Neglect, 31*, 801-827.
- 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*, 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*, 250-259.
- DuMont, K.A., Mitchell-Herzfeld, C., Greene, R., Lee, E., Lowenfels, A., Rodriguez, M., Dorabawila, V. (2008). Healthy Families New York Randomized Trial: Effects on early child abuse and neglect. *Child Abuse & Neglect, 32*, 295-315.
- DuMont, K.A., Rodriguez, M., Mitchell-Herzfeld, S., Walden, N., Kirkland, K., Greene, R., Lee, E. (2008). Effects of Healthy Families New York on Maternal Behaviors: Observational Assessments of Positive and Negative Parenting. Rensselaer, NY: New York State Office of Children and Family Services. Retrieved from <http://www.ocfs.state.ny.us/main/reports/>.

- DuMont, K.A., Spera, C., Mitchell-Herzfeld, S., Lee, E., Greene, R., Lowenfels, A. & Baum, D. (2005). Findings from a randomized control trial of the Healthy Families New York Program. Presented at the Society for Prevention Research: Washington, D.C.
- Dunn, L.M. & Dunn, D.M. (2007). *Peabody Picture Vocabulary Test (4th ed.)*. Bloomington, MN: Pearson Assessments.
- Eckenrode, J.J. & Runyan, D. (2004). The prevention of child abuse and neglect. *Annales Nestle*, 62, 32-42.
- Eddy, J.M., Leve, L.D. & Fagot, B.I. (2001). Coercive family processes: A replication and extension of Patterson's Coercive Model. *Aggressive Behavior*, 27, 14-25.
- Egeland, B. & Sroufe, L.A. (1981). Developmental sequelae of maltreatment in infancy. In R. Rizley & D. Cicchetti (Eds.), *New directions for child development: Developmental perspectives in child maltreatment*, (pp. 77-92). San Francisco, CA: Jossey-Bass.
- Eron, L.D., Huesmann, L.R. & Zelli, A. (1991). The role of parental variables in the learning of aggression. In D. Pepler & K. Rubin (Eds.), *The Development and treatment of childhood aggression*, (pp. 169-188). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Fagot, B.I. & Gauvain, M. (1997). Mother-child problem solving: Continuity through the childhood years. *Developmental Psychology*, 33, 480-488.
- Farrington, D.P. (1991). Childhood aggression and adult violence: Early precursors and later-life outcomes. In D.J. Pepler & K.H. Rubin (Eds.), *The development and treatment of childhood aggression*, (pp. 5-29). Hillside, NJ: Erlbaum.
- Farrington, D.P. (2005). Childhood origins of antisocial behavior. *Clinical Psychology and Psychotherapy*, 12, 177-190.

- Feldman, S., Downey G. (1994). Rejection sensitivity as a mediator of attachment behavior in young adults exposed to family violence. *Developmental Psychopathology*, 6, 231–247.
- Fergusson, D.M., Horwood, L.J. & Lynskey, M.T. (1996). Childhood sexual abuse and psychiatric disorders in early adulthood. II. Psychiatric outcomes of childhood abuse. *Journal of the American Academy of Childhood and Adolescent Psychiatry*, 35, 1365-1374.
- Fromm, L. (2001). Total estimated cost of child abuse and neglect in the United States: Statistical evidence. Chicago, IL: Prevent Child Abuse America. Retrieved from http://www.preventchildabuse.org/learn_more/research_docs/cost_analysis.pdf.
- George, R.M. & Lee, B.J. (1997). Abuse and neglect of the children. In R. Maynard (Ed.), *Kids having kids*, (pp. 205-230). Washington, DC: The Urban Institute Press.
- Glazner, J., Bondy, J., Luckey, D., Olds, D. (2004). Effect of the Nurse Family Partnership on government expenditures for vulnerable first-time mothers and their children in Elmira, New York, Memphis Tennessee, and Denver Colorado. Retrieved from http://www.acf.hhs.gov/programs/opre/welfare_employ/economic_analysis/reports/effect_nursefam/nursefam_toc.html.
- Gomby, D.S. (2007). The promise and limitations of home visiting: Implementing effective programs. *Child Abuse & Neglect*, 31, 793-799.
- Guterman, N.B. (2001). Stopping child maltreatment before it starts: Emerging horizons in early home visitation services. Thousand Oaks, CA: Sage Publications.
- Hall, S.M., Delucchi, K.L., Velicer, W.F., Kahler, C.W., Ranger-Moore, J., Hedeker, D., Tsoh, J. Y. & Niaura, R. (2001). Statistical analysis of randomized trials in tobacco treatment:

- Longitudinal designs with dichotomous outcome. *Society for Research on Nicotine and Tobacco*, 3, 193–202.
- Harding, K., Galano, J., Martin, J., Huntington, L. & Schellenbach, C. (2007). Healthy Families America effectiveness: A comprehensive review of outcomes. *Journal of Prevention & Intervention in the Community*, 34, 149-179.
- Harding, K., Reid, R., Oshana, D. & Holton, J., (2004). Initial results of the HFA implementation study. Chicago, IL: National Center on Child Abuse Prevention Research, Prevent Child Abuse America.
- Hawkins, J.D. (2003). *Seattle social development project: Intergenerational project measurement dictionary: G3 child interviews*. University of Washington.
- Hawkins, J.D., Herrenkohl, T., Farrington, D.P., Brewer, D., Catalano, R.F. & Harachi, T.W. (1998). A review of predictors of youth violence. In R. Loeber & D.P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions*, (pp. 106-146). Thousand Oaks, CA: Sage Publications.
- Herrenkohl, T.I., Huang, B, Tajima, E.A. & Whitney, S.D. (2003). Examining the link between child maltreatment and youth violence: An analysis of mediating mechanisms. *Journal of Interpersonal Violence*, 18, 1167-1188.
- Hollis, S. & Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomized controlled trials. *British Medical Journal*, 319, 670-674.
- Institute on Medicine. (1989). *Research on children and adolescents with mental, behavioral, and developmental disorders*. Washington, D.C.: National Academies Press.
- Karoly, L.A., Kilburn, M.R. & Cannon, J.S. (2005). *Early childhood interventions: Proven results, future promise*. Santa Monica, CA: RAND Corporation.

- Kempe, H. (1976). *Child abuse and neglect: The family and the community*. Ballinger Publishing Company: Cambridge, MA.
- Kessler, S., Nixon, A. & Nelson, C. (2008). Don't throw out the baby with the bathwater: A novel way of evaluating the Healthy Families America programs. *American Journal of Evaluation, 29*, 288-300.
- Kitzman, H., Olds, D.L., Henderson, C.R., Hanks, C., Cole, R., Tatelbaum, R., McConnochie, K. M., Sidora, K., Luckey, D., Shaver, D., Engelhardt, K., James, D. & Barnard, K. (1997). Effects of prenatal and infancy home visitation by nurses on pregnancy outcomes, childhood injuries, and repeated childbearing: A randomized controlled trial. *Journal of the American Medical Association, 278*, 644-652.
- Knutson, J.F., DeGarmo, D., Koepl, G. & Reid, J.B. (2005). Care neglect, supervisory neglect and harsh parenting in the development of children's aggression: A replication and extension. *Child Maltreatment, 10*, 92-107.
- Kokoski, M. (2010). Getting real: Price deflation of nominal values. *Journal of Benefit-Cost Analysis, 1*(1), Article 7. Retrieved from <http://www.bepress.com/jbca/vol1/iss1/7>.
- Korfmacher, J. (1999). The Kempe Family Stress Inventory: A review. *Child Abuse & Neglect, 24*, 129-140.
- Korfmacher, J., Kitman, H.K. & Olds D.L. (1998). Intervention processes as conditioners of home visitation program effects. *Journal of Community Psychology, 26*, 49-64.
- Landsverk, J., Carrilio, T., Connelly, C., Ganger, W., Slymen, D., Newton, R., Leslie, L. & Jones, C. (2002). *Healthy Families San Diego clinical trial: Technical report*. San Diego, CA: Child and Adolescent Services Research Center and San Diego Children's Hospital and Health Center.

- Lee, S., Aos, S. & Miller, M. (2008). *Evidence-based programs to prevent children from entering and remaining in the child welfare system: Benefits and costs for Washington*. Olympia: Washington State Institute for Public Policy.
- Lee, E., Mitchell-Herzfeld, S., Lowenfels, A., Greene, R., Dorabawilla, V. & DuMont, K.A. (2009). *Improving birth outcomes through prenatal home visitation: A randomized trial of Healthy Families New York*. Unpublished, University at Albany.
- Leventhal, J. M. (2005). Getting prevention right: Maintaining the status quo is not an option. *Child Abuse & Neglect*, 29, 209-213.
- Lipsey M.W. & Derzon J.H. (1998). Predictors of violent and serious delinquency in adolescence and early adulthood: A synthesis of longitudinal research. In R. Loeber & D.P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions*, (pp. 86–105). Thousand Oaks, CA: Sage Publications.
- Loeber, R., Green, S. & Lahey, B.B. (2003). Risk factors for adult antisocial personality. In D.P. Farrington & J.W. Coid (Eds.), *Early prevention of adult antisocial behavior*, (pp.79-108). Cambridge University Press.
- Love, J.M., Kisker, E.E., Ross, C., Raikes, H., Constantine, J., Boller, K., Brooks-Gunn, J., Chazan-Cohen, R., Tarullo, L., Brady-Smith, C., Fuligni, A., Schochet, P. Z., Paulsell, D. & Vogel, C. (2005). The effectiveness of Early Head Start for 3-year-old children and their parents: Lessons for policy and programs. *Developmental Psychology*, 41, 885-901.
- MacMillan, H.L., Thomas, B.H., Jamieson, E., Walsh, C.A., Boyle, M.H., Shannon, H.S. & Gafni, A. (2005). Effectiveness of home visitation by public-health nurses in prevention of the recurrence of child physical abuse and neglect: A randomized controlled trial. *Lancet*, 365, 1786-1793.

- Maxfield, M.G. & Widom, C.S. (1996). The cycle of violence: Revisited six years later. *Archives of Pediatrics and Adolescent Medicine, 150*, 390-395.
- McCarton, C.M., Brooks-Gunn, J., Wallace, I.F., Bauer, C.R., Bennett, F.C., Bernbaum, J.C., Broyles, R.S., Casey, P.H., McCormick, M.C., Scott, D.T., Tyson, J., Tonascia, J. & Meinert, C.L. (1997). Results at age 8 years of early intervention for low-birth-weight premature infants. The Infant Health and Development Program. *Journal of the American Medical Association, 277*, 126-132.
- Mercy, J., Butchart, A., Farrington, D. & Cerdá M. (2002). Youth violence. In E. Krug, L.L. Dahlberg, J.A. Mercy et al. (Eds.), *The world report on violence and health*, (pp. 25–56). Geneva (Switzerland): World Health Organization.
- Miller, M.E., Morgan, T.M., Espeland, M.A. & Emerson, S.S. (2001). Group comparisons involving missing data in clinical trials: A comparison of estimates and power (size) for some simple approaches. *Statistics in Medicine, 20*, 2383–2397.
- Mischel, W., Shoda, Y. & Rodriguez, M.L. (1989). Delay of gratification in children. *Science, 244*, 933-938.
- Mitchell-Herzfeld, S., Izzo C., Greene, R., Lee, E. & Lowenfels, A. (2005). *Evaluation of Healthy Families New York (HFNY): First Year Program Impacts*. Rensselaer, NY: New York State Office of Children & Family Services. Retrieved from http://www.ocfs.state.ny.us/main/prevention/assets/HFNY_FirstYearProgramImpacts.pdf
- Moffitt, T. (1993). Adolescent-limited and life-course persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100*(4), 674-701.

- Musil, C.M., Warner, C.B., Yobas, P.K. & Jones, S. L. (2002). A comparison of imputation techniques for handling missing data. *Western Journal of Nursing Research*, 24, 815-829.
- National Research Council and Institute of Medicine. (2006). *A study of interactions: Emerging issues in the science of adolescence workshop summary*. Washington, DC: The National Academies Press.
- National Research Council. (1993). *Understanding child abuse and neglect*. Washington DC: National Academies Press.
- New York State Average Benefits by Eligibility Type and Year. Average benefit obtained by dividing the expenditures for that category by the number of eligible in same category for that time period. Retrieved from <http://www.health.state.ny.us/nysdoh/medstat/quarterly/ssd/quarterly.htm>.
- New York State Office for Children and Family Services (OCFS). (June 2008). *New York State Child and Family Services Plans and Annual Progress and Service Reports (2005- 2009)*. Retrieved from <http://www.ocfs.state.ny.us/main/reports/Final%20Title%20APSR%20IV-B%20Plan.pdf>.
- O'Conner, T.G., Deater, D.K., Rutter, M. & Plomin, R. (1998). Genotype-environment correlations in late childhood and early adolescence: Antisocial behavioral problems and coercive parenting. *Developmental Psychology*, 34, 970-981.
- Olds, D.L., Eckenrode, J.J., Henderson, Jr., C.R., Kitzman, H., Powers, J., Cole, R., Sidora, R., Morris, P., Pettitt, L. & Luckey, D. (1997). Long-term effects of home visitation on maternal life course, child abuse and neglect and children's arrests: Fifteen-year follow-up of a randomized trial. *Journal of the American Medical Association*, 278, 637-643.

- Olds, D.L., Henderson, C.R. Jr., Tatelbaum, R. & Chamberlin, R. (1988). Improving the life-course development of socially disadvantaged mothers: a randomized trial of nurse home visitation. *American Journal of Public Health, 78*(11), 1436-45.
- Olds, D.L., Henderson, C.R. Jr., Phelps, C., Kitzman, H. & Hanks, C. (1993). Effect of prenatal and infancy nurse home visitation on government spending. *Medical Care, 31*, 155-174.
- Olds, D.L., Henderson, C.R. Jr., Kitzman, H.J., Eckenrode, J.J., Cole, R. & Tatelbaum, C. (1999). Prenatal and infancy home visitation by nurses: Recent findings. *The Future of Children, 9*, 44-65.
- Olds, D.L., Henderson, C.R., Kitzman, H. (1994). Does prenatal and infancy nurse home visitation have enduring effects on qualities of parental caregiving and child health at 25 to 50 months of life? *Pediatrics, 93*, 89-98.
- Olds, D.L., Kitzman, H.J., Cole, R., Robinson, J., Sidora, K., Luckey, D.W., Henderson, Jr., C.R., Hanks, C., Bondy, J. & Holmberg, J. (2004). Effects of nurse home-visiting on maternal life course and child development: Age 6 follow-up results of a randomized trial. *Pediatrics, 114*, 1150-1159.
- Olds, D.L., Kitzman, H.J., Hanks, C., Cole, R., Anson, E., Sidora-Arcoleo, K., Luckey, D.W., Henderson, Jr., C.R., Holmberg, J., Tutt, R. A., Stevenson, A.J. & Bondy, J. (2007). Effects of nurse home visiting on maternal and child functioning: Age-9 follow-up of a randomized trial. *Pediatrics, 120*, 832-845.
- Olds, D., Robinson, J., O'Brien, R., Luckey, D., Pettit, L., Henderson, C., Ng, R., Sheff, K., Korfmacher, J., Hiatte, S. & Talma, A. (2002). Home visiting by paraprofessionals and nurses: A randomized, controlled trial. *Pediatrics, 110*(3), 486-496.

- Olds, D.L., Sandler, L. & Kitzman, H. (2007). Programs for parents of infants and toddlers: Recent evidence from randomized trials. *Journal of Child Psychology and Psychiatry*, 48, 355-391.
- Olson, S.L., Bates, J.E. & Bayles, K. (1984). Mother-infant interaction and the development of individual differences in children's cognitive competence. *Developmental Psychology*, 20, 166-179.
- Pagani, L., Tremblay, R.E., Vitaro, F., Boulerice, B. & McDuff, P. (2001). Effects of grade retention on academic performance and behavioral development. *Developmental Psychopathology* 13, 297-315.
- Patterson, G. (1982). *Coercive family process*. Eugene, OR: Castalia Publishing Company.
- Pearlin, L.I. & Schooler, C. (1978). The structure of coping. *Journal of Health and Social Behavior*, 19, 2-21.
- Pianta, R.C., Smith, N. & Reeves, R.E. (1991). Observing mother and child behavior in a problem-solving situation at school entry: Relations with classroom adjustment. *School Psychology Quarterly*, 6, 1-15.
- Plotnick, R.D. & Deppman, L. (1999). Using benefit-cost analysis to assess child abuse prevention and intervention programs. *Child Welfare*, 78, 381-407.
- Prevent Child Abuse America. (2001). *Critical elements: Rationale and supporting research*. Retrieved July 16, 2009 from http://www.healthyfamiliesamerica.org/downloads/critical_elements_rationale.pdf.
- Radloff, L.S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1, 385-401.

- Reynolds, A.J., Ou, S.R. & Topitzes, J.W. (2004). Paths of effects of early childhood intervention on educational attainment and delinquency: A confirmatory analysis of the Chicago child-parent centers. *Child Development, 75*, 1299-1328.
- Rodriguez, M.L., Dumont, K., Mitchell-Herzfeld, S.D., Walden, N.J., Greene, R (2010). Effects of Healthy Families New York on the promotion of maternal parenting competencies and the prevention of harsh parenting. *Child Abuse & Neglect, 34*, 711-723.
- Sainani, K.L. (2010). Statistically speaking: Making sense of intention-to-treat. *The American Academy of Physical Medicine and Rehabilitation, 2*, 209-213.
- Schatz, J.N., Smith, L.E., Borkowski, J.G., Whitman T.L. & Keogh D.A. (2008). Maltreatment risk, self-regulation, and maladjustment in at-risk children. *Child Abuse and Neglect, 32*, 972-982.
- Schmitt, S.K., Sneed, L. & Phibbs, C.S. (2006). Costs of newborn care in California: A population-based study. *Pediatrics, 117*, 154-160.
- Sethi, A., Aber, J.L., Shoda, Y., Rodriguez, M.L. & Mischel, W. (2000). The role of strategic attention deployment in development of self-regulation: Predicting preschooler's delay of gratification from mother- toddler interactions. *Developmental Psychology, 36*, 767-777.
- Sewell, T.E., Palmo, A.J. & Manni, J.L. (1981). High school dropout: Psychological, academic, and vocational factors. *Urban Education, 16*, 65-76.
- Shaw, D.S. & Gross, H.E. (2008). What we have learned about early childhood and the development of delinquency. In A.M. Liberman (Ed.), *The long view of crime: A synthesis of longitudinal research*, (pp. 79-127). New York, NY: Springer.
- Shook-Slack, K, Berger, L.M., DuMont, K., Youn Yang, Kim, Ehrhard-Dietzel (submitted).
Risk and Protective Factors for Child Neglect during Early Childhood: A Cross-Study

- Comparison. *Child Services Review*.
- Smith, C. & Thornberry, T.P. (1995). The relationship between childhood maltreatment and adolescent involvement in delinquency. *Criminology*, 33, 451-481.
- Smith, K., Landry, S. & Swank, P. (2000). Does the content of mothers' verbal stimulation explain differences in children's development of verbal and nonverbal cognitive skills? *Journal of School Psychology*, 38, 27-49.
- Sternberg, K.J., Lamb, M.E., Guterman, E. & Abbott C.B. (2006). Effects of early and later family violence on children's behavior problems and depression: A longitudinal multi-informant perspective. *Child Abuse & Neglect*, 30, 283-306.
- Stier, D.M., Leventhal, J.M., Berg, A.T., Johnson, L. & Mezger, J. (1993). Are children born to young mothers at an increased risk of child maltreatment? *Pediatrics*, 91, 642-648.
- Straus, M.A., Hamby, S.L., Finkelhor, D., Moore, D.W. & Runyan, D. (1998). Identification of child maltreatment with Parent-Child Conflict Tactics Scales. Development and psychometric data for a national sample of American parents. *Child Abuse & Neglect*, 22, 249-270.
- Tremblay, R.E., Japel, C., Perusse, D., Boivin, M., Zoccolillo, M., Montplaisir, J. & McDuff, P. (1999). The search for the age of "onset" of physical aggression. Rousseau and Bandura revisited. *Criminal Behavior and Mental Health*, 9, 24-39.
- U. S. Advisory Council. (1990). *Child Abuse and neglect: Critical first steps in response to a national emergency*. Washington, D.C.: U.S. Department of Health and Human Services, Government Printing Office.
- U.S. Department of Health and Human Services (2001, 2003). *Child Maltreatment*. Washington, D.C.: Government Printing Office.

- U.S. Department of Health and Human Services, Administration on Children, Youth and Families (ACF). (2003). Emerging practices in the prevention of child abuse and neglect. Washington, D.C.: Government Printing Office.
- U.S. Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Families, Children's Bureau (2010). *Child Maltreatment 2008*. Retrieved from http://www.acf.hhs.gov/programs/cb/stats_research/index.htm#can.
- U.S. Department of Labor, Bureau of Labor Statistics. (2010). Consumer Price Index All Urban Consumers New York-Northern New Jersey-Long Island, NY-NJ-CT-PA. All items (1982-84=100). Washington, DC: U.S. Bureau of Labor Statistics. Retrieved from <http://www.bls.gov/ro2/nycpi9102.pdf>.
- Valla, J., Bergeron, L., Berube, H., Gaudet, N. & St. Georges, M. (1994). A structured pictorial questionnaire to assess DMS-III-R-based diagnosis in children (6-11 years): Development, validity, and reliability. *Journal of Abnormal Child Psychology*, 22, 403-423.
- Valla, J., Bergeron, L. & Smolla, N. (2000). The Dominic-R: A pictorial interview for 6- to 11-year-old children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39, 85-93.
- Widom, C.S. (1989). Child abuse and neglect, and violent criminal behavior. *Criminology*, 27, 251-271.
- Widom, C.S. & White, H.R. (1997). Problem behaviors in abused and neglected children grown-up: Prevalence and co-occurrence of substance abuse, crime, and violence. *Criminal Behavior and Mental Health*, 7, 287-310.

Williams, Stern & Associates (2005). Healthy Families Florida Evaluation Report, Miami, Fl.

Yoshikawa, H. (1995). Long-term effects of early childhood programs on social outcomes and delinquency. *The Future of Children*, 5, 51-75.

Zeilinski, D.S., Eckenrode, J. & Olds, D.L. (2009). Nurse home visitation and the prevention of child maltreatment: Impact on the timing of official reports. *Development & Psychopathology*, 21, 441-453.

Zuravin, S.J. (1988). Child maltreatment and teenage first births: A relationship mediated by chronic sociodemographic stress? *American Journal of Orthopsychiatry*, 58, 91-103.