

HHS’s Maternal, Infant, and Early Childhood Home Visiting Program:

Which Program Models Identified by HHS As “Evidence-Based” Are Most Likely To Produce Important Improvements in the Lives of Children and Parents?

Board of Advisors

Robert Boruch
University of Pennsylvania

Jonathan Crane
Coalition for Evidence-Based Policy

David Ellwood
Harvard University

Judith Gueron
MDRC

Ron Haskins
Brookings Institution

Blair Hull
Matlock Capital

Robert Hoyt
Jennison Associates

David Kessler
Former FDA Commissioner

Jerry Lee
Jerry Lee Foundation

Dan Levy
Harvard University

Diane Ravitch
New York University

Howard Rolston
Abt Associates

Isabel Sawhill
Brookings Institution

Martin Seligman
University of Pennsylvania

Robert Solow
Massachusetts Institute of Technology

Nicholas Zill
Westat, Inc.

President

Jon Baron
jbaron@coalition4evidence.org
202-683-8049

1725 I Street, NW, Suite 300
Washington, DC 20006
202-349-1130
www.coalition4evidence.org

The nonprofit, nonpartisan Coalition for Evidence-Based Policy strongly supports the new HHS Home Visiting program, which incorporates key evidence-based approaches that we have promoted through our work with Congressional and Executive Branch policy officials. The program – whose total funding over FY 2010-2014 is \$1.5 billion – awards grants to states to implement early childhood home visiting programs, and is required by law to reserve at least 75 percent of its funding for “evidence-based” program models.

However, evidence suggests that the program’s overall effectiveness will depend critically on which program models are implemented by the states. HHS has identified eight models as currently meeting the standard for “evidence-based” in the program’s authorizing statute, based on an HHS-sponsored evidence review carried out by Mathematica Policy Research. However, the statutory standard focuses on whether rigorous evaluations have found that the model produces *statistically-significant* effects, but not on whether these effects have *policy or practical importance*. As discussed below, we believe the evidence shows wide variation among the eight models in the importance of their effects, and that the program’s overall impact in improving people’s lives therefore depends critically on which models are implemented by the states.

We have therefore rated the eight models on the following criterion: **How much confidence does the evidence provide that, if a state were to replicate the model faithfully in a similar population, it would produce important improvements in the lives of at-risk children and parents?** Our ratings build on Mathematica’s findings about (i) which studies are of “high” quality (i.e., most likely to provide unbiased estimates of the model’s effects); and (ii) whether those studies show statistically-significant effects (i.e., meet the statutory standard). We add the further condition of whether those effects are important (e.g., sizable decrease in children’s hospitalizations, or sustained increase in school success).

Our ratings – and reasoning behind them -- are summarized briefly in the attachment (8½ pages plus references). The following table provides a guide:

Model	Level of confidence that the model will produce important life improvement
Nurse-Family Partnership	STRONG
Early Intervention Program	MEDIUM
Family Check-Up	MEDIUM
Early Head Start – Home Visiting	LOW
Healthy Families America	LOW
Healthy Steps	LOW
Parents as Teachers	LOW
Home Instruction Program for Preschool Youngsters	Insufficient evidence

We note that the Coalition is a neutral, objective party in assessing the evidence. We are a foundation-supported nonprofit organization with broad experience reviewing evidence for Congress and the federal agencies. We have no affiliation with any program models in home visitation or any other policy area.

1. Early Head Start – Home Visiting (EHS):

A. **Level of Confidence: LOW.** A large randomized evaluation found the program has few if any effects on key outcomes, but study weaknesses limit reliability of this result.

B. **Program description.** EHS is a program for low-income pregnant women and families with children from birth through age 3 years. EHS includes a variety of program approaches that vary by site, including a home-based program, a center-based program, and a combination of home- and center-based programs. The EHS home-based program is the model that HHS has identified as eligible for participation in the federal Maternal, Infant and Early Childhood Home Visiting Program, and is therefore the focus of both Mathematica’s and our evidence review. EHS home-based services include (i) weekly 90-minute home visits by program staff experienced in child development and health, adult learning, and related areas; and (ii) two group socialization activities per month for parents and their children.

C. **Summary of the evidence, based on findings from one large randomized controlled trial of moderate study quality.**¹ The EHS home-based program was evaluated in one large randomized controlled trial, with a sample of 1385 families at seven sites nationwide.² The study received Mathematica’s “moderate” (as opposed to “high”) rating for study quality, and meets our criteria for a well-conducted trial³ with the limitations noted below. We summarize the findings here, in the absence of any EHS studies with the “high” rating.

- **Key findings:** At the age-3 follow-up – approximately 21 months after program entry – the study found few or no effects on (i) child outcomes (e.g., cognitive/language development, social/emotional development, health);⁴ (ii) parenting outcomes (e.g., emotional support, stimulation of language and learning, negative parenting behaviors);⁵ (iii) parent mental and physical health;⁶ or (iv) family employment, welfare receipt, and rapid subsequent births.⁷ However, the study found that the program did produce statistically-significant increases of about 15-20% in various measures of parental participation in education and job training.

At the age-10 (5th grade) follow-up, the study found few or no effects on any outcome area, including (i) child social-emotional outcomes; (ii) child academic outcomes; (iii) parenting and child home environment, (iv) family well-being and mental health, or (v) parent self-sufficiency.⁸

- **Study limitations:** At the age-3 follow-up, the study had moderate-to-high sample attrition – outcome data were obtained for only 54-69% of the original sample, depending on the outcome measure. At the age-10 follow-up, the study had high sample attrition – for most outcomes, data were obtained for only 54% of the original sample. Such attrition could have undermined the equivalence of the program and control groups, and therefore reduces confidence in the findings.⁹

2. Early Intervention Program (EIP):

A. **Level of Confidence: MEDIUM.** A rigorous evaluation found sizable reductions in infant hospitalizations, but replication of this finding would be desirable to confirm its validity.

B. **Program description.** EIP is a home visitation program for healthy teenagers pregnant with their first child. Key program elements include (i) four motherhood classes provided to the expectant mothers during their third trimester, and (ii) two home visits by an experienced nurse during the second or third trimester, and up to 15 additional home visits during the child’s first year of life. Home visits cover topics such as maternal and infant health, family planning, life skills, and social

support. The visiting nurses also refer mothers to other services as needed (e.g., mental health counseling, child care).

C. Summary of the evidence, based on findings from one well-conducted randomized controlled trial.¹⁰ The program was evaluated in one trial with a sample of 144 pregnant teenagers, most of whom were Hispanic, low-income, and unmarried. The trial, summarized as follows, received Mathematica’s “high” rating for study quality¹¹ and meets our criteria for a well-conducted trial.³

- **Key findings:** At the six-week postpartum follow-up, the study found two statistically-significant effects: (i) infants in the EIP group were hospitalized an average of 27% fewer days than infants in the control group, and (ii) EIP-group mothers were more likely to be attending high school or junior college, or to have graduated from high school (the effect size is not reported). The study found no statistically-significant effects on outcomes in four other areas (prenatal health, birth outcomes, maternal substance use, and parenting behaviors); however, in these areas there was often a pattern of non-significant effects favoring the EIP group (e.g., on incidence of premature birth).¹²

At the age-2 follow-up, the study found similar effects, including a statistically-significant 58% reduction in average number of non-birth-related hospitalizations per infant. However, we believe these findings are less reliable than those at the earlier follow-up, for reasons discussed immediately below.

Study limitations: (1) At the longer-term (age-2) follow-up, the study had moderate-to-high sample attrition, which differed between the two study groups – outcome data were obtained for 75% of the EIP group versus 65% of the control group. Such attrition may have undermined the equivalence of the two groups, and therefore reduces confidence in the findings at age 2. (2) The study had a relatively small sample and was conducted in a single site (one county in California). Replication of these findings in a second trial, conducted in another setting, would be desirable to rule out the possibility that the findings occurred by chance, and confirm that the program is effective in other settings where it would normally be implemented.

3. Family Check-Up (FCU):

A. Level of Confidence: **MEDIUM. Rigorous studies have found promising – but not yet strong – evidence of effects on child behavior.**

B. Program description. Family Check-Up (FCU) is a three- to nine-session home visitation program for families with young children at risk for behavior problems, which provides parenting and family problem-solving advice and helps families identify needed services. The program is delivered by a trained, master’s- or Ph.D.-level parent consultant with experience in family interventions.

C. Summary of the evidence, based on findings from two well-conducted randomized controlled trials. The program was evaluated in two trials, summarized as follows, that received Mathematica’s “high” rating for study quality, and meet our criteria for a well-conducted trial.³

Study 1 – Pennsylvania, Oregon, and Virginia.¹³ This was a multi-site trial with a sample of 731 low-income mothers of toddlers (average age 2) identified as being at-risk of future behavior problems.

- **Key findings:** At follow-up two years after random assignment, children in the FCU group were 14-19% less likely than children in the control group to have behavior problems severe enough to require clinical care, based on their mothers' reports. Also, at the one-year follow-up, the program was found to produce an 11% reduction in the likelihood that mothers suffered from clinical depression.¹⁴ See the endnote for more detail on the effects.¹⁵
- **Study limitations:** (1) The study does not report whether the effects at the two-year follow-up point, described above, were statistically significant. It does report that the *change* in behavior over the two years (i.e., "slope") differed significantly between the FCU and control groups. Preferably, the study would have reported both, to rule out the possibility that the effects found at two years were due to chance. (2) The study relied exclusively on mothers' reports to measure children's behavioral outcomes – reports which can sometimes be affected by maternal depression and other factors. In future studies, corroboration of the effects by other raters (e.g., teachers) would be desirable to strengthen confidence in the findings.

Study 2 – Pittsburgh, Pennsylvania.¹⁶ This trial had a sample of 120 low-income mothers of boy toddlers (average age 2) identified as being at-risk of future behavior problems.

- **Key findings:** At the age-4 follow-up, FCU mothers were rated by outside observers as being significantly more involved with their sons – in talking, structuring child's play, and so on – than control group mothers during a home assessment.¹⁷ However, the study found that FCU had no statistically-significant effects on mothers' ratings of their child's aggression and destructive behavior. (The non-significant pattern of effects suggests that FCU may have had small positive effects on child behavior, but these effects could also be due to chance.)
- **Study limitations:** Because of its small sample, this study may not have been able to detect a modest improvement in child behavior if such improvement had occurred. (By contrast, study 1 had a larger sample capable of detecting more modest effects.)

4. Healthy Families America (HFA):

- Level of Confidence: LOW. Rigorous studies have found varying effects across local HFA programs – some show no effects, others show small positive effects.**
- Program description.** Healthy Families America is a flexible program model whose elements vary somewhat across state or local HFA programs. The program offers weekly home visits, conducted by trained paraprofessionals, to families at-risk of child maltreatment, beginning prenatally or within the first three months after a child's birth and continuing through the first three to five years of life. The program seeks to help families manage life's challenges and, in addition to home visits, may include parent support groups, job training, and other services.
- Summary of the evidence, based on findings from five well-conducted randomized controlled trials.** HFA programs at the state or local level have been evaluated in five trials, summarized as follows, that received a "high" rating for study quality from Mathematica and meet our criteria for a well-conducted trial.³

Study 1 – Hawaii Healthy Start: Few if any effects found on child or parent outcomes.¹⁸ This was a multi-site randomized controlled trial, with a sample of 685 families with a newborn child.

- **Key findings:** Over the three years after program entry, the study found few or no effects on (i) mother-reported parenting practices or substantiated Child Protective Services reports of

maltreatment (see endnote for details¹⁹); (ii) parental risk factors for child maltreatment (maternal mental health, substance use, partner violence²⁰); or (iii) the rate of mothers' Rapid Repeat Births. The study did find a statistically-significant effects on three (out of 20) measures of intimate partner violence, which would be important if valid, but could well be a chance finding as discussed immediately below.

- **Study limitations:** The effects on intimate partner violence may have appeared by chance given the numerous effects measured in this trial – more than 50 across the various reports, the large majority of which were not statistically significant.²¹ In addition, there was a statistically-significant *pre-program* difference between the program and control group in a key measure of intimate partner violence, favoring the program group, which may help explain the observed effect on intimate partner violence at follow-up.²²

Study 2 – Healthy Families Alaska: Evidence suggests possible, but limited, effects on child outcomes.²³ This was a multi-site randomized controlled trial, with a sample of 364 women who were pregnant or had recently given birth.

- **Key findings:** At the age-2 follow-up, the study found few or no effects on (i) child maltreatment (e.g., hospitalizations, official reports of abuse or neglect,); (ii) parental risks for child maltreatment (e.g., substance use, partner violence); or (iii) parental attitudes and disciplinary strategies. However, at age 2 the study did find statistically-significant positive effects on some measures of child development and behavior (e.g., 58% of program group children scored in the normal range of child mental development, versus 48% of the control group children; 82% of the program group scored in the normal range in externalizing behavior – such as aggression and rule-breaking – versus 77% of the control group).
- **Study limitations:** Although suggestive, the positive findings might have appeared by chance given the large number of effects measured in this study (of the 105 effects the study measured in total at age 2, 12 were statistically significant²¹). Thus, replication of these findings in a second trial would be desirable to confirm that they are valid.

Study 3 – Healthy Families Georgia: No effects found on child maltreatment, as measured by confirmed Child Protective Services (CPS) cases.²⁴ This randomized controlled trial had a sample of 249 families with newborn children. Although the study reported findings for a number of outcomes, Mathematica's review and ours identified only one such finding as scientifically valid – the program's effect on cases of child maltreatment confirmed by CPS.²⁵

- **Key findings:** During the 12 months after random assignment, the program had no statistically-significant or non-significant effects on CPS-confirmed cases of child maltreatment.
- **Study limitation:** CPS cases, as a measure of child maltreatment, may not be fully reliable for reasons discussed in the endnote.²⁶

Study 4 – Healthy Families New York: Evidence suggests possible, but limited, effects on child and parent outcomes.²⁷ This was a multi-site randomized controlled trial, with a sample of 1,254 women who were pregnant or had recently given birth.²⁸

- **Key findings on parenting practices and child maltreatment:** At various follow-up points during the seven years after random assignment, the study found a few statistically-significant positive effects on parenting practices, as follows: (i) at year 3, independent observers found positive effects on mothers' use of positive parenting strategies (e.g., listening, praising), but no

reduction in negative parenting behaviors (e.g., use of threats, blaming, fighting); and (ii) at year 7, the home-visited children were 8% less likely to report that their parents used minor physical aggression than control-group children; however, on five other child-reported parenting practices, the study found no significant effects or pattern of outcomes favoring the program group.

The study found larger effects on *mothers' reports* on their own parenting behaviors (including maltreatment) over the seven years and, by contrast, few or no effects on *official CPS reports* of child abuse and neglect. However, these two measures may not be as reliable as those summarized above, for reasons discussed in the endnotes.^{29, 26}

- **Key findings on children's behavior, academic success, and mental health outcomes:** At the age-7 follow-up, the study found few or no effects on either mother- or child-reported outcomes in these areas.³⁰
- **Key findings on birth outcomes:** For the subgroup of women enrolled in the study by 30 weeks gestation,³¹ the study found a statistically-significant reduction in the incidence of low birth weight newborns (from 10% to 5%), but no significant effect on the rate of pre-term births or children being born small for gestational age.
- **Study limitations:** Although the reduction in incidence of low birth weight newborns is of potential policy importance, replication of this finding in a second trial would be desirable to rule out the possibility that it was a chance finding resulting from the study's examination of many outcomes and subgroups.²¹

Study 5 – Healthy Families San Diego: Few if any effects found on child or parent outcomes.³²

This randomized controlled trial had a sample of 515 women who had recently given birth.

- **Key findings:** At the age-3 follow-up, the study found few or no effects on (i) maternal life course (e.g., high school degree); (ii) home environment (e.g., substance use); (iii) maternal mental health; (iv) partner violence; (v) parenting behaviors; (vi) child immunization rates, medical checkups, and health insurance coverage; (vii) family welfare receipt; (viii) child behavior; (ix) child cognitive outcomes; and (x) family use of community services (e.g., job training).³³ The study did find a few positive, statistically-significant effects in mothers' reports on their own parenting behaviors (including maltreatment), but such self-reports may not be a fully reliable outcome measure for reasons discussed in the endnote.²⁹
- **Study limitations:** None other than those noted just above.

5. Healthy Steps (HS):

- Level of Confidence: **LOW.** A large rigorous evaluation found small short-term effects, but no long-term effects, on parenting practices and child behavior/safety.**
- Program description.** Healthy Steps is a home visitation program for families with a newborn child, designed to strengthen parents' knowledge, attitudes, and behaviors in ways that promote child health and development. The program is administered by a team of trained specialists (e.g., early childhood educators, nurses, social workers) partnered with medical practitioners at pediatric or family medicine practices. The specialists provide two to five home visits between childbirth and age 30 months, and also work with the families in enhanced well-child office visits, parent group meetings, and other venues. Program participation is not limited to high-risk families.

C. Summary of the evidence, based on findings from a large, well-conducted randomized controlled trial.³⁴ The program was evaluated in one randomized controlled trial, with a sample of 2,235 families at six sites across the country, that received a “high” rating for study quality from Mathematica,³⁵ and meets our criteria for a well-conducted trial.³

- **Key findings:** In the short-term follow-up (2-4 months after random assignment), the study found statistically-significant, but small, positive effects on four parenting practices – namely 2 to 4 percentage point differences between HS and control-group families in (i) the percent putting their infants in the wrong sleep position; (ii) the percent incorrectly giving their infant water to drink; (iii) the percent bringing their infant to a well-child doctor visit at age 1-month; and (iv) the percent of infants receiving appropriate vaccinations at 2 months.

In the long-term follow-up (child age 5.5 years), the study found no meaningful effects – statistically significant or non-significant – on the 13 child and parenting outcomes that were measured (e.g., child behavior, development, and social skills; parenting practices such as making their child wear a bike helmet; and child hospitalizations in the past year).

- **Study limitations:** (1) In the long-term follow-up, the study had high sample attrition – outcome data were obtained for only 60% of HS families and 57% of control-group families. Such attrition could have undermined the equivalence of the two groups, and therefore reduces confidence in the findings at age 5.5.³⁶ (2) In both follow-ups, most of the parenting practices were measured through parent self-reports, which may not be fully reliable for reasons discussed in the endnote.²⁹ Immunizations and well-child visits, however, were measured through children’s medical records.

6. Home Instruction Program for Preschool Youngsters (HIPPY):

A. Level of Confidence: N/A. There is currently insufficient evidence to reliably assess HIPPY’s effectiveness.

B. Program description. HIPPY is an early education program, designed to provide parents who have limited education with assistance in preparing their young children (ages 3-5) for elementary school. The program offers weekly activities for 30 weeks of the year, alternating between home visits by a trained paraprofessional and group meetings (two one-on-one home visits per month and two group meetings per month). Program duration varies from two to three years across different HIPPY sites.

C. Summary of the evidence, based on findings from one randomized controlled trial.³⁷ The program was evaluated in one small trial, with a sample of 52 low-income, Hispanic families, that received Mathematica’s “high” rating for study quality, and meets our criteria for a well-conducted trial with the exceptions noted below.³

- **Key findings:** 15 weeks after random assignment (i.e., mid-way through the program’s first year), the study found a statistically-significant positive effect on one of three standardized measures of child vocabulary and school readiness, but no significant or non-significant effects on the other two measures.³⁸ The study also found a statistically-significant positive effect on parent-reported involvement in improving their child’s language skills.³⁹
- **Study limitations:** (1) Because of the study’s small sample (52 families) and short follow-up period (15 weeks), we believe it provides insufficient evidence to assess whether the program produces any meaningful gains in children’s educational outcomes. (2) Parental involvement in

teaching language skills at home was measured through parent self-reports, which may not be fully reliable for reasons discussed in the endnote.²⁹

- **Other studies:** Given the limited evidence from the study above, we also examined the two studies that received Mathematica’s “moderate” (as opposed to “high”) rating for study quality. For reasons discussed in the endnote, we believe these studies do not provide reliable evidence.⁴⁰

7. Nurse-Family Partnership (NFP):

- A. **Level of Confidence: STRONG.** An expert review panel convened by the Coalition for Evidence-Based Policy⁴¹ has found that NFP meets the Congressional “Top Tier” evidence standard: *Interventions shown in well-designed and implemented randomized controlled trials, preferably conducted in typical community settings, to produce sizable, sustained benefits to participants and/or society.* (The standard is based on legislative provisions from Public Laws 110-161 and 111-8.)
- B. **Program description.** NFP provides nurse home visits to pregnant women with no previous live births, most of whom are (i) low-income, (ii) unmarried, and (iii) teenagers. The nurses visit the women approximately once per month during their pregnancy and the first two years of their children’s lives. The nurses teach (i) positive health related behaviors, (ii) competent care of children, and (iii) maternal personal development (family planning, educational achievement, and participation in the workforce).
- C. **Summary of the evidence, based on findings from three well-conducted randomized controlled trials.**⁴² The Top Tier panel’s report, briefly describing the findings from each study as well as any limitations, is posted at http://toptierevidence.org/wordpress/?page_id=168. What follows is the “highlights” section of the panel’s report, as well as another short excerpt that summarizes the panel’s views.

“HIGHLIGHTS:

- **Intervention:** A nurse home visitation program for first-time mothers – mostly low-income and unmarried – during their pregnancy and children’s infancy.
- **Evaluation Methods:** Three well-conducted randomized controlled trials, each carried out in a different population and setting.
- **Key Findings:** Pattern of sizable, sustained effects on important child and maternal outcomes in all three trials. The specific types of effects differed across the three trials, possibly due to differences in the populations treated. Effects found in two or more trials include (i) reductions in child abuse/neglect and injuries (20-50%); (ii) reduction in mothers’ subsequent births (10-20%) during their late teens and early twenties; (iii) improvement in cognitive/educational outcomes for children of mothers with low mental health/confidence/intelligence (e.g., 6 percentile point increase in grade 1-6 reading/math achievement).”

* * *

“Importantly, the three trials – each carried out in a different population and setting – all found the program to produce sizable, sustained effects on important mother and child outcomes. This provides confidence that this program would be effective if faithfully replicated in other, similar populations and settings. However, the specific types of effects often differed across the three studies. For

example, two of the trials found a reduction in mothers' receipt of welfare, whereas the third trial did not. These differences may be caused by (i) differences in the study populations across the three trials (e.g., different rates of pre-program welfare participation); (ii) differences in the time periods when the trials were conducted (e.g., whether before or after the major 1996 welfare reforms); or (iii) other, unknown factors. Thus, although the study results provide confidence of overall effectiveness, they offer less confidence that a faithful replication of this program will reproduce the specific effects found in any one trial.

“The specific effects that were replicated, with no countervailing findings, in two or more of the trials – and thus are the most likely to be reproducible in a program replication – are: (i) reduction in measures of child abuse and neglect (including injuries and accidents), (ii) reduction in mothers' subsequent births during their late teens and early twenties, (iii) reduction in prenatal smoking among mothers who smoked at the start of the study, and (iv) improvement in cognitive and/or academic outcomes for children born to mothers with low psychological resources (i.e., intelligence, mental health, self-confidence).”

Study limitations: See above, regarding varying effects across the three trials (the [full panel summary](#) provides further detail).

8. Parents as Teachers (PAT):

- A. **Level of Confidence: LOW.** Rigorous studies have found few if any effects on child or parent outcomes.
- B. **Program description.** PAT provides home visits by trained parent educators (with a bachelor's or master's degree) to mostly low-income women starting during their pregnancy or child's infancy. Local sites decide on the intensity of home visits, ranging from weekly to monthly, and the duration of program participation, which may extend until kindergarten entry. The home visits are designed to improve parenting knowledge and practice, detect developmental delays, prevent child maltreatment, and increase children's school readiness. In addition to home visits, the program provides health and developmental screenings, group meetings, and a resource network.
- C. **Summary of the evidence, based on findings from two well-conducted randomized controlled trials.** The program was evaluated in two trials, summarized as follows, that received Mathematica's "high" rating for study quality, and meet our criteria for a well-conducted trial.³

Study 1 – Wagner, et. al. (1999).⁴³ This trial had a sample of 497 families – primarily Hispanic and low-income – with an infant age 0-6 months.

- **Key findings:** At the age-3 follow-up, the study found few or no effects on a broad range of outcomes, including parent knowledge, attitudes, and behaviors; maternal education and household economic status; child cognitive, social, and physical development; and child health. Of 47 effects measured, 4 were statistically significant – 1 positive (i.e., favoring the PAT group), 3 adverse (i.e., favoring the control group). Such findings could well have appeared by chance given the large number of outcomes measured in this study.²¹
- **Study limitations:** None other than the note about chance findings immediately above.

Study 2: Drotar et. al, 2009.⁴⁴ This trial had a sample of 527 families of diverse socio-economic status, with a healthy infant age 0-9 months. The trial evaluated PAT as implemented with its new curriculum (Born to Learn) designed to stimulate early brain development.

- **Key findings:** At the age-3 follow-up, the study found few or no effects on a broad range of child outcomes, including cognitive development, behavior, task persistence/competence, language, school readiness, and social skills. Of 12 effects measured, only 1 was statistically significant (an increase in child competence in playing with a new toy). The other 11 effects were not statistically significant and showed no pattern of superior outcomes for the PAT group. Given the sizable number of outcomes measured, the one statistically-significant effect could well have appeared by chance.²¹
- **Study limitations:** At the age-3 follow-up, the study had moderate-to-high sample attrition – outcome data were obtained for only 56-65% of the original sample, depending on the outcome measure.⁴⁵ Such attrition could have undermined the equivalence of the program and controls groups, and therefore reduces confidence in the findings.

References

-
- ¹ Love, J., Kisker, E., Ross, C. M., Schochet, P. Z., Brooks-Gunn, J., Paulsell, D., et al. (2002). *Making a difference in the lives of infants and toddlers and their families: The impacts of Early Head Start. Volumes I-III: Final technical report [and] appendixes [and] local contributions to understanding the programs and their impacts*. Washington, DC: U.S. Department of Health and Human Services, Head Start Bureau. Vogel, Cheri A., Yange Xue, Emily M. Moiduddin, Ellen Eliason Kisker, and Barbara Lepidus Carlson (2010). *Early Head Start Children in Grade 5: Long-Term Followup of the Early Head Start Research and Evaluation Study Sample*. OPRE Report # 2011-8, Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- ² Our summary focuses on the results from the 7 sites in the EHS Research and Evaluation Project (out of 17 total) that implemented the EHS home-based, as opposed to center-based or mixed, program approach. We note that, at the age-3 follow-up, the effects found for all 17 study sites combined were more positive than those found for the 7 sites implementing the home-based approach.
- ³ The criteria we used to assess whether the trial was well-conducted are summarized in the [RCT Checklist](#), and include such items as: (i) the program and control groups were similar in their pre-program characteristics; (ii) the study had low sample attrition, and similar attrition rates for the program versus control group; (iii) the study measured outcomes for all individuals assigned to the program group, regardless of whether or how long they participated in the program; (iv) study outcomes were assessed with valid measures; and (v) where appropriate, research staff collecting outcome data were unaware of which sample members were in the program versus control group.
- ⁴ Out of 15 child outcomes the study measured, it found statistically-significant effects on 1. For the other 14, there was a slight (but non-significant) pattern of superior outcomes for the program group.
- ⁵ Out of 23 parenting outcomes the study measured, it found statistically-significant effects on 1. For the other 22, there was no pattern of superior outcomes for the program group.
- ⁶ Out of 6 parent mental and physical health outcomes the study measured, it found statistically-significant effects on 1. For the other 5, there was a slight (but non-significant) pattern of superior outcomes for the program group.
- ⁷ Out of 36 outcomes the study measured in these areas, it found statistically-significant effects on 1. For the other 35, there was no pattern of superior outcomes for the program group.
- ⁸ Out of 56 outcomes the study measured at the age-10 follow-up, it found statistically-significant effects on 2. For the other 54, there was a slight (but non-significant) pattern of superior outcomes for the program group in some areas, such as child behavior.
- ⁹ Sample attrition rates were approximately the same for the program versus control group at both the age-3 and age-10 follow-up, providing partial reassurance that attrition did not undermine the equivalence of the two groups.
- ¹⁰ Koniak-Griffin, D., Anderson, N. L., Verzemnieks, I., & Brecht, M. L. (2000). A public health nursing early intervention program for adolescent mothers: Outcomes from pregnancy through 6 weeks postpartum. *Nursing Research*, 49(3), 130–138. Koniak-Griffin, D., Anderson, N. L., Brecht, M. L., Verzemnieks, I., Lesser, J., & Kim, S. (2002). Public health nursing care for adolescent mothers: Impact on infant health and selected maternal outcomes at 1 year postbirth. *Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine*, 30(1), 44–54. Koniak-Griffin, D., Verzemnieks, I. L., Anderson, N. L., Brecht, M. L., Lesser, J., Kim, S., et al. (2003). Nurse visitation for adolescent mothers: Two-year infant health and maternal outcomes. *Nursing Research*, 52(2), 127–136.

¹¹ Mathematica limited its “high” rating for study quality to the short-term (six-week postpartum) follow-up, and gave the longer-term (age 2) follow-up a “moderate” quality rating due to high sample attrition. We summarize the results from both follow-ups, but also discuss the attrition issue at age 2 as a study limitation.

¹² The study found one statistically-significant adverse effect: the “external social competence” of EIP group mothers (i.e., a measure of their social skills) increased at a slower rate than that of the control group mothers over the course of the study. However the study also found that, at the start of the study, the control group had a significantly lower level of external social competence than the EIP group. Thus the greater gains in social competence experienced by the control group during the study could well reflect regression to the mean, rather than a true adverse effect of the program.

¹³ Dishion, T. J., Shaw, D., Connell, A., Gardner, F., Weaver, C., & Wilson, M. (2008). The family check-up with high-risk indigent families: Preventing problem behavior by increasing parents’ positive behavior support in early childhood. *Child Development*, 79(5), 1395–1414. Shaw, D. S., Connell, A., Dishion, T. J., Wilson, M. N., & Gardner, F. (2009). Improvements in maternal depression as a mediator of intervention effects on early childhood problem behavior. *Development and Psychopathology*, 21, 417–439.

¹⁴ It does not appear that maternal depression was measured in year two.

¹⁵ At the two-year follow-up, unadjusted effect sizes for the three child behavior outcomes measured (CBCL internalizing, CBCL externalizing, and Eyberg behavior problems) were 0.17-0.19, and at one-year follow-up, the unadjusted effect on maternal depression was 0.15.

¹⁶ Shaw, D. S., Dishion, T. J., Supplee, L., Gardner, F., & Arnds. (2006). Randomized trial of a family-centered approach to the prevention of early conduct problems: 2-year effects of the family check-up in early childhood. *Journal of Consulting and Clinical Psychology*, 74(1), 1–9.

¹⁷ The effect was statistically significant; its size was 0.48 standard deviations.

¹⁸ Duggan, A. K., McFarlane, E. C., Windham, A. M., Rohde, C. A., Salkever, D. S., Fuddy, L., et al. (1999). Evaluation of Hawaii’s Healthy Start program. *Future of Children*, 9(1), 66–90; discussion 177–178. Duggan, A., Fuddy, L., Burrell, L., Higman, S. M., McFarlane, E., Windham, A., et al. (2004). Randomized trial of a statewide home visiting program to prevent child abuse: Impact in reducing parental risk factors. *Child Abuse & Neglect*, 28(6), 623–643. Duggan, A., McFarlane, E., Fuddy, L., Burrell, L., Higman, S. M., Windham, A., et al. (2004). Randomized trial of a statewide home visiting program: Impact in preventing child abuse and neglect. *Child Abuse & Neglect*, 28(6), 597–622. El-Kamary, S. S., Higman, S. M., Fuddy, L., McFarlane, E., Sia, C., & Duggan, A. K. (2004). Hawaii’s Healthy Start home visiting program: Determinants and impact of rapid repeat birth. *Pediatrics*, 114(3), e317–e326. Bair-Merritt, M. H., Jennings, J. M., Chen, R., Burrell, L., McFarlane, E., Fuddy, L., et al. (2010). Reducing maternal intimate partner violence after the birth of a child: A randomized controlled trial of the Hawaii Healthy Start home visitation program. *Archives of Pediatrics and Adolescent Medicine*, 164(1), 16–23.

¹⁹ Of the 20 such parenting practice and child maltreatment outcomes measured over the three-year follow-up period, the study found 2 statistically-significant effects (on frequent use of corporal/verbal punishment and neglectful parenting behaviors). As the study authors observed, these effects could well have appeared by chance given the sizable number of outcomes measured. For the other 18 outcomes measured, there was no consistent pattern of positive or adverse effects.

²⁰ Of the 8 parental risk outcomes measured, there were no statistically-significant effects.

²¹ By design, each test for statistical significance has a 1 in 20 chance of giving a false positive answer. Thus, if a study measures a large number of effects, it is very likely to produce at least a few findings of statistical significance if only by chance.

²² Specifically, at the start of the study, 52% of control-group mothers had experienced three or more incidents of partner violence in the past year, compared to 43% of program-group mothers. This difference was statistically significant ($p=0.02$). See Duggan 2004 and El-Kamary 2004, cited above.

²³ Caldera, D., Burrell, L., Rodriguez, K., Crowne, S. S., Rohde, C., & Duggan, A. (2007). Impact of a statewide home visiting program on parenting and on child health and development. *Child Abuse & Neglect*, 31(8), 829–852. doi:10.1016/j.chiabu.2007.02.008. Duggan, A., Caldera, D., Rodriguez, K., Burrell, L., Rohde, C., & Crowne, S. S. (2007). Impact of a statewide home visiting program to prevent child abuse. *Child Abuse & Neglect*, 31(8), 801–827.

²⁴ Chambliss, J. W. (1998). An experimental trial of a home visiting program to prevent child maltreatment (Doctoral dissertation, Georgia State University, 1998). *Dissertation Abstracts International*, 61(03B), 152–1628. (AAI9967277).

²⁵ Consistent with Mathematica’s review, we found that most of the effects measured in the study (other than confirmed CPS cases) suffered from a limitation that greatly reduces confidence in their validity – a high rate of sample attrition, which differed between the program and control groups (53% for the program group versus 40% for the control group). Based on the study report, we believe this may have occurred because the study did not seek outcome data for the families assigned to the program group who did not participate fully in the program – a violation of the “intent-to-treat” principle, which could have distilled the program group down to its more motivated members and thereby inflated the study’s estimates of the program’s effects.

²⁶ Official CPS reports of child maltreatment may be unreliable (biased against the program group) if HFA home visitors report suspected child maltreatment to CPS that would otherwise go unreported in the control group. This “surveillance bias” could potentially mask a true program effect on maltreatment.

²⁷ DuMont, K., Mitchell-Herzfeld, S., Greene, R., Lee, E., Lowenfels, A., Rodriguez, M., et al. (2008). Healthy Families New York (HFNY) randomized trial: Effects on early child abuse and neglect. *Child Abuse & Neglect*, 32(3), 295–315. Dumont, Kimberly A, Susan Mitchell-Herzfeld, Rose Greene, Eunju Lee, Ann Lowenfels and Monica Rodriguez (June 2006). Healthy Families New York (HFNY) Randomized Trial: Impacts on Parenting after the First Two Years, *Office of Children and Family Services Working Paper #1*. Lee, Eunju, PhD., et. al. (2009). Reducing low birth weight through home visitation. *American Journal of Preventive Medicine*, 36 (1), 154-60. Rodriguez, M.L., Kimberly Dumont, Susan Mitchell-Herzfeld, N.J. Walden, R Greene (2010). Effects of Health Families New York on the promotion of maternal parenting competencies and the prevention of harsh parenting. *Child Abuse and Neglect*, 34, 711-723. DuMont, Kimberly, Kristen Kirkland, Susan Mitchell-Herzfeld, Susan Ehrhard-Ditezell. (October 31, 2010). *A Randomized Trial of Healthy Families New York (HFNY): Does Home Visiting Prevent Child Maltreatment?* NIJ Grant 2006-MU-MU-0002.

²⁸ Most of the follow-up reports on this study – in addition to measuring the program’s effects on the full sample – also examined effects on the subgroup of first-time mothers under age 19 enrolled by 30 weeks gestation. For this subgroup, the effects were approximately the same as those we have described for the full sample – perhaps slightly more positive but still rarely reaching statistical significance. At the age-7 follow-up, the study examined effects on another subgroup: women with an official report of abuse or neglect against them in the five years prior to random assignment. For this subgroup, the study found a statistically-significant reduction in 2 of 12 officially-reported measures of child abuse and neglect. However, this subgroup analysis appears to have been planned after the study began; therefore, we suggest that its results need confirmation in another trial to ensure they did not occur by chance during examination of a range of potential subgroups.

²⁹ Mothers’ reports on their own parenting behaviors may be unreliable (biased in favor of the program group) because the program specifically taught the home-visited mothers what is considered appropriate parenting behavior, so these mothers may have reported parenting outcomes they knew to be socially acceptable.

³⁰ Of the 22 child behavior, academic success, and mental health outcomes that the study measured at the age-7 follow-up, the study found 2 statistically-significant effects. For the other 20 outcomes measured, there was no

consistent pattern of positive or adverse effects. Given the sizable number of outcomes measured, the two significant effects could well have appeared by chance.

³¹ The study measured birth outcomes for the subgroup of women enrolled in the study by 30 weeks gestation (rather than the full sample), since this is the subgroup for whom the program might have plausibly affected birth outcomes.

³² Landsverk, J., Carrilio, T., Connelly, C. D., Ganger, W., Slymen, D., Newton, R., et al. (2002). *Healthy Families San Diego clinical trial: Technical report*. San Diego, CA: The Stuart Foundation, California Wellness Foundation, State of California Department of Social Services: Office of Child Abuse Prevention.

³³ Out of 61 outcomes that the study measured in these areas, the study found 3 statistically-significant effects.

³⁴ Minkovitz, C., Strobino, D., Hughart, N., Scharfstein, D., Guyer, B., & Healthy Steps Evaluation Team (2001). Early effects of the Healthy Steps for Young Children Program. *Archives of Pediatrics & Adolescent Medicine*, 155(4), 470–479. Guyer, B., Barth, M., Bishai, D., Caughy, M., Clark, B., Burkom, D., Tang, C. (2003). The Healthy Steps for Young Children Program National Evaluation. Baltimore: Women’s and Children’s Health Policy Center, Department of Population and Family Health Sciences, Johns Hopkins Bloomberg School of Public Health. Minkovitz, C. S., Strobino, D., Mistry, K. B., Scharfstein, D. O., Grason, H., Hou, W., Guyer, B. (2007). Healthy Steps for Young Children: Sustained results at 5.5 years. *Pediatrics*, 120(3), 658–668.

³⁵ Mathematica limited its “high” quality rating to the short-term follow-up, and gave the longer-term follow-up a “moderate” quality rating due to high sample attrition at that follow-up. We summarize both follow-ups, but also discuss the attrition issue as a study limitation.

³⁶ The program and control group families in follow-up sample at 5.5 years were still highly similar in their observable pre-program characteristics, which provides partial reassurance that sample attrition did not undermine the equivalence of the two groups.

³⁷ Necoechea, D. M. (2007). Children at-risk for poor school readiness: The effect of an early intervention home visiting program on children and parents. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 68 (6-A), 2311. (Dissertation Abstract: 2007-99230-512).

³⁸ Specifically, the study found no effect on the Picture Peabody Vocabulary Test, measuring receptive language skills (standardized effect size = -0.02); a statistically-significant positive effects on the Expressive One-Word Picture Vocabulary Test, measuring children’s ability to name pictures of common objects (standardized effect size = 0.35); and no effect on the Developmental Skills Checklist, measuring school readiness tasks such as letter identification, and ability to write one’s name (standardized effect size = -0.03).

³⁹ The standardized effect size was 0.87.

⁴⁰ One of these studies, a randomized controlled trial conducted in New York with a sample of 247 families, suffered from a limitation that greatly reduces confidence in its findings – it did not seek outcome data for the families assigned to the HIPPY group who did not participate fully in the program. So, for example, within the first month of the program’s start, 31% of the HIPPY group members were lost from the sample, compared to 22% of control group members. This is a violation of the “intent-to-treat” principle, which could well have distilled the HIPPY group down to its more motivated members and thereby inflated the study’s estimates of the program’s effects. The second study was a comparison-group studies conducted in Arkansas with a sample of 226 families that suffered from a similar limitation: HIPPY group members self-selected themselves into the program and participated for at least one month, whereas the comparison group was comprised of matched families from the community who did not self-select in this way. This could well have created a difference in family motivation favoring the HIPPY group and thereby inflated the study’s estimates of the program’s effects. Baker, A. J. L., & Piotrkowski, C. S. (1996). *Parents and children through the school years: The effects of the home instruction program for preschool youngsters*. New York: National Council of Jewish Women, Center for the Child. Baker, A. J. L., Piotrkowski, C. S., & Brooks-Gunn, J. (1999). The Home Instruction Program for Preschool Youngsters (HIPPY). *Future of Children*, 9(1), 116–133.

⁴¹ The expert panel was convened as part of Top Tier Evidence initiative, established by the nonprofit, nonpartisan Coalition for Evidence-Based Policy to assist Congress and the Executive Branch in identifying social programs meeting the Congressional Top Tier evidence standard. See www.toptierevidence.org.

⁴² All three trials received Mathematica’s “high” study-quality rating. Mathematica also identified a fourth trial that met its “high” rating -- Nguyen, J. D., Carson, M. L., Parris, K. M., & Place, P. (2003). A comparison pilot study of public health field nursing home visitation program interventions for pregnant Hispanic adolescents. *Public Health Nursing, 20*(5), 412. doi:10.1046/j.1525-1446.2003.20509.x. This trial had a sample of 225 low-income adolescent Hispanic girls. The study’s follow-up at time of delivery found a reduction in incidence of premature births (4.3% for the NFP group versus 8.2% for the control group) and low birth weight (5.6% for the NFP group versus 10.6% for the control group), which would be important if valid. However, the study does not report any tests of statistical significance (or standard errors) to provide confidence that the findings are not due to chance. The study found no effect on mean gestational age.

⁴³ Wagner, M., Clayton, S., Gerlach-Downie, S., & McElroy, M. (1999). *An evaluation of the northern California Parents as Teachers demonstration*. Menlo Park, CA: SRI International.

⁴⁴ Drotar, D., Robinson, J., Jeavons, L., & Lester Kirchner, H. (2009). A randomized, controlled evaluation of early intervention: The Born to Learn curriculum. *Child: Care, Health & Development, 35*(5), 643–649.

⁴⁵ Sample attrition rates were approximately the same for the program versus control group, which provides partial reassurance that attrition did not undermine the equivalence of the two groups.