



SCHOOL of
PUBLIC POLICY

26

Syracuse Family Development Research Program

Douglas J. Besharov
Peter Germanis
Caeli A. Higney
and
Douglas M. Call

September 2011



Maryland School of Public Policy
Welfare Reform Academy
www.welfareacademy.org

Part of a forthcoming volume
Assessments of Twenty-Six Early Childhood Evaluations
by Douglas J. Besharov, Peter Germans, Caeli A. Higney, and Douglas M. Call

Note: This report is open to public comments, subject to review by the forum moderator. To leave a comment, please send an email to welfareacademy@umd.edu or fill out the comment form at http://www.welfareacademy.org/pubs/early_education/chapter26.html.

26

Syracuse Family Development Research Program

The Syracuse Family Development Research Program (FDRP), which operated in Syracuse, New York from 1969 to 1975, “attempted to break the well-documented link between low-education, low-income households and children’s later educational difficulties.” The program was designed “to influence the permanent environment of the child, the family and the home.” The program was parent-centered, and “weekly contact with mothers and other family members in the home of each child was stressed as the key intervention component.”¹ This contact was achieved through weekly home visits, guided by the approach of “nonjudgmental family advocacy.”² Although the evaluation did not use random assignment, it is included in this volume because it is often cited as proof that early childhood programs work.

Peter L. Mangione, J. Ronald Lally and Alice S. Honig,³ researchers at Syracuse University (the “Syracuse team”) conducted the evaluation using a comparison-group design (not random assignment), comparing children who participated in the program from 1969–1971 to “similar” children in other neighborhoods. The Syracuse team reported significant improvements in school performance for girls and reductions in juvenile delinquency among both boys and girls in adolescence. FDRP has often been cited as an effective crime prevention program.⁴ Severe

¹J. Ronald Lally, Peter L. Mangione, and Alice S. Honig, “The Syracuse University Family Development Research Program: Long-Range Impact on an Early Intervention with Low-Income Children and Their Families,” in *Parent Education as Early Childhood Intervention: Emerging Directions in Theory, Research and Practice*, ed. Douglas Powell (Norwood, NJ: Ablex Publishing Corporation, 1988), 80.

²Lally, Mangione, and Honig, 1988, 80.

³ Peter L. Mangione is currently the co-director of the Center for Child and Family Studies at WestEd, a nonprofit research, development, and service agency. J. Ronald Lally is currently the director of the Center for Child and Family Studies at the Far West Laboratory for Educational Research and Development in San Francisco. Previously he was at Syracuse University, where he chaired the Department of Child and Family Studies and directed the Family Development Research Program. Alice Honig is professor emerita of child development at Syracuse University.

⁴See, for example, Sanford Newman, T. Berry Brazelton, Edward Zigler, Lawrence W. Sherman, William Bratton, Jerry Sander, and William Christeson, *America’s Child Care Crisis: A Crime Prevention Tragedy* (Washington, DC: Fight Crime: Invest in Kids, 2000).

attrition and possible selection bias, however, seriously undermine these findings.

Program Design

Program group. From 1969 to 1971, the FDRP recruited pregnant women in their last trimester of pregnancy from “very deprived” families. All families that enrolled in the program had incomes below \$5,000 in 1970 dollars (or \$25,168 in 2005 dollars). The mothers tended to be young (mean age of eighteen years), black, and single. Most had less than a high school diploma and little or no work experience. Unlike many other early childhood programs operating at the time, the FDRP provided services to both parent and child.

Services. The FDRP provided comprehensive education, nutrition, health, and safety services to parents beginning during the prenatal program and continuing until children entered school. The emphasis of the program was reaching children through their parents: The staff were to act in support of the parents, not as substitutes for them, in promoting the development of their children.

Families received weekly home visits from paraprofessionals, who provided advice to the parents, with the goal of improving family functioning and increasing cohesiveness. Parenting skills emphasized during these visits included an affectionate mother-child relationship, yielding to children’s needs for self-comforting activities, and positive reinforcement. In addition, the paraprofessionals provided the parents with adult-oriented services, including advice on finances and employment, as well as referrals to community services as needed. Children from six to fifteen months were provided center-based care for a half-day, five days per week; children from fifteen months to five years were provided center-based care full-time.⁵ The curriculum in the center stressed active child participation, language development, and the development of trust and initiative.⁶ Lastly, the paraprofessionals acted as a liason between the center staff and the parents.

The Evaluation. FDRP enrolled 108 families. When the program children were three years old, a comparison group of 108 similar children was selected. They were matched “in pairs with Center children with respect to sex, ethnicity, birth ordinality, age, family income, family marital status, maternal age, and maternal education status (non high-school diploma) at the time of the infant’s birth.”⁷

⁵As an indication of the intensity of the program, compare the 8,250 hours of preschool intervention services a child in this project could have received to the 720 hours received in a typical Head Start program.

⁶Lally, Mangione, and Honig, 1988, 81.

⁷Lally, Mangione, and Honig, 1988, 90.

The evaluation used a variety of data sources. Parents were interviewed at three and five years into the program.⁸ Both children and parents were interviewed again when the children were age fifteen. Various IQ and other tests were administered during the preschool years. Subsequent school performance was measured from school records. Children's behavior was measured from teacher interviews and information about juvenile delinquency was obtained from court and probation records.

Major Findings

The Syracuse team reported significant improvements in school performance for girls and reductions in juvenile delinquency among both boys and girls in adolescence.⁹ FDRP has often been cited as an effective crime prevention program.

Cognitive. When the program children were three years old, they outscored the comparison children by 19 points on an IQ test (an average of 110 vs. 91), but by age five, these differences were no longer statistically significant, nor were there any significant differences on other measures of cognitive development and intellectual ability.

School readiness/performance. Although Lally, Mangione, and Honig found few significant impacts on school performance overall, girls tended to benefit from the intervention. In the seventh and eighth grade, the girls who had participated were less likely to have below a C average (24 percent vs. 47 percent), to be failing (0 percent vs. 16 percent), or to have excessive (more than 20) school absences in the previous year (0 percent vs. 31 percent). There were no statistically significant differences in either grade retention or special education placement.

Socioemotional development. The Social-Emotional Observer Rating of Children was used to assess children's socioemotional functioning. At thirty-six months old, program children exhibited "superior functioning" compared to the children in the comparison group. This pattern continued through the first grade in terms of interactions with other children, but the program group's behavior toward adults was mixed, with more positive and more negative behaviors. Lally, Mangione, and Honig hypothesize that the negative behavior could have resulted because the children received less personalized attention than they expected, leading some to react in negative ways.

Health. Data apparently either not collected or not reported.

⁸Peter L. Mangione, Syracuse University, e-mail message to Peter Germanis, February 13, 2003, states: Interviewers were graduate students in developmental psychology, child development, and social work. The graduate students had worked as professionals in social services before pursuing their advanced degrees.

⁹Lally, Mangione, and Honig, 1988.

Behavior. Program girls were rated by teachers as having more positive attitudes toward themselves and toward other people relative to girls in the comparison group. Similarly, teachers reported that girls also had better achievement in school and control of their impulses with respect to other people. There were, however, no differences for boys.

Crime/delinquency. The strongest findings were reductions in juvenile delinquency. Despite the absence of many cognitive or school-related impacts, relatively large reductions in juvenile delinquency were reported when the children were fifteen years old. Only 6 percent of FDRP children were assessed as probation cases by the County Probation Department compared to 22 percent of comparison group children. Moreover, the crimes committed by the comparison youth were more serious, including burglary, robbery, and physical and sexual assault.

Early/nonmarital births. None of the girls in either the program or comparison follow-up sample had an early nonmarital pregnancy by age fifteen.¹⁰

Economic outcomes. Data apparently either not collected or not reported.

Effects on parents. The findings also show that the program “had practically no impact on family income and career advancement of the parents.”¹¹

Benefit-cost findings. Apparently a benefit-cost analysis was not performed. The program cost about \$7,345 per year (in 2005 dollars) per program family, averaged over the five-year intervention program.¹²

Overall Assessment

The Syracuse FDRP was evaluated using a comparison-group design (not random assignment), comparing children who participated in the program to “similar” children in other neighborhoods. The Syracuse team reported significant improvements in school performance for girls and reductions in juvenile delinquency among both boys and girls in adolescence. Severe attrition and possible selection bias, however, seriously undermine these findings.

Program theory. With regards to the program’s theory, Lally, Mangione, and Honig note:

The major thrust of the intervention was to influence and have impact on the more

¹⁰Peter L. Mangione, Syracuse University, e-mail message to Peter Germanis, February 13, 2003.

¹¹Lally, Mangione, and Honig, 1988, 102.

¹²J. Ronald Lally, Peter L. Mangione, and Alice S. Honig, Syracuse University, e-mail message to Peter Germanis, November 12, 2001.

permanent environment of the child, the family, and the home . . . The pursuit of this goal led to an intervention strategy that viewed parent contact as the primary intervention, with child care as supplementary, rather than, as most child centered programs of the time were structured, enriched child care as the core of the program and parent contact as outreach. In actual operation, however, both components became crucially important and integrated aspects of the comprehensive and long term intervention.¹³

Within this context, the evaluation is appropriate.

Program implementation. No implementation problems were reported and nearly 80 percent of the program children participated until program completion. No information, however, was provided regarding the child care and other experiences of the comparison group children.

Assessing the randomization. The groups were not randomly assigned.

Assessing statistical controls in experimental and nonexperimental evaluations. The FDRP was not evaluated using random assignment, so the findings should be interpreted with caution. When the program children were thirty-six months old, more than three years after the intervention began, a comparison group was selected. The children were matched on a basis of a number of socioeconomic characteristics, including sex, race, birth order, age, family income, marital status of the mother at the birth of the child, maternal age, and maternal education status (no high school diploma) at the time of infant's birth. The Syracuse team, however, provide relatively little information to assess the initial comparability of the research groups.

It is unclear whether there were any statistical controls for differences in baseline socioeconomic characteristics beyond the initial matching procedure. More important, since the groups were not randomly assigned and participation in FDRP was voluntary, there may have been important unobserved differences between the groups. For example, the parents who enrolled their children in the program may have cared more about their children's future outcomes than the parents of children in the comparison group. Lally, Mangione, and Honig, however, note that the extensive outreach efforts may have muted this concern:

Once a program family was identified as meeting the requirements, they were repeatedly visited until they agreed to participate. There were many disorganized families that required up to 6 months of persistent visits before they agreed to participate. It is highly unlikely that only more motivated families joined the program group. An alternative hypothesis is that the carrot of free child care with the child being picked up and dropped off at home each day actually brought in more disorganized families to the program group

¹³Lally, Mangione, and Honig, 1988, 80.

than to the control group.¹⁴

Nevertheless, the possibility of unmeasured differences between the groups makes potential selection bias a major concern.

Sample size. The full sample consisted of 216 children, evenly divided between the program group and the comparison group. Subsequent attrition, however, reduced the groups considerably. With a small sample, large impacts are needed for statistically significant findings. Thus, the absence of impacts in some areas does not mean that the program did not affect some outcomes, but that the impacts may have been too small to be detected with the sample size. A small sample also means that differences in baseline characteristics would also have to be very large to be statistically significant, making it more difficult to assess the comparability of the program and control groups.

Attrition. The Syracuse FDRP had a very high attrition rate. Of the 108 children who started the program, 82 completed the full five-year intervention, but only 74 of the matched comparison group children remained in the sample this long. (The attrition rate for the comparison group is especially high, considering they were only selected when the children were three years old.)

When the final follow-up was conducted ten years later (when the children were age fifteen), sixty-five program families, representing 60 percent of the original sample, and fifty-four control families, representing just 50 percent of the original sample, were found and consented to have data collected from school records, court records, and probation department records. An even smaller number of parents participated in the survey (fifty-one and forty-two) and still smaller number of children (forty-nine and thirty-nine), representing attrition rates of 50 to 65 percent.

Lally, Mangione, and Honig compared the characteristics of the program families at the time of follow-up to the characteristics of those in the sample at the close of the intervention (when the children were five years old).¹⁵ They did the same for the comparison group. (The characteristics were: child's Stanford-Binet score at forty-eight months old; mother's years of education at the sixty-month interview; mother's age at the birth of the study child; the presence or absence of a father figure in the house; and the family's annual income level when the study child was sixty months old.) They found no statistically significant differences, suggesting that those who remained in the program were similar to those who left. Some attrition, however, had already occurred by the time the program closed. Since program children were assigned before

¹⁴J. Ronald Lally, Peter L. Mangione, and Alice S. Honig, Syracuse University, e-mail message to Peter Germanis, November 12, 2001.

¹⁵Lally, Mangione, and Honig, 1988, 93.

birth, more information on baseline characteristics, characteristics of those leaving, and the final follow-up samples must be provided in order to rule out potential biases due to attrition.

In addition, the degree of attrition varied by outcome and data source. For example, teacher ratings of children's behaviors were limited to 119 of the 216 children in the study and complete questionnaires were obtained on behalf of only 101 children. Similarly, criminal records were available for 119 sample members. No information, however, is provided on how attrition affected the comparability of the program and comparison groups (as well as the original and follow-up groups) using these data sources.

Data collection. The data collection relied on a various standardized tests and parent and teacher surveys. The data sources are appropriate for the questions being studied.

Measurement issues. The interviewers were students who were not informed which group the family was in (program or control group). The use of students rather than professional interviewers could result in higher levels of interviewer bias and other errors. It may also have contributed to the low response rate.

Generalizability. The generalizability of the findings are limited because the intervention was operated as a model program in just one site. Further replication would be needed to see if similar results could be achieved when implemented under normal circumstances.

Replication. The program in its entirety has not been replicated, but various components have been reflected in other programs. For example, many elements of the program are reflected in the design of the Early Head Start program (see chapter 6).

Evaluator's description of findings. Lally, Mangione, and Honig conclude, "The Syracuse Family Development Research Program clearly had a positive impact on the children and families who participated in the intervention."¹⁶ The various shortcomings in the evaluation described above—especially selection bias, small sample size, and attrition—should have led to a more cautious appraisal than suggested by the phrase "clearly had a positive impact." Indeed, the impacts should be considered highly uncertain.

Evaluator's independence. The evaluation was conducted by an independent evaluation team, headed by Peter L. Mangione, with consultation provided by Ronald Lally and Alice Honig.

Statistical significance/confidence intervals. Statistical significance was measured and reported at the 5 and 10 percent levels.

¹⁶Lally, Mangione, and Honig, 100.

Effect sizes. Apparently effect sizes were either not calculated or not reported.

Sustained effects. The evaluation examined impacts through age fifteen, about ten years after program completion.

Benefit-cost analysis. Apparently not performed.

Cost-effectiveness analysis. Apparently not performed.

Commentary

Peter L. Mangione, J. Ronald Lally, and Alice S. Honig*

In their summary and critique of the Syracuse University Family Development Research Program (FDRP), Douglas Besharov and his colleagues state that a fundamental limitation of this research is the lack of random assignment of families to intervention and control groups. However, in some intervention contexts, random assignment is not always the optimal research strategy. It is important to note that extensive consultation with Donald Campbell occurred related to the selection of a quasi-experimental design used in the FDRP. He advised that the comparison group design would be more applicable than a random design because an intensive intervention in the low-income neighborhoods of Syracuse, New York would probably lead to sharing of information and childrearing strategies among intervention and control group members if the selection of the control group was not staggered. Because of not having a randomized sample, two precautions were taken to prevent bias in the direction of the intervention group.

First, intervention and control groups were meticulously matched at thirty-six months on numerous variables that could be used as rival explanations of the results. As reported by Lally and Honig,¹ the intervention and control group members were matched on the following variables:

1. family income of \$5,000 or less at birth of child (\$25,168 in 2005 dollars);
2. mother with less than a high school education at birth of child;
3. mother with no work history or an unskilled or semi-skilled work history at birth of child;
4. father with high school education or less if he was living in the home at birth of child;
5. maternal age at birth of child;
6. sex of child;
7. birth order;
8. birth date of child (within 6 months); and

*Peter L. Mangione is currently the co-director of the Center for Child and Family Studies at WestEd, a nonprofit research, development, and service agency. J. Ronald Lally is currently the director of the Center for Child and Family Studies at the Far West Laboratory for Educational Research and Development in San Francisco. Previously he was at Syracuse University, where he chaired the Department of Child and Family Studies and directed the Family Development Research Program. Alice Honig is professor emerita of child development at Syracuse University.

¹J. Ronald Lally and Alice S. Honig, *The Family Development Research Program: A program for prenatal, infant and early childhood enrichment. Final Report* (Syracuse, NY: Syracuse University, 1977).

9. marital status of parents at birth of child.

Second, potential intervention group mothers who fit the criteria of low income, low education, and none or semi-skilled work history were wooed by home visitors for months until they agreed to participate.

The program aimed to seek out potential high-risk families and serve them. If there was any selection bias, it was in the direction of finding and keeping families in the intervention group that needed the services. The program had less drawing power with prospective highly disorganized control families. Without being able to offer possible control families free day care and family support for five years, the recruiters reported only being able to enroll families at relatively less risk into the control group and have their children tested.

With regard to attrition at age fifteen, if there was any bias, it appeared to be in favor of the control group. During its operation, the FDRP became like a family. Participants formed bonds and stayed in contact with one another. The network of program families helped the follow-up team find many participants who had less successful lives, but were still willing to be interviewed by someone representing FDRP. For the control group, no such network of contacts or allegiance developed. Many of the control families vanished from the community, and the research team had the impression that many of the control families it found who refused to participate in the study did so because they did not want to report negative outcomes. If there was an undetected attrition bias at age fifteen, it is likely to have been in the opposite direction from the direction Besharov and his colleagues suggest.

In the chapter, no mention is made of the interview findings from program parents and children at the fifteen-year follow-up study. The interview responses of program parents and youth were consistent with the low incidence of juvenile delinquency in that group. When asked about what advice they would give to young people today, a significantly higher percentage of program parents than control parents expressed feeling proud about the prosocial attitudes and behaviors of their children. The follow-up study also found that, as compared to control children, program children feel more positively about themselves, envision education being part of their life, and would more likely handle problems directly and actively.² It is noteworthy that, in other studies of early intervention the greater the level of education attained by at-risk youth, the greater the likelihood that they become employed and productive, tax paying members of society.

Although no benefit-cost findings were reported in the fifteen-year evaluation, data on the relative costs of juvenile court services for the program and control groups illustrate the potential

²J. Ronald Lally, Peter L. Mangione, and Alice S. Honig, "The Syracuse University Family Development Research Program: Long-Range Impact of an Early Intervention With Low-Income Children and Their families," in *Parent Education in Early Childhood Intervention: Emerging Directions in Theory, Research and Practice*, ed. Douglas R. Powell (Norwood, NJ: Ablex Publishing Corporation, 1988).

long-range savings of the FDRP. The juvenile court costs for the program group at age fifteen were \$12,111, and those for the control group were \$107,192.³ These costs, of course, exclude those related to the crimes committed by the youths, such as hospitalizations, property damage, and lost work time by victims. A complete benefit-cost analysis was attempted, but was impossible to conduct because of the lack of a reliable estimate of the cost of early childhood services utilized by control families.

The ten-year follow-up study of the FDRP resulted in a substantial number of positive findings for the intervention group in important domains of child functioning. At age fifteen, the program youth had statistically significant fewer convicted juvenile delinquencies. Of the four program youth who were convicted, three were PINS cases rather than the more grave convictions of the control group youth such as armed robbery, rape, and assault. In addition, more control youth than program youth had been convicted as repeat offenders. More program females were on target in school grades and school attendance as compared to control females. Teachers also rated the program girls as functioning better than control girls in the areas of self esteem, feelings toward others, control of aggression toward others, and achievement-related skills. More parents of program youth spoke positively about their youngsters as compared to control parents. And more program youth expressed that they would continue with their education a few years onward, in comparison with control youths' predictions for themselves.

The positive findings at age fifteen for program youth as compared to control youth were consistent across domains. It is noteworthy that there was not one positive finding for the control group children. Yet at birth, these two groups of children started out in similar situations. The program and control groups were matched on important socioeconomic variables at birth, and the follow-up samples were not statistically different from the original program and control group samples. We do not know if, at enrollment, the program parents cared more for their children's future than the control parents. Only random assignment to program and control groups would have completely ruled out that possibility. But we do know that the majority of parents who ultimately enrolled in the program did not rush to FDRP's door. It often took recruiters months to convince parents to participate. The majority of program families were in a state of crisis and not attending to their children's future.

An issue was raised about the qualifications of the interviewers. For the child and family interviews at the ten-year follow up, we used graduate students who were already trained and experienced helping professionals, a widely accepted practice in behavioral science research. We did not inform the students about which group the family was in (program or control) to prevent bias on the part of the interviewers.

We agree with the conclusion that the findings' generalizability is limited because the

³Lally, Mangione, and Honig, 1987.

intervention operated as a model in just one site. Further replication is needed, with a randomization if feasible and appropriate, to see if similar results could be achieved in other low income communities where many youth fail in school and often engage in juvenile delinquency.

Finally, Besharov and his colleagues summarize the finding that, in first grade, as compared to control children, program children's behavior toward adults was mixed, with more positive and more negative behaviors. Honig, Lally, and Mathieson hypothesized that the negative behavior may have resulted because the children received less personalized attention than they expected, leading some to react in negative ways.⁴ This type of finding points to a larger issue. As Brooks-Gunn states in a SRCD policy brief on early intervention: "After an early childhood intervention program ends, poor children are very likely to go to schools that are not conducive to learning. They are likely to live in neighborhoods with relatively few resources. . . . Their parents are more likely to experience discrimination in housing and jobs as well as transportation difficulties."⁵ This description characterizes the type of neighborhood that the FDRP served and where FDRP children spent their middle childhood and early adolescence. In such a context, it is remarkable that only 6 percent of the program children had been convicted of a juvenile crime, and none of the program girls were failing school.

The FDRP was thoughtfully designed for a real world intervention in a low-income community that was to take place for five continuous years. Given that condition, the possibility of contamination of randomized controls by intervention participants in the same community was very high. Therefore, a quasi-experimental design was deemed appropriate, and measures were taken through detailed matching and persistent recruitment to ensure numerous rival hypotheses that might explain away the results were controlled. What resulted from this study based on the reports of families, teachers and the courts was evidence of the development of "nice kids." For that reason alone, the FDRP should be replicated.

Note: This report is open to public comments, subject to review by the forum moderator. To leave a comment, please send an email to welfareacademy@umd.edu or fill out the comment form at http://www.welfareacademy.org/pubs/early_education/chapter26.html.

⁴Alice S. Honig, J. Ronald Lally, and Deborah H. Mathieson, "Personal and Social Adjustment of School Children After Five Years in the Family Development Research Program," *Child Care Quarterly*, 11 (2) (1982): 136–146.

⁵Jeanne Brooks-Gunn, "Do You Believe in Magic?: What We Can Expect From Early Childhood Intervention Programs," *Social Policy Report*, XVII (1): 3–14.