DOCUMENT RESUME

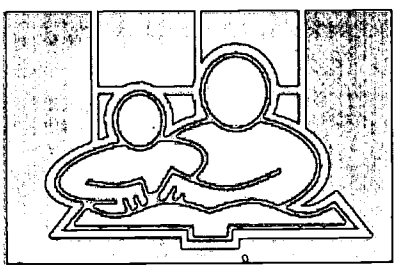
ED 418 827 RC 021 425

AUTHOR Meehan, Merrill L.
TITLE Evaluation of the Monongalia County Schools, Even Start Family Literacy Program. Monongalia County Even Start Program. Parents Are Their Children's First and Most Important Teachers.
INSTITUTION Appalachia Educational Lab., Charleston, WV.
PUB DATE 1997-09-00
NOTE 139p.
PUB TYPE Reports - Evaluative (142)
EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS Adult Literacy; Agency Cooperation; *Family Literacy; Family Programs; Literacy Education; Models; Parent Education; Parents as Teachers; *Preschool Children; Preschool Education; Program Evaluation; *Rural Education; *School Readiness
IDENTIFIERS *Even Start; *Monongalia County Schools WV

ABSTRACT

The Monongalia County (West Virgina) Even Start Program begun in 1993 to address family literacy and to prepare children for entrance into school, consists of three interrelated approaches: adult education, parenting education, and early childhood education. In June 1996, the program was serving 62 families (mostly low-income) with an average of 2.2 children. The unique feature of the Monongalia program is that it is mainly a rural, home-based program--home visits are the primary service delivery mechanism. This detailed evaluation of the program covers the years 1994-96. The introduction includes the history of the national Even Start program, including its emphasis on families, services to eligible families, and the first and second national evaluations; site specifics, approach, and interagency collaboration of the Monongalia program; objectives of the evaluation; and audience. The section on methodology addresses national evaluation population characteristics, the Monongalia sample study, data collection techniques (adult and child), and the data analysis process. Program component findings cover program model, recruitment, home visits, early childhood education, adult education, parenting education, interagency collaboration, evaluation, and staff development. Among the participants' educational and developmental outcomes, the program was found to have a tremendous impact on school readiness; large impact on adult reading literacy, mathematics literacy, and parenting; and some impact on children's auditory comprehension, expressive communication, and total language development. Recommendations relate to program continuation, improvements, and evaluations. Contains 20 references. Appendices include 1995-1996 program statistics, the home visit report sheet, the innovation configuration matrix (program evaluation model), a staff self-assessment report; and the completed program evaluation standards checklist. (SAS)

Evaluation of the Monongalia County Schools Even Start Family Literacy Program



Monongalia County Even Start Program

Parents are their children's first and most important teachers

Merrill L. Meehan

AEL
APPALACHIA EDUCATIONAL LABORATORY INC

Evaluation of the Monongalia County Schools, Even Start Family Literacy Program

by:

Merrill L. Meehan
Senior Research and Evaluation Specialist

September 1997

Appalachia Educational Laboratory
Planning, Research, and Evaluation Unit
Post Office Box 1348
Charleston, West Virginia 25325

AEL's mission is to link the knowledge from research with the wisdom from practice to improve teaching and learning. AEL serves as the Regional Educational Laboratory for Kentucky, Tennessee, Virginia, and West Virginia. For these same four states, it operates both a Regional Technology Consortium and the Eisenhower Regional Consortium for Mathematics and Science Education. In addition, it serves as the Region IV Comprehensive Technical Assistance Center and operates the ERIC Clearinghouse on Rural Education and Small Schools.

Information about AEL projects, programs, and services is available by writing or calling AEL.

Appalachia Educational Laboratory
Post Office Box 1348
Charleston, West Virginia 25325-1348
304/347-0400
800/624-9120 (toll-free)
304/347-0487 (Fax)
aelinfo@ael.org
http://www.ael.org

This publication is based on work sponsored wholly or in part by the Monongalia County Schools, West Virginia. Its contents do not necessarily reflect the views of Monongalia County Schools.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF FIGURES

EXECUTIVE SUMMARY

The Even Start Family Literacy Program, started in 1989, is a national literacy program that identifies and addresses the developmental and educational needs of targeted families. The developmental and educational needs of eligible Even Start families are addressed in three different ways. First, the literacy and educational needs of the adults in the family are identified and addressed. Second, effective parenting skills for the adults are taught. Third, the preschool readiness needs of the children in the family are addressed with at-home and/or center-based early childhood education sessions. States receive their funding based on a formula related to the allocation they receive from the federal government for their Title I program. States then solicit grant applications from local agencies and make awards on a competitive basis.

The Monongalia County (WV) Even Start Program began in 1993. At the law requires, the program consists of three interrelated approaches to address family literacy and to prepare children for entrance into schooling: adult education, parenting education, and early childhood education. The goals of the program are:

Adult Education -	to provide and strengthen opportunities that will increase the literacy and education of parents;
Parenting Education -	to help parents gain competence in effective parenting practices that will maximize their child's development and assist them in becoming full partners in their child's education; and
Early Childhood Education -	to assist children in reaching their full potential as learners.

The unique feature of the Monongalia County Even Start Program is that it is mainly a home-based program. While the law requires that all Even Start programs conduct **some** home visits, in the Monongalia County program, home visits are the **primary** service delivery mechanism. Three Family Educators work with their assigned families to develop and schedule a plan of service. Most families are visited every other week in the homes by the Family Educator. Typically, these bi-weekly home visits last about one and a half to two hours each and include some adult education, some parenting education, and some early childhood education. A home visit usually begins with the collection of books and associated toys loaned during the prior visit and then moves into a learning activity with the child or children. Next, while the children are occupied, the Family Educator discusses child development issues or parenting concerns, such as discipline. This is followed by some adult education. The home visit concludes with an inquiry about other needs the family might have that might be hindering progress in educational or other areas and scheduling the next visit.

Evaluation Design

The Monongalia County Even Start Program staff contracted with the Appalachia Educational Laboratory (AEL) in Charleston, West Virginia, to conduct an independent, third-party evaluation

of their program. This AEL evaluation was designed to analyze much quantitative data already collected by program staff and the randomly-drawn Sample Study. Also, since the national evaluations seemed to report mainly on nonrural, center-based Even Start programs, a part of this evaluation was to determine, depict, and describe the unique aspects of Monongalia County's mostly rural and home-based Even Start program.

Four objectives guided this Even Start program evaluation. They were:

1. to identify (1) the essential components of the Monongalia County Even Start Program, and (2) the implementation variations of each component;
2. to evaluate the implementation of each of the Monongalia County Even Start Program's essential components;
3. to assess the adult and parenting education outcomes of the program's Sample Study adults, and
4. to assess the early childhood education outcomes of the program's Sample Study children.

The Sample Study being completed by Abt Associates consists of 57 randomly-selected projects from the national evaluation's Universe Study of 490 state-administered Even Start projects. More detailed participation and outcome data were collected from parents and children in the Sample Study than from parents and children in the Universe Study. The Monongalia County program was one of the rural sites in the Sample Study. Families in the national sample had to meet additional criteria: This dramatically reduced the Monongalia County subsample to 16 families out of the 62 in the program. Data were collected over two years from each family. Eight of those subsample families enrolled in September-November of the 1994-1995 program year and the other eight families enrolled in September-November of the 1995-1996 program year. There was some attrition in the Monongalia County subsample.

Data for this evaluation were collected by a variety of ways, including program records, face-to-face interviews, observation of events, and seven participant outcome measures. Program records included home visit reports, Even Start Information System reports, staff self-assessment ratings, and records of all staff development sessions attended. Interviews with program staff, collaborators in other agencies, and clients were completed in a one-week site visit, as were observations of two program activities. The Comprehensive Adult Student Assessment System (CASAS) Life Skills test was used to assess adults' reading literacy and mathematics literacy, while the Home Screening Questionnaire (HSQ) was used to assess their parenting skills. The PreSchool Inventory (PSI) was employed to assess school readiness skills of children and the Preschool Lanuage Scale-3 (PLS-3) was selected to assess their language development. The PLS-3 yields three scale scores: Auditory Comprehension, Expressive Communications, and Total Language. For each family in the subsample, one adult (all mothers) and one child were subjects in this evaluation.

Data analyses were conducted on both the qualitative and quantitative types of data. The Innovation Configuration Matrix (ICM) was constructed from program records, interviews, and

observations. Next, the eight essential components in the final ICM were evaluated with those same data sources. Even Start service intensity data were computed from staff-completed records, including (1) home visits, (2) adult education contact hours, (3) parenting education contact hours, and (4) early childhood education days attended. Descriptive statistics were obtained for (1) each of the three program years in the study, (2) those three program years combined, and (3) the two treatment years for each family. Treatment effects data were the three obtained scores (pretest, posttest, and follow-up) on the three adult measures and the four child measures. Statistical analyses of the outcome measures were t-tests for dependent means by the matched pairs technique. Next, effect sizes were calculated for each t-test. Then, items or points per treatment month gained (or lost) were computed. Graphic displays of the outcome measures were designed. Last, the local subsample's results were compared to those of the national sample.

Findings

The findings from this evaluation are presented in two sections--program components and participants' educational and developmental outcomes. The findings narrative is supported by 21 tables and 8 graphic displays. Highlights from these two sections follow.

The completed ICM for the Monongalia County Even Start Program consists of eight essential components, each with four possible variations of implementation (A through D). The eight essential components are (1) identification and recruitment, (2) home visits, (3) early childhood education, (4) parenting education, (5) adult education/literacy, (6) collaboration, (7) evaluation, and (8) staff development. The program was judged to be operating at the ideal, or "A," variation on all eight components. The essential components of the program can be arranged in six concentric circles to depict the model of the program. The third circle contains components #3, #4, and #5.

For the CASAS Reading Literacy instrument, all of the mean scores across the three administrations were above the national high school level. Still, the subsample adults gained one and a half standard score points in the first treatment year and almost three and a half points in the second treatment year. More dramatically, over the two treatment years combined, the mean score gained nine points (for fewer adults with all three scores), which was statistically significant at the .001 level. At 0.818, this effect size was "large," while it was "small" for each year individually.

On the CASAS Mathematics Literacy instrument, the subsample's pretest mean score was about 10 points below the national high school level, but it rose 6.75 points in the first treatment year, statistically significant at the .001 level. There was a 4.25 point gain in the second year (not significant) and a 12.38 point gain over the two-year treatment period, which was statistically significant at the .001 level. The effect sizes for the three gains were "medium," "small," and "large."

Regarding adults' parenting outcomes, the mean HSQ score increased 3 points (56-point maximum) and was statistically significant at the .001 level. There was a 1.43 gain in the second year, which was not statistically significant. However, over the two-year treatment period, there was a

4.14 point increase, which was statistically significant at the .002 level. The effect size for each of the treatment years individually was "small," but for both years combined, it was 0.805, which was "large."

Moving to child outcomes, the 32-point PSI instrument assessed children's readiness for school. In the first year, the mean score increased 8 points to just 4 points below the maximum. This gain was statistically significant at the .001 level, and its effect size of 1.409 was labeled "very large." The gain in the second year was less than half of the first year, but it was statistically significant at the .002 level and its effect size was "medium." Over the two treatment years, the gain was almost 11 points, which was statistically significant at the .0001 level. At 1.840, the two-year effect size was labeled "extremely large." Also, the PSI points gained per treatment month (0.764) was almost double that expected for normal growth.

For the PLS-3 Auditory Comprehension scale, the mean was under the national norm, but at follow-up, it was 4 points above the norm. The mean score increased 3.46 points in the first treatment year, which was not statistically significant and its effect size was "small." The 6-point gain in the second treatment year was above the national norm, but again, it was not statistically significant and its effect size was "small." When combined (for children with three scores), both treatment years yielded a 10.1 gain, and although it was not statistically significant, the effect size of 0.602 was "medium." The two-year gain was .710 points per treatment month on the 150-point maximum score.

With respect to the PLS-3 Expressive Communications scale, the pretest mean was less than 1 point under the 100-point national norm, but by follow-up, it was 7 points above the national norm. The mean score increased 7.62 points in the first treatment year, which was not statistically significant and had a "small" effect size. Unexpectedly, the mean score dropped 3.36 points in the second year (the only decline in the evaluation), which was not statistically significant. When combined, the two treatment years yielded a 4-point gain to above the national norm, but it was not statistically significant and its effect size was "small."

Regarding the PLS-3 Total Language scale, the mean was a little more than 4 points under the norm at pretest, but moved up to more than 8 points above the norm at follow-up. The mean score increased nearly 4 points in the first year, which was not statistically significant and had a "small" effect size. In the second year, the mean score gained 5.6 points and, although not statistically significant, the 0.513 effect size was "medium."

Finally, another way to summarize the adult and child outcomes in this evaluation was to average the two-year (pretest to follow-up) effect sizes. The average two-year effect size for the three adult outcome measures was 0.893, which is "large" in social science research. The average two-year effect size for the four child outcome measures was 0.823, which also is "large."

Conclusions

Thirteen conclusions were drawn from the evaluation findings. First, the final Innovation Configuration Matrix of the Monongalia County Even Start Program can be employed to describe

the program to others but, moreover, to differentiate it from others on the basis of where other programs are operating with respect to the four variations for each of the eight essential components.

Second, it is concluded that the graphic depiction of the eight essential components in six concentric circles is an accurate portrayal of the Monongalia County Even Start Program.

Third, it is concluded that the Monongalia County Even Start Program had a large impact on adults' Reading Literacy, especially over the two-year treatment period.

Fourth, it is concluded that the Monongalia County Even Start Program had a large impact on adults' Mathematics Literacy, especially in the first year in comparison to the second year, and when the two years were combined.

Fifth, it is concluded that the Monongalia County Even Start Program had a large impact on adults' parenting outcomes, especially in the first year in comparison to the second year, and when the two years were combined.

Sixth, it is concluded that the Monongalia County Even Start Program had tremendous impact on children's readiness for school outcomes, especially in the first year in comparison to the second year, but even more so when the two treatment years were combined.

Seventh, it is concluded that the Monongalia County Even Start Program had some impact on children's Auditory Comprehension outcomes, especially over the two-year treatment period rather than either year individually.

Eighth, considering the up and down movement of the mean scores, it is concluded that the Monongalia County Even Start Program had some impact on children's Expressive Communications outcomes over the two-year treatment period.

Ninth, it is concluded that the Monongalia County Even Start Program had some impact on children's Total Language development outcomes over the two-year treatment period and, further, more so over the second year in comparison to the first year.

Tenth, it is concluded that the Monongalia County Even Start Program impacted adults' CASAS Mathematics Literacy and children's readiness for school (PSI) and PLS-3 Expressive Communications outcomes the most for the first treatment year. For the second treatment year, the program impacted the children's PLS-3 Auditory Comprehension and Total Language outcomes the most; and for both years combined, the program impacted adults' CASAS Mathematics Literacy and children's readiness for school (PSI), and PLS-3 Total Language outcomes the most.

Eleventh, it is concluded that the locally- and nationally-developed instruments employed in this evaluation of the Monongalia County Even Start Program were both useful and effective. However,

it was noted that three of the subsample children reached the ceiling score on the PSI at the end of their second year in the program.

Twelfth, it is concluded that, overall, the Monongalia County Even Start Program is a comprehensive, rural, home-based family literacy program that produced many positive impacts on the adults and children studied in this effort. It is concluded that the Monongalia County staff have designed and implemented an effective rural, home-based Even Start program for its clients.

Thirteenth, in the area of needed improvements, it is concluded that there is insufficient data on the fathers of the families in the Monongalia County Even Start Program to make any statements of impact on them. There is not enough service delivery nor outcome data on the fathers to analyze at the level of the mothers in the families.

Recommendations

Recommendations from this evaluation were presented in four major headings of (1) program continuation, (2) program improvements, (3) evaluation utilization, and (4) new evaluations.

Regarding program continuation, it was recommended that the Monongalia County Even Start Program be continued in future years; that the collection of adult and child outcome measure data be continued; that the staff's self-assessment activities continue; that the staff consider expanding their program to other families, if additional resources can be found; and that staff should continue to identify, implement, and evaluate new/different diagnostic and outcome measures with adults and children.

In the area of program improvements, it was recommended that program staff should strive to improve the children's PLS-3 Auditory Comprehension scores; strive to get more fathers in the families involved in the events, sessions, activities, and other components; continue to develop original curriculum materials for the program; implement the "collaboration" part of their own action plan first, and consider using the 64-item PSI instrument instead of the 32-item version.

With respect to evaluation utilization, it was recommended that this evaluation report be submitted to the national ERIC system for accession; that this report be disseminated locally, regionally, and nationally both within and outside the Even Start system; that the ICM and graphic model of the program be disseminated widely; and that local staff should use this evaluation report or relevant sections to seek some sort of national recognition for the Monongalia County Even Start Program.

Finally, regarding new evaluations, it was recommended that program staff collect and report on case histories of Even Start participants who have been especially successful; if the opportunity arises, staff should participate in the next national evaluation; and this evaluation should be repeated in two years with two additional cohorts of Even Start families and the same or comparable instrumentation.

INTRODUCTION

Large numbers of families and their children are at risk of not succeeding in this rapidly-changing, information-age society. Many complex and interrelated factors have placed these families and children at risk of not making their way in the new-world, global-economy society. Whatever those factors are, the government of the United States has proposed to overcome them by the next millennium. For example, in its ambitious Goals 2000 Project, the U. S. President and Congress have committed to moving the country to be first in the world in several areas related to education, including mathematics and science achievement (Public Law 103-446, Section 102, paragraph 5).

However laudatory the goal of being first in the world in mathematics and science achievement is, families and children at risk have much more basic and fundamental problems to overcome. Two other goals in the Goals 2000 project address these needs: (1) all children in America will enter school ready to learn, and (2) every adult will be literate and will possess the skills necessary to compete in a global economy and exercise the rights and responsibilities of citizenship (Public Law 103-761). Among the various government programs designed to meet needs of families and children at risk, one special program was designed to work with the family as a unit via a family literacy improvement model. That program is the Even Start Family Literacy Program.

This section presents an overview of the National Even Start Program, including its basis in the law, a description of the Monongalia County Even Start Program, the purpose and objectives of this evaluation, and an explanation of the audience for this report.

The National Even Start Program

The Even Start Family Literacy Program is a national literacy program that identifies and addresses the educational and developmental needs of targeted families. The educational and developmental needs of eligible Even Start families are addressed in three different ways. First, the educational and literacy needs of the adults in the family are identified and addressed. Second, effective parenting skills for the adults are taught. Third, the developmental and preschool readiness needs of the children in the family are addressed with at-home and/or center-based early childhood education sessions.

The national Even Start program began in 1989. Legislative authority for the program is in the Hawkins-Stafford Elementary and Secondary School Improvement Amendments of 1988 to the Elementary and Secondary Education Act of 1965. Specifically, the program is Part B, Chapter 1 of Title I in Public Law 100-297 (Tao, Swartz, St. Pierre, & Tarr, 1997). Two years later, the original Even Start legislation was amended with the passage of the National Literacy Act (Public Law 102-73). Reauthorization for the Even Start program was provided by the Improving America's Schools Act as Part B of Title I. Tao, et al. (1997, p. 2), cite the main intention of the Even Start program in the law as:

> ...help break the cycle of poverty and illiteracy by improving the educational opportunities of the Nation's low-income families by integrating early childhood education, and parenting education into a unified family literacy program....The program shall (1) be implemented through cooperative projects that build on existing community resources to create a new range of services, (2) promote achievement of the National Education Goals, and (3) assist children and adults from low-income families to achieve the challenging state content standards and challenging state student performance standards (P.L. 103-382, Sec. 1201).

Emphasis on Families

As stated in the legislation, the purpose of the Even Start program is to "help break the cycle of poverty and illiteracy" by working with low-income ***families*** to improve their educational opportunities. Then, the legislation spells out those educational opportunities as being adult basic education or adult literacy, parenting education, and early childhood education. Further, the Even Start legislation requires that these three educational areas be integrated "...into a unified family literacy program" (P.L. 103-382, Sec. 1201). The Even Start program, by legislative mandate, represents an innovative mixture of adult literacy or adult basic education, parenting education, and early childhood education.

With parents and their children as the family unit, this Even Start program has three main goals:

1) to help parents improve their literacy or basic skills,

2) to help parents become full partners in the education of their children, and

3) to assist children in reaching their full potential as learners (Tao et al., 1997, p. 2).

Criteria for eligibility into the Even Start program was established in the legislation. To be eligible for participation in the program, families had to meet two criteria. Participants had to be (1) a parent or parents who (A) are eligible for participation in an adult basic education program under the Adult Education Act, or (B) are within the state's compulsory school attending age ranges; and (2) the child or children, from birth through age seven, of the parent or parents described in (1). The (B) part of the parent eligibility requirement were new in 1995-96, which expanded the Even Start program to include teen parents for the first time.

Program Services

Even Start program services provided to eligible families can be divided into two types. Those two types of services to parents and their children are (1) "core" educational services, and

(2) ancillary support services. As mandated in the Even Start legislation, the "core" educational services are in three interrelated areas of:

(1) high-quality instructional programs that promote adult literacy,

(2) high-quality instructional programs that empower parents to support the educational growth of their children, and

(3) developmentally-appropriate early childhood educational services and preparation of children for success in regular school programs (P.L. 103-761, Sec. 1205).

The legislation also requires that Even Start programs provide these core services to parents and children jointly and, also, to provide some home-based services (more on this below).

Even Start programs are required to provide support services to participating families. The twin support services of child care and transportation for the purposes of enabling parents and their children to participate in Even Start programs are named specifically in the law. Local programs typically provide a range of other support services to their participating families including health care, vision and hearing screening, employment assistance, mental health referrals, cross-agency assistance, nutrition education, referrals for screening or treatment for chemical dependency, and public library card attainment assistance. Through its interagency collaborative arrangements, other support services are available to Even Start families.

In their 1995 Interim Report of the *National Evaluation of the Even Start Family Literacy Program* (1997, p. 6) authors Tao et al. described the history of the program. Following the passage of the law authorizing it, Even Start initially began as a demonstration program administered by the U. S. Department of Education (USDE). The USDE funded 76 Even Start projects in fiscal year 1989 for a total of $14.5 million. These first Even Start projects were four-year discretionary grants to local school districts to implement the new family literacy programs. In 1991, the Even Start program expanded to the level of $49.7 million, still administered by the USDE. Then, in 1992, 100 new Even Start projects were started, totaling $70.0 million. And, according to the law, since the total budget for the program exceeded $50 million, its administration was transferred to the states. Tao et al. (1997, p. 6) report that about 80 new projects were begun in both 1993 and 1994, now administered by the states. This brought the total number of Even Start projects in the nation to approximately 450 in 1993 and over 500 in 1994. The number of projects is expected to level off at about 550 after 1995 (p. 6).

The states receive their funding based on a formula related to the allocation they receive from the federal government for their Title I Program. States then solicit grant applications and make awards on a competitive basis. The law provides a portion of the state's allocation for state-level administration and management. The law also stipulates that the minimum award for an Even Start project is $75,000 (except one project per state can be for less than that dollar amount). California had the most Even Start projects participating in the national evaluation study with 44, followed by

Texas with 35 and New York with 31. Delaware, District of Columbia, and Hawaii each had just one Even Start project (Tao et al., 1997, Appendix B-1,2). Additionally, the USDE continues to administer special programs for migrant families, Indian tribes, and women and children in prison with set-aside funds authorized in the law.

First National Evaluation

Evaluation of the Even Start program has been required in the authorizing legislation. Two national-level evaluations of the Even Start program have been awarded—one has been completed and the second is underway. Section 1209 of the current Even Start statue is titled "Evaluation" and states:

> From funds reserved under Section 1202(b)(1), the Secretary shall provide for an independent evaluation of programs assisted under this part - (1) to determine the performance and effectiveness of programs assisted under this part; and (2) to identify effective Even Start programs assisted under this part that can be duplicated and used in providing technical assistance to federal, state, and local programs. (P.L. 103-761, Sec. 1209).

The first Even Start national evaluation began in January 1990 and ended in April 1994. The USDE awarded the evaluation contract to Abt Associates as the primary organization responsible for the evaluation, although assistance was provided on special aspects of the evaluation by RMC Research Corporation and Pelavin Associates as subcontractors. The first national evaluation consisted of four major components: (1) the National Evaluation Information System for all Even Start projects, (2) an in-depth study of ten selected projects, (3) other local evaluations conducted by individual Even Start projects, and (4) assistance to local projects to apply for and enter the (then-existing) National Diffusion Network (St. Pierre, Swartz, Gamse, Murray, Deck, & Nickel, 1995, pp. 12-13).

As required by the law, annual reports of the first Even Start national evaluation were submitted to the USDE and to Congress (St. Pierre, Swartz, Murray, Deck, & Nickel, 1993). The was completed in 1994 and its report was published in 1995 (St. Pierre et al.). This report describes the various types of Even Start projects funded, the various services that they provided to participating families, the collaborative interagency activities they completed, and the numbers and types of problems and obstacles that the local Even Start programs met in their implementation efforts. The first evaluation report described the characteristics of participating Even Start families, the nature and number of different services the families received, and the effects of the program parents' literacy, parenting skills, children's readiness for school, and family stability and resources.

For example, among their responses to the research question of, "How well does the basic Even Start model work?," the authors wrote:

- Participation in adult education services led many Even Start adults to attain a GED certificate.
- Participation in early childhood education led Even Start children to attain school readiness skills sooner than they otherwise would have.
- Greater parents' participation in parenting education is related to an increase in their children's vocabularies.
- Even Start did not produce any measurable effects on parenting skills, family resources, income, or employment. (St. Pierre et al., 1995, pp. 255-256).

The descriptive results in the final report of the first Even Start national evaluation were based on all 340 projects funded from 1989 to 1992, while the program effects results were based on data from 120 projects funded in the first two years only.

Second National Evaluation

The second Even Start national evaluation began in March 1994 when USDE awarded the contract to Fu Associates as the primary evaluator, with Abt Associates as the subcontractor. Like the first evaluation, the second national evaluation was four years (1993-1997), with interim reports and, also, it consisted of four major questions. These major evaluation questions are:

(1) Who is served by the program and what services do they receive? Is the program reaching the appropriate target population?

(2) How is the federal funding spent on the program? How many of the projects are well implemented?

(3) How well does the Even Start basic model work? Do participants perform better on key measures than similar persons who do not participate?

(4) What are effective practices and programs? (Tao et al., 1997, pp. 11-12)

Also like the first national evaluation, the second evaluation is made up of four major components—two similar and two different from the first evaluation. The first component is the Even Start Information System (ESIS) designed to collect relevant data from all projects and which was a refinement of a data collection system implemented in the prior national evaluation. The

second component is called the Sample Study, which is designed to collect data from 57 randomly-selected projects. The third and fourth components are site visits planned for migrant education and Indian tribe Even Start projects. These site visits to special USDE administered-projects were not completed in the first evaluation study.

Two interim reports of the second national Even Start program have been produced at this writing. The first interim report, for the 1993-1994 year, focused on data collected at the project-level only (439 out of the 490 projects). This report (for 1994) provided much detailed information on the projects, but not on families, because the new data collection system was not implemented fully (Tao et al., 1997, p. 16).

The second interim report of the second national evaluation, for program year 1994-1995, did contain information from both projects and participating families. Of the total of 513 "regular," migrant, and Indian/Tribal education Even Start projects, the 1995 interim report (published in January 1997) contained data from 476 (93%) of them. These projects serviced 28,534 families; 36,752 parents; and 52,757 children (Tao et al., 1997, p. 17). Those numbers were for the population of projects. The Sample Study consists of 57 Even Start projects, for which additional outcome measures were administered. The outcome measures were administered on a pretest-posttest basis to approximately 450 children and parents in the Sample Study (although the parent's outcomes in adult education areas were split by the choice of which one of two possible instruments the project selected for them to complete. However, the number of adults and children that had ***both*** pretest and posttest data on their outcome measures was about half of those that completed the pretests. All of the adult and child outcomes reported in the 1995 interim report are relevant to this particular evaluation effort and, thus, will not be discussed in detail in this section, but will be later. Generally, though, the authors of the 1995 report concluded that the Even Start services to families "are helping to improve parents' academic skills and children's language development" (1997, p. x).

The relevance and importance of the 1995 interim report of the second national Even Start program to the current evaluation is very direct. This is because (as shall be explained in more detail below) this evaluation is of one of the 57 projects selected at random to be in the national evaluation's Sample Study. Therefore, most of the adult and children outcomes data collected and submitted to the national evaluation also were utilized for this evaluation of the local Even Start project. This evaluation benefitted also by having the pretest-posttest data in the national evaluation report for comparison purposes. For this, we are grateful and acknowledge our appreciation here. On the other hand, the posttest and follow-up data results for the national evaluation were not available at this writing, which hindered our local project results to national sample results comparisons. Nonetheless, we are able to display and discuss the within-group differences from pretest and posttest to follow-up test over two treatment years for the adults and children in the local project.

The Monongalia County Even Start Program

Program Site Information

Monongalia County, West Virginia, is located in the north central part of the state, bordering the Pennsylvania state line. The population of the county, according to the 1990 U. S. Census, was 75,599 in its 417.85 square miles, or 180.92 residents per square mile. The county seat is the city of Morgantown, the fifth largest city in West Virginia. Approximately 34% of the total county population lives in Morgantown, with the remaining 66% living in the outlying, mostly rural and mountainous communities (Even Start Staff, 1995, p. III-3). Morgantown also is the home of West Virginia University, the largest university in the state.

Monongalia County School System had 10,079 students in the 1996-1997 school year in grades K to 12. There are 28 schools in the system and 598 teachers (Quality Education Data, 1996, p. 14). In 1995, 3,615 (35%) of the system's children qualified for free and reduced lunches (Even Start Staff, 1995, p. III-3). Like most of the state, the vast majority (93%) of the school children are Caucasian. It is reported that 54% of the system's students are college-bound, while the dropout rate of 17.5% is very close to the state's average of 17.7% in 1995. The total number of computers available for instructional use was reported as 939, of all brands (Quality Education Data, 1996, p. 14). Each school does have a Local School Improvement Council and a Faculty Senate. The predominant organizational arrangement for the 20 elementary schools is K-5 or 6, although there is one K-3 school, two K-4 schools, and one school with no kindergarten. There are two middle schools (one grades 5-8 and one grades 6-8), two middle or junior high schools of grades 7-9, one small grade 7-12 school of 415 students, two large high schools (one grade 9-12 and one grade 10-12) of over 1,100 students each, and one vocational-technical education center. The adult education center is housed in the vocational-technical center. The Monongalia County Even Start Program is housed in the Dorsey Education Center, a renovated school building that also houses the administration offices for the Title I Program, the Head-Start Program, the Early Head Start Program, and two Head Start classrooms.

There is poverty, unemployment, and illiteracy in Monongalia County. The district's average of low-income families is 34%. The total caseload of families receiving health and/or human services in the county was reported as 5,174 in July 1994 (Even Start Staff, 1995, p. III-3). The number of persons unemployed in 1995 was reported as 2,580, which placed the county in the top seven in the state in this category. Data from the 1990 U. S. Census show that 22,562 adults in the county have not completed high school and, further, there are 12,003 persons 18 years and over with less than a 9th grade level of education. The Head Start Program has identified 236 unserved families who meet Headstart's income eligibility guidelines and could benefit from participating in the Even Start program.

Program Information

The Monongalia County Even Start Program began in 1993 and has operated since then. As the law requires, the program consists of three interrelated approaches to address family literacy and to prepare children for entrance into schooling. These three approaches are adult education, parenting education, and early childhood education. The Even Start staff (1995, p. III-1) explain their goals as:

Adult Education -	to provide and strengthen opportunities that will increase the literacy and education of parents;
Parenting Education -	to help parents gain competence in effective parenting practices that will maximize their child's development and assist them in becoming full partners in their child's education; and,
Early Childhood Education -	to assist children in reaching their full potential as learners.

The Even Start program in Monongalia County is designed to provide simultaneous education services to participating parents and their children. The program recognizes that many families lack the background, skills, and support necessary to help their children learn and be ready to enter school. The program has been designed to inform and educate parents on how important their involvement is in their child's education and then to teach parents how to achieve this goal. The Monongalia County program has received training in the Parents As Teachers (PAT) and has adapted it for its parenting education curriculum. The PAT is a nationally-validated home-school partnership program designed for 0-5 year old children. Its purpose is to provide the training and curriculum to support parents in giving their children the best possible start in life. PAT's philosophy is that parents are a child's first and most influential teachers (Even Start Staff, 1995, p. III-1).

In addition to the PAT curriculum, the Monongalia County Even Start program has developed and implemented its own family literacy model. This locally-developed family literacy model emphasizes children' literature in fostering cognitive and educational growth in the family unit. A unique component of this model is that the program has developed an extensive lending library of children's books; developmentally-appropriate toys; and other learning materials such as crayons, paints, markers, and art paper. These literacy development materials are loaned to Even Start families during regular, planned home visits. Parents are encouraged to help select age-appropriate items for the Family Educator to bring on the next home visit. For example, a young boy might choose a book such as *Tommy The Truck* to read and a Tonka truck might be brought for him to play with before the next home visit by the Family Educator.

The unique feature of the Monongalia County Even Start program differentiating it from most other programs is that it is primarily a home-based program. While the law requires that all Even Start programs conduct ***some*** home visits, in the Monongalia County Program, home visits are the ***primary*** service delivery mechanism. The three Family Educators work with their assigned families

to develop and schedule a plan of service. Most families are visited every other week in their homes by the Family Educator. Visits includes some adult education, some parenting education, and some early childhood education. Typically, these bi-weekly home visits last about two hours each, but they can vary according to the family's needs and other factors. As an example, in those families where the adult is attending adult education instruction at the county's adult education learning center, then the amount of such instruction in the home visits may be reduced. During these bi-weekly home visits, the Family Educator provides meaningful instruction to both the child or children present and to the parent. Typically, a home visit begins with the collection of books and toys loaned during the prior visit, and moves into a learning activity with the child or children, that focuses on literacy. Next, while the children are occupied, the Family Educator discusses child development issues or addresses parenting concerns, such as discipline (Even Start Staff, 1995, p. III-2).

Another aspect of the home visits is that the Family Educator completes assessments of other family needs. Here, the occasion of being in the home is used to inquire and talk about other needs that the family might have that might be hindering progress in educational or other areas. Examples include vision and hearing screening; assistance in securing family benefits; assistance in looking for and obtaining jobs; help in receiving treatment for chemical dependency; and help in enrolling and transporting children to Head Start, kindergarten, or other school programs. When appropriate, the Family Educators will make referrals to other agencies in the county to which the Even Start program is linked—many in a formal, collaborative manner. Another purpose of the home visits in the Monongalia County model is to foster communications between the schools and the participating families. Also, whenever it can be arranged, the Family Educators encourage and foster communications across the participating families and sponsor events and activities for all or parents or subgroups of parents with similar needs. A group meeting at the local public library, where parents received their own library card, is one example.

Interagency Collaboration and Evaluation

Interagency cooperation and collaboration is another feature of the Monongalia County Even Start Program. This program has formed an Even Start Advisory Council to connect with outside agencies in the county. Each fall, collaborators review the past year's program and plan for the coming year. Partnerships with school-related groups include Monongalia County Schools, Head Start, Early Head Start, Title I, Special Education, Technical Education Center, and the Adult Education Center. Close connections to all of the school system's federal projects for children (Head Start, Early Head Start, Title I, and Even Start) are assured because one administrator supervises all of them and also, because all of them are headquartered in the same building. The Even Start program also collaborates with the Monongalia County Interagency Council, the Valley-Comprehensive Mental Health Program, the Monongalia County Health Department, the Shack Neighborhood House Happy School, the Rock Forge Neighborhood House Happy School, Right From The Start, Christian Help, Stepping Stones, Family Service Association, The Caring Program for Children, Literacy Volunteers, The Department of Health and Human Resouces, and the Sabraton Adult Reading Class. The program also collaborates with the West Virginia Department of Health and Human Resources, West Virginia University, and the Morgantown Public Library.

Evaluation is another important aspect of the Monongalia County Even Start Program. In its first two years, program staff contracted with doctoral students at West Virginia University to conduct two independent evaluations of the program. Both of these were process-oriented evaluations and involved mostly qualitative-type data and employed data collection techniques such as in-depth interviews with adults, children, and staff and observations of home visits made by the Family Educators. This present evaluation of the program also is being completed by a third-party evaluator, independent of the Even Start program. This evaluation differs markedly from the prior outside evaluations in that it is product-oriented and involves primarily quantitative-type data, although one important portion of this evaluation required face-to-face interviews.

All Even Start programs are asked to submit evaluation data as part of the national evaluation's Universe Study being administered by Fu Associates under contract to the USDE. But, the Monongalia County program was selected at random to be one of the 57 programs to comprise the Sample Study being conducted by Abt Associates under a subcontract to Fu Associates. Participation in the Sample Study requires the collection of additional outcomes-oriented data from families that meet several special criteria. Monongalia County program staff decided to collect these additional data from all its families, in addition to those meeting the criteria to be in the national Sample Study. Staff have received special training for administering and scoring the extra data collection instruments. Finally, the program staff consistently have evaluated their own program over the years in order to improve its operations, outcomes, and impact in the county.

Program information at the end of 1995-96. As part of its recordkeeping and reporting functions, the Monongalia County Even Start Program summarizes and reports on the program at the end of each project year (June). Staff prepare a report of their data, with numerous tables and charts, for their own and others' use. Staff's report of their program at the conclusion of the 1995-1996 project year appears in full as Appendix A. The following paragraphs are drawn from the data supplied in staff's report (Even Start Staff, 1996) and serve to describe the program, its families, and their characteristics at the time this evaluation began in the fall of 1996.

In June of 1996 there was a total of 62 families in the Monongalia County Even Start program that participated in at least two of the program components (parenting education and early childhood education). Of those families, 54 (87%) participated in all three program components (adult education, parenting education, and early childhood education). From those 62 families, 99 adults (62 mothers and 37 fathers) participated in at least one program component. The composition of families in the program are 34 (55%) couple households and 28 (45%) single households. Collectively, these families have 139 children participating in the program. Of those, 122 (88%) are children between 0-7 years old and 17 (12%) are between 8-18 years old. The average number of children for all families was 2.2, but there was a difference across family type in that the average number of children in the couple households was 2.5, while the average in the single households was 1.9 children.

The distribution of the ages of the 122 children between 0 and 7 years of age was rather even at 8% to 13% for ages 1, 2, 4, 5, and 6. At 20%, there were more children between 0-1, followed closely by 15% in the 3-year-old category. With 5%, the 7-year-old category had the fewest children. The number of 4-year-olds participating in the Even Start program increased from 10 in the 1993-1994 year to 17 anticipated for the 1996-1997 project year. The percentage of those 4-year-olds enrolled in Head Start is viewed as one measure of this program's success. Staff figures show a steady increase in the percentage of 4 year olds enrolled in Head Start from 30% in 1993-1994, to 57% in 1994-1995, to 67% in 1995-1996, and 83% projected in 1996-1997.

The average yearly income for all 62 families in the Monongalia County Program in 1995-1996 was $7,583. However, this varied a lot by family type. For couple households in the program, the average family income was $10,383; but for single households, the average family income was $4,040. The family incomes ranged from a low of $2,206 up to a high of $28,000. For all enrolled families, 53% of their incomes were under $5,000. The sources of income varied with the majority (61%) being AFDC and 32% from employment. The three other sources of income were disabled/retired (3%), parental support (2%), and child support (2%).

Participating parents' schooling completed at enrollment and participation in adult education was reported by program staff. With respect to schooling completed upon enrollment, 51% of the fathers and 33% of the mothers had a GED or high school diploma prior to enrolling in Even Start. Staff reports that 5 (12%) of the remaining mothers received their GED during the 1995-1996 program year. For mothers, 51 (82%) participated in some form of adult education with the most (27 or 44%) in adult basic education with some Even Start direct activity and less (18 or 29%) only in Even Start home-based adult education. For fathers, 7 (19%) participated in some form of adult education with the most (5 or 71%) in adult basic education with some Even Start direct activity. One reason why fathers did not participate in adult basic education as much as mothers was because 17 (57%) were high school graduates, had a GED before enrolling in Even Start, or had some college experience.

Objectives for This Evaluation

The Monongalia County Even Start Program staff contracted with the Appalachia Educational Laboratory (AEL) to conduct an independent, third-party evaluation of their Even Start program. Since the program had completed its fourth year and since two prior process-oriented evaluations had been completed, the new evaluation was to be more product-oriented and involve the analysis of much quantitative-type data already collected by program staff as a function of being involved in the national evaluation Universe Study and, also, the national evaluation Sample Study. Too, given that the national evaluations seemed to report mainly on nonrural, center-based Even Start programs, a part of this evaluation was to determine, depict, and describe the unique aspects of Monongalia County's mostly rural and home-based Even Start program.

Given the above, the main purposes of this evaluation were (1) to describe and evaluate the unique components of the rural, home-based Monongalia County Even Start program; and (2) to evaluate the impact of the program on the Sample Study families in terms of their educational and developmental outcomes. These were the two broad goals that guided this evaluation.

Four objectives were designed to address the two main evaluation goals of describing the unique, rural and home-based program and to evaluate its impact on participating adults and children. Specifically, the four objectives of this AEL evaluation were:

1. to identify (1) the essential components of the Monongalia County Even Start Program, and (2) the implementation variations of each component;

2. to evaluate the implementation of each of Monongalia County Even Start Program's essential components;

3. to assess the adult and parenting education outcomes of the program's Sample Study adults; and

4. to assess the early childhood education outcomes of the program's Sample Study children.

Audience for This Report

The primary audience for this evaluation report is the Monongalia County Even Start Program staff and their host agency, the Monongalia County School District. This evaluation is seen as a product- and outcomes-oriented study of the Even Start program after four continuous years of operation. As such, this evaluation is expected to be useful to program staff to (1) explain their rural, home-based Even Start model to clients, funders, peers, and others; and (2) to document and describe the impact of their program on the adults and children in the Sample Study. Secondary audiences would include state administrators of the Even Start program, administrators and researchers interested in family literacy models, educators and other scholars interested in rural education and rural programs, and evaluators interested in evaluating family literacy programs.

The next section presents and explains the methodologies and techniques employed in this evaluation. This is followed by two sections of findings from those various evaluation techniques, including tables and charts. One section of findings is of the components in the local program and the other is of the participants' educational and developmental outcomes. Next, important conclusions drawn from those findings are presented in a separate section, followed by recommendations in the last section. Relevant exhibits and other information are provided in the appendices. An executive summary of this evaluation effort preceded this section.

METHODOLOGY

This section of the report focuses on the methodology and techniques employed by AEL staff to evaluate the Monongalia County Even Start Program. This section first presents the national evaluation with its Universe Study and Sample Study, followed by a description of the Monongalia County subsample of the national sample. This is followed by longer explanations of the various data collection techniques, including the two adult instruments (with three outcome measures) and the two instruments (and four outcome measures) for children. The last part of this section explains the data analysis processes, including the process used to develop the Innovation Configuration Matrix from the interviews and observations and the multiple methods for analyzing each of the seven outcome measures.

National Evaluation Population Characteristics

As described in the first section, the second national evaluation of Even Start covers the 1993-1997 program years. Fu Associates was awarded the primary contract for this national evaluation with Abt Associates as the subcontractor for the Sample Study. The four components of the national evaluation are:

(1) Even Start Information System (ESIS) to collect data from all projects;
(2) the Sample Study to collect data from 57 selected projects;
(3) site visits to Migrant Education Even Start projects; and
(4) site visits to Even Start projects for Indian tribes and tribal organizations (Tao, Swartz, St. Pierre, & Tarr, 1997, p. 12).

The first two of these components are relevant to this local Even Start evaluation because they describe the population of all such projects in the nation and, also, the 57 randomly-selected projects in the Sample Study.

The authors of the 1995 interim report of the national evaluation (published in January 1997) estimated that the 513 Even Start projects served about 31,000 families nationally; however, data were available on 28,534 families. They report that the number of those families with two parents has remained rather steady at 50% (Tao et al., 1997, p. 19). The remaining families are divided between single parent families (37%-39%) and extended families (12%-13%). The average number of persons per household in 1994-1995 was 5.5, including parents and children.

Regarding their ages, Even Start parents typically are in the 22-29 or 30-39 categories. Their ethnicity has changed dramatically over the years. Tao et al. (1997, p. 19) report that the percentage of Hispanic parents has risen from 22% in 1989-1990 to 36% in 1994-1995, the largest ethnic group. In that same period, the percentage of Caucasian parents decreased 6% and African-Americans decreased 13%. As expected, the number of parents who spoke Spanish also increased in those same years from 15% to 29%.

The levels and sources of income have been rather low and consistent over the years. For example, in 1994-1995, 72% of the Even Start families had total annual incomes of less than $11,999. Since the 1992-1993 program year, just under 50% of the participating parents reported that either government assistance or wages from a job was their primary source of income. But, about three-fourths of the parents reported that they did not have a job at the time of their intake to the Even Start program (Tao et al., 1997, p. 20). Also at intake, the educational level of the parents has declined from 1992-1993, when about 21% had high school diplomas or the equivalent to about 11% in 1994-1995 who had either.

The average age of children in Even Start program has changed over the years. In 1989-1990, the average age of children was 4.3, but that figure declined to 3.7 in the 1992-1993 program year. Then, in 1994-1995, it rose to 4.4 years. This might have been a result of ages deemed eligible to participate in Even Start activities. The percent of Even Start children who had not participated in formal educational programs prior to their enrollment in Even Start has been relatively stable at 55% in 1992-1993 and 57% in 1994-1995 (Tao et al., 1997, p. 20).

Sample Study and Monongalia County Subsample

The Sample Study being completed by Abt Associates consists of 57 randomly-selected projects from the national evaluation's Universe Study of 490 state-administered Even Start projects. The main purpose of the Sample Study was to collect additional outcomes data from families not being collected from the Universe Study families (Tao et al., 1997, p. 14). More detailed participation and outcomes data were sought from the parents and children in the 57 Sample Study than from the parents and children in the Universe Study. This reduced the participation and outcomes data collection burdens on the projects.

The 57 Even Start projects in the Sample Study were chosen to be representative of both urban and rural areas and, also, all regions of the United States. Another criterion in the selection process was the projects in the sample were not new Even Start projects—they were projects that started in either the 1992 or 1993 program years. Thus, they were in their third or fourth year of operation by the 1994-1995 program year. The Sample Study was designed to collect participation and outcomes data over two years from each family. There were two different groups or cohorts of families in the sample, those that started in the fall of 1994 and those that started in the fall of 1995.

The systematic collection of assessment data from parents and children was the primary feature of the Sample Study. For parents' attainment of the adult education goals, participating projects could choose one of two nationally-normed instruments: the Comprehensive Adult Student Assessment System or the Tests of Adult Basic Education. The local project selected the former instrument. Assessing the parenting education outcomes was accomplished with the Home Screening Questionnaire. The outcomes selected to be assessed for the Sample Study children were school readiness and language development. These outcomes were measured with the PreSchool Inventory

and the Preschool Language Scale-3, respectively (Tao et al., 1997, p. 15). These instruments were to be administered to all parents and children in each family, with the notable exception of the Home Screening Questionnaire, which was targeted to a single child in the family. The 1995 interim report shows that the number of parents with posttest scores dropped to nearly half of those with pretest scores (which was about 450). About the same number of children completed pretests, but approximately 350 had posttest scores (1997, p. 18).

The Monongalia County Even Start Program was one of the 57 local project sites to be selected at random to be in the national Sample Study. Monongalia County is one of the rural sites in the sample. Also, this site is one of two sites representing the state of West Virginia. As part of the Sample Study, the local project staff were trained and did administer the additional outcome assessment measures to the parents and children in the participating families.

Participating families in the Monongalia County Even Start Program met the eligibility requirements established by Even Start. First, adults in the family unit were defined as being 16 years old or beyond the age of compulsory school attendance. Parents were adults who lacked the basic literacy skills or basic job skills for the current labor market. Children were defined as being under eight years old, although older children could receive Even Start services but they were not tested as part of the Sample Study. Per the law, the local project targeted defined attendance areas and, for Monongalia County, it was the Riverside and Dorsey Head Start areas. Combined, these two attendance areas included five schools: Riverside and Cass for the former and Central, Cool Springs, and Summers for the latter (Even Start Staff, 1995, p. III-19).

Even though the Monongalia County program administered all of the required outcome assessment measures to all 62 of their participating families, families in the national sample had to meet additional criteria. This dramatically reduced the Monongalia County subsample to 16 families out of the 62 in the program. The additional criteria employed to create the subsample for this evaluation included only families that were new to the Even Start program and, also, only families in which at least one adult and one child spoke either English or Spanish. This was because the outcome measures were available in just those two languages. Another criterion for being in the Sample Study and, thus, the Monongalia County subsample, was that families had to have started in the Even Start program in the three-month window of September, October, and November of the program year. Further, they had to complete their intake and pretest measures within 30 days of their enrollment. Additionally, families in the local subsample had to be in the program for at least three months before their posttests were administered.

There were 16 families in the Monongalia County subsample. Eight of those families enrolled in September-November of the 1994-1995 program year and the other eight families enrolled in September-November of the 1995-1996 program year. In many ways, these 16 subsample families shared many of the characteristics as the other 46 families in the Monongalia County program. This information was provided earlier in the report and will be supplemented with just a few characteristics of the subsample. For example, the average number of children in the subsample families was 2.25. Although both parents in the family were administered the outcome measures in the couple-type

households, only the mothers in the subsample received two or more of the Even Start components on a regular basis. ***Therefore, the adult outcome measures in the subscample were from all females.*** Only one child in each family was targeted to be the subject of the Home Screening Questionnaire measure and, so, only that child was included in this evaluation study. Of those 16 children in the subsample, 3 were in the 0-3 age category and 13 were in the 3-6 age category. The gender of the children selected to be in this evaluation, on the basis of being selected to be the target for the one instrument, was 12 males and 4 females.

The information provided in the previous paragraph was drawn from the 16 families in the Monongalia County subsample at the time of their enrollment in the Even Start program. As experienced at the national level, there was attrition in the sample as families left the program for a variety of reasons. Four families dropped out of the Monongalia County subsample sometime during the three years included in this evaluation. Thus, the number of adults and children with data varied for almost every outcome measure. Too, in a few cases, a parent may not have completed an assessment measure or a child may not have verbalized sufficiently to yield a scorable response.

Data Collection Techniques

Data for this evaluation of the Monongalia County Even Start Program were collected by a variety of ways including program records, face-to-face interviews, observations of events, and various outcome measures. This evaluation benefitted from the local project being in the national Sample Study in that all of the adult and parent outcome measures were administered, scored, and recorded, and were available for analysis. Each of the major data collection procedures is described in the following subsections, starting with the qualitative-type then moving to the quantitative-type.

Program Records

Several different program records were utilized in this evaluation; some indigenous to the local project site and some part of the national evaluation. Since the home visit was the primary service delivery mechanism employed by the Monongalia County project, the Even Start staff designed and implemented an Even Start Home Visit Report Form (see Appendix B). This Home Visit Report Form provided room for family name, date, number of visits, members present, and theme of visit at the top of the form. Next, room for parents' comments/concerns, handouts of items loaned or given to the family, and materials used in the lesson was provided. About one third of the form was devoted to the early childhood part of the visit, including developmental focus, benchmarks or performance criteria, and materials and activities. The remaining section of the form was for the adult goals portion of the home visit that also started with the main focus to be checked and adult benchmarks or performance criteria. Then, the next two areas were for writing in follow-up notes to the Family Educator and the parent. The last part was labeled "Comments," which typically included ideas for follow-up.

The completed Even Start Home Visit Report Forms were filed in individual family folders. These folders, with all the accumulated forms and information about the family, were made available to the AEL evaluator after guaranteeing the confidentiality of the information. Included in these family folders maintained by the assigned Family Educator was all the pertinent information about the subject family members from enrollment to the time they exited or completed their participation in the Even Start program. This included recruitment, enrollment, health records, vision and/or hearing screening reports, referrals to other agencies, and any other relevant reports. Also included was a form with all the outcome assessment administration dates and scores. Much of the important information in these folders also was available in the computerized database developed for the national evaluation, but these family folders did provide unique information to the instant evaluation not located elsewhere, such as number of books and other learning materials circulated to the family during the program year. This information was collected as part of a week-long site visit to the program on December 2-6, 1996 (see Appendix C for the site visit schedule).

The major program records used in this evaluation were the Form E reports of the Even Start Information System. All projects in the national evaluation completed Forms A, B, and C as part of the Universe Study, but only projects in the Sample Study completed Form E for their families. Form E recorded all the service delivery and educational/developmental outcomes information for the adults and children in the sample. There were two parts to Form E: the Adult form and the Child form. In the Adult version, staff input basic identifying information; the number of contact hours of adult education and parenting education; and the assessment measure administration dates and scores for reading literacy, mathematics literacy, and parenting education. The Monongalia County Even Start staff, following advice given to them, were very careful in entering the contact hours information to separate the adult education hours from the parenting education hours to avoid double counts. The Form E software program had provisions for three administrations of each outcome measure: pretest, posttest, and follow-up. The Child version of Form E was similar to the Adult version except, of course, that it solicited service delivery and outcome measures information germane to the participating child. The child service delivery information was input as the early childhood education days scheduled and days attended, by program year months. The child outcome measures were the school readiness and language development instruments. Like the Adult form, the Child form had provisions for the pretest, posttest, and follow-up administrations of each instrument.

Two other program records, developed by the Even Start staff, were utilized in this evaluation. The first was the set of completed self-assessment rating sheets that staff completed in monthly meetings over the 1996-1997 program year. Using the quality program indicators published in the *Guide to Quality for Even Start Family Literacy Program* (Dwyer, 1995), staff discussed each indicator in terms of where their program was and rated themselves on the 1 to 5 Likert-type scale, with 1 being "very descriptive." The indicators were in ten different Even Start program areas such as integration of components and home visiting. The number of individual indicators in each area discussed and rated by the staff ranged from a low of 9 to a high of 19. The second program record was the compilation of all staff development sessions attended by the Even Start staff from the beginning of the program to the end of the 1996-1997 program year. Program staff compiled this information into one listing that showed the names of the staff development sessions, their sites, the

names of program staff who attended, and the number of hours for the session. For the first of these two records, simple means and standard deviations were computed; for the second, the extended hours by staff attending figures were calculated.

Interviews

Three different structured interview forms were developed and utilized in this evaluation. These forms were for Even Start program staff, adult education collaborators, and Head Start collaborators. The program staff interview form was vastly different than the other two forms because it was designed specifically to solicit responses that would be used to develop the draft Innovation Configuration Matrix (ICM). The AEL evaluator selected sight components in the Monongalia County Even Start Program that served as potential components for the draft ICM. A series of three or four questions related to the potential ICM were placed in the interview form to assess the criticality of each possible component, for a total of 25 questions. The last question on the program staff interview form was open-ended, soliciting any further comments or thoughts from the interviewee. This interview form was printed on six white pages.

The adult education collaborator and Head Start collaborator interview forms also started with basic identification information (names, dates, location, and time). Then, three background information questions were asked (e.g., length of time in present job), followed by questions asking about the respondent's connection to the Even Start program. Next, 14 questions about their role and involvement in the Even Start program were asked, followed by an open-ended question. The total of 18 questions in both the adult education and Head Start collaborator interview forms fit on four pages. These two forms were printed on yellow and pink paper, respectively, to assist in identifying them during the December 2-6, 1996, site visit.

Observations

Direct observations of selected Even Start activities were completed by the AEL evaluator to help assess those processes. Specifically, the evaluator observed the annual fall collaborators breakfast meeting and work session and one of the program staff's self-assessment sessions with the quality program indicators. Both of these observations were completed during the week-long site visit in December 1996. The self-assessment observation was completed on Tuesday morning and the collaborators breakfast was completed on Friday morning. The purpose of these observations was to witness both activities firsthand in order to understand them better and, also, to relate them to other components in the Even Start project. Following recommendations for qualitative data collection, the identity of the evaluator was made known to participants and questions were encouraged. At the same time, participants were informed that their identities would not be revealed in any evaluation reports.

Adult Instruments

To study this outcome of their adult education component on participants, the Monongalia County program selected to administer the Comprehensive Adult Student Assessment System (CASAS). To study the outcomes of their parenting education component, this program administered the Home Screening Questionnaire. Each of these adult instruments is discussed in the following subsections.

CASAS Reading and Mathematics Literacy. The Comprehensive Adult Student Assessment System is an adult-oriented, functionally-relevant assessment system that measures a broad range of adult literacy skills and their application in real life domains. All of the CASAS tests are paper-and-pencil instruments. Appraisal tests are initial assessment instruments in the CASAS system. These appraisal tests help place the adult into the correct level of the Life Skills tests. These Life Skills tests assess the adult's basic skill levels in reading and mathematics literacy (CASAS, 1993). The CASAS has been used for over 15 years and in 27 states.

Results from the various CASAS assessment instruments are reported on a common scale. This common, numerical scale with its corresponding competency indicators has become a standard means for reporting learning outcomes at local, state, and national levels (CASAS, 1997). CASAS scores range from 150 to 260 points. This numerical scale has been divided into five levels, with descriptions for each. The five CASAS levels are: A, Pre-Literacy (150-200); B, Basic Skills (201-210 = Beginning Basic Skills, and 211-220 = Intermediate Basic Skills); C, Advanced Basic Skills (221-235); D, Adult Secondary (236-245); and E, Advanced Adult Secondary (246-260). The score of 225 is the national high school level. The Appraisal tests are used as level test locators and are not reported as scores.

The CASAS common numerical scores apply to all the various tests, including both the reading and mathematics Life Skills tests used in this evaluation. Both the Reading Literacy and Mathematics Literacy tests consist of 20 items and both take about 20 minutes to complete, but neither is administered in a timed manner.

The *CASAS Technical Manual* (1993) contains one chapter (#4) on the reliability of the instruments and another chapter (#5) on their validity. However, Tao et al. (1997, p. 152) opine that "There is only minimal data on the psychometric characteristics of the CASAS." The technical manual discusses both traditional and item-response type reliabilities in chapter 4. The traditional Kuder-Richardson 20 coefficients are all above .80 for the 70 series CASAS tests used in this evaluation. These K-R 20 reliability coefficients ranged from .81 up to .89 for the six tests in the 70 series, with half of those at .88 or .89 (CASAS, 1997, p. 26). The CASAS manual then presents unique standard errors of measurement for each ability level in scale scores for the reading literacy tests. Also, Tao et al. (1997, p. 152) report a test-retest stability reliability coefficient of .86 for adults in the first Even Start national evaluation. Regarding validity of the CASAS tests, the technical manual's chapter 5 discusses this, at length, relative to the Item Response Theory, used in the construction, calibration, and banking of the items. Other forms of validity for the CASAS tests are

presented in terms of adults' CASAS scores in relation to the placement of students into GED instruction and, also, to the student performance levels developed in the National Mainstream English Training Project. The *CASAS Technical Manual* authors conclude that their system possesses sufficient external validity.

Home Screening Questionnaire. The Home Screening Questionnaire (HSQ) was developed by researchers at the University of Colorado Medical School for use by educators and health professionals who are interested in promoting, studying, and evaluating child development. The HSQ is a screening instrument of factors with a young child's home environment that have been found to be related to the child's growth and development (Coons, Gay, Fandal, Ker, & Frankenburg, 1981). Items on the HSQ were selected from the more lengthy home environment assessment instrument titled Home Observation for Measurement of the Environment (HOME) developed by Caldwell and Bradley (1978). While instruments developed by Caldwell and Bradley require a home visit by a trained observer to complete, the HSQ does not require a home visit and it is administered to the parents of the subject child. The primary purpose of the HSQ is to screen the environments of homes for children between birth and six years of age.

The HSQ is a questionnaire for parents written at the third or fourth grade reading level (Coons et al., 1981). The instrument consists of yes/no, multiple choice, and constructed-response items. There are two versions of the HSQ: one for children 0-3 years old and a second for children 3-6 years old. The HSQ for children 0-3 consists of 30 items to be completed by the parent and a list of 50 different toys and other home materials (pots and pans, crayons, etc.) to be checked. The HSQ for children 3-6 consists of 34 items to be answered and the same list of 50 toys and other home materials. Both versions of the questionnaire are printed on both sides of one piece of 11" x 17" paper, folded in the middle, creating four 8½" x 11" pages of questions/lists. The 0-3 HSQ is printed on blue paper and the 3-6 HSQ on white paper, allowing easy identification. Each version of the HSQ takes about 15-20 minutes to complete. Scoring completed HSQs is done by following detailed instructions in the HSQ reference manual. The maximum possible score for the 0-3 version is 43 points and for the 3-6 version the maximum is 56 points; there are no subscales for either version of the HSQ.

Reliability and validity information about the HSQ is presented in the reference manual (Coons et al., 1981). Internal consistency reliabilities of the HSQ were computed with the Kuder-Richardson Formula 20 procedure. The K-R 20 internal consistency coefficients were .74 for the 0-3 HSQ and .80 for the 3-6 HSQ. Stability (test-retest) reliabilities were computed for samples of parents in each instrument's age group (0-3 and 3-6). These stability reliability coefficients were .62 for the 0-3 HSQ and .86 for the 3-6 HSQ. But, the authors report that the stability reliability coefficient for the younger children increased to .82 when the scores from the parents of children aged 0-1 years old were dropped from the sample. To assess the validity of the HSQ, the Colorado researchers completed several analyses with the HSQ and the HOME instrument scores. The agreement of low HOME and HSQ scores and high HOME and HSQ scores were computed for each version of the HSQ. The developers found 81% agreement in the low scores of both instruments for the 0-3 HSQ and 86% agreement for the 3-6 HSQ. Regarding high scores, they found 66% and 55% agreement for the two HSQ versions, in order. The authors conclude that the HSQ "identifies a large proportion

of children whose HOME scores would be low" (1997, p. 6). Content validity was assumed because HSQ items were drawn from the HOME (Caldwell & Bradley, 1978) original items.

Child Instruments

The outcomes identified for the early childhood education component of the Even Start program were school readiness and language development. The PreSchool Inventory was selected to assess school readiness skills of children and the Preschool Language Scale-3 was selected to assess their language development. Each instrument is discussed below.

PreSchool Inventory. The PreSchool Inventory (PSI) originally was developed by Betty Caldwell as a measure of the skills "regarded as necessary for success in school" (CTB/McGraw Hill, 1970). The first version of the PSI consisted of 85 items, but later it was revised to 64 items. That version was shortened again to 32 items for use in the Head Start Planned Variation study. The authors of that large-scale evaluation reported that their 32-item version possessed adequate reliability and validity and was less burdensome to administer by programs and requires less time of parents (Walker, Bank, & Bryk, 1973). Later, Abt Associates, with permission from the Educational Testing Service, created the present version by combining the English and Spanish versions into one form.

The PSI is administered individually to the child by asking questions, some involving materials such as checkers and crayons. The PSI takes about 15 minutes to administer. Example items include: "Show me your shoulder," "What does a dentist do?," and "Color the square purple." The measure is appropriate for children up to five years old. Each item counts as a point and only a total score is computed; there are no subscales. The PSI is printed on the front and back of one piece of buff, 8½" x 11" paper. This one-page PSI includes both the English and Spanish version of each item, in addition to typical identifying information at the top of the first side (Abt Associates, 1994). This version of the PSI has been used in numerous large-scale evaluations.

Descriptive statistics, reliability, and validity information regarding the PSI is presented in its administration manual (Abt Associates, 1994). This manual displays the means and standard deviations by age groups for the PSI total scores from four large-scale evaluations. They report that the mean scores are very consistent across these different samples of children and the standard deviations are relatively large and homogenous across age groups and samples. "In addition, there are no apparent floor or ceiling effects" (p. 25). Internal consistency reliability coefficients (Alpha) were reported for ten administrations of the PSI in six large evaluations. These coefficients ranged from .77 to .87, with just one in the .70s. Stability (test-retest) reliability coefficients were reported for four major evaluations, with the interval between administrations ranging from two weeks to seven months. These coefficients ranged from .66 to .95, with just two in the .60s. To assess concurrent validity, the PSI was correlated with four other measures of cognitive abilities with the results of .48 to .74 for the seven reported correlations. The PSI was factor analyzed as part of construct validity and those results confirmed that it did contain a unitary factor (no subscales) (pp. 25-32).

Preschool Language Scale-3. The precursor to the present version of this measure was the Preschool Language Scale developed in 1969. This first version was based on the best available information on language development in young children from the fields of human development, psycholinguistics, and speech-language pathology (Zimmerman, Steiner, & Pond, 1992, p. 2). The first version was revised in 1979 to update its content, provide clearer administration directions, and provide a simplified scoring system. Next, it was revised a second time in 1992 and retitled Preschool Language Scale-3 (PLS-3) by its developers.

The PLS-3 measures both receptive and expressive language skills of young children and yields scores for each plus a combined total score. The receptive language skills are measured by the Auditory Comprehension subscale of the PLS-3, which assesses children's ability to process and understand the language that they hear, including skills in the areas of listening; the meaning of words, content, and syntax; and integrating thinking skills (Zimmerman et al., 1992, pp. 6-10). Expressive language skills are measured by the Expressive Communications subscale of the PLS-3, which assess children's ability to produce language, including skills in vocal development, social communication, use of words and concepts, syntax, and integrative thinking skills (pp. 10-15). The PLS-3 Total Language score is computed by adding the standard score of each subscale and converting that sum to the same standard scale as the individual scales.

The PLS-3 is appropriate for children of ages 2 weeks to 6 years and 11 months, and is administered by a trained examiner. The amount of time required to complete the instrument varies by the age of the child, but about 30-40 minutes is needed, typically. The PLS-3 examiner's kit includes an easel-backed picture manual, a spiral-bound *Examiner's Manual* with administration and scoring directions, and the 16-page Record Form. To administer the PLS-3, the picture manual; the *Examiner's Manual*; the Record Form; and a variety of materials such as a teddy bear, a tennis ball, keys on a ring, three plastic cups and spoons, and eight play blocks are required. Also, age-appropriate books and toys are recommended as icebreakers for the individually-administered PLS-3.

Reliability and validity information is presented in Chapter 4 of the *Examiner's Manual* (Zimmerman et al., 1992). The internal consistency reliabilities (Alpha) for a sample of 1,200 children in 14 age groups ranged from .47 to .94 across the two subscale and total instrument scores. Of the 52 Alphas for the three scores, 8 were in the .90s, 18 were in the .80s, 11 were in the .70s and the remainder were below .70. All of the lowest Alphas were reported for age groups of less than one year old. The stability (test-retest) reliability coefficients were assessed with a subsample of 1,200 children in three age groups of 3 to 3½ years, 4 to 4½ years, and 5 to 5½ years. With the test-retest interval ranging from two days to two weeks, the nine stability coefficients (3 subscales x 3 age groups) were all .82 or higher with six being in the .90s (pp. 88-90). Content validity was determined by reviewing the language tasks with the items in the PLS-3, with the developers concluding that the instrument provides a thorough and balanced sample of language behaviors. Concurrent validity was assessed by correlating PLS-3 scores with scores of samples of children on the previous version of the instrument and on the Clinical Evaluation of Language Fundamental-Revised instrument. The resultant correlation for the six (3 subscales x 2 instruments) ranged from .66 to .88, with the two highest being with expressive communications and total instrument scores of the previous version.

Construct validity was evaluated by assessing, via discriminant analysis, how the PLS-3 scores correctly identified children as having language disorders or not. Small samples of children in three age groups were involved in this study, with the results being that 66% of the 3-year olds, 80% of the 4-year olds, and 70% of the 5-year olds were identified correctly as language-disordered or not (pp. 91-93).

Data Analysis Processes

All data analyses for this evaluation were completed by the AEL evaluator at the AEL facilities. Data analyses were conducted on both the qualitative and quantitative types of data. The qualitative data from interviews and observations were analyzed by hand while the quantitative data were analyzed by a statistical software program. Each of the three major analysis processes is explained below.

Innovation Configuration Matrix

The first major process in data analysis was to construct the draft Innovation Configuration Matrix (ICM) from program records, interviews, and observations. Using program records as input sources, the evaluator prepared the initial list of essential components in the Monongalia County Even Start Program. This list of nine essential components was discussed with the Even Start coordinator who suggested that one (retention) was not a problem in their rural program, because attrition is low in their home-based program. However, it is an important component in many other Even Start programs in the nation. The remaining eight components of the local program were retained temporarily and used as the chief organizing scheme for the site visit observations and face-to-face interviews with program staff. Three or four questions regarding the implementation of each component were included in the interviews for the four program staff. The adult educator and interagency collaborator interview forms differed in that they asked more specific questions about the components with which the interviewee worked.

Following the site visit in December 1996 (see Appendix B), the completed interviews and the observation notes were analyzed (1) to confirm the essential components in the ICM and (2) to identify the major variations in the implementation of those components as described by program staff, the adult educators, and the interagency collaborators. This analysis led to development of the draft ICM, which consisted of eight components with four variations for each. Thus, the draft ICM was 24 cells of information describing the essential components and the major variations in the implementation of each within the Monongalia County Even Start Program. Also, three different graphic depictions of the essential components in the Monongalia County Even Start Program were prepared in draft form for staff to review and revise in preparation for a final selection for displaying the program's model to others. Both the draft ICM and the associated draft graphic depictions were mailed to program staff to review near the end of January 1997.

During the next two months, program staff and the evaluator interacted about the ICM and graphic depiction of the program. Program staff critically reviewed the eight essential components and the detailed descriptions of their implementation variations. AEL staff made those changes that were easy to understand and initiated telephone calls to discuss those suggestions that required more elaborations and examples based on the staff's daily work with the project. Also, staff made slight adjustments to one of the three graphic models for AEL to prepare a camera-ready master for use in the evaluation report and by project staff. By early April, the ICM and the graphic depiction of the eight components were finalized by the two parties. Next, the final copy of each was given to the graphic designer at AEL, who completed camera-ready copies of each item. These final, print-quality masters were sent to the local project staff on April 15, 1997, for immediate use and also were retained by the evaluator for use in this report.

Treatment Hours or Days

Service intensity data for each family in the Monongalia County subsample were taken from the ESIS, Forms D and E output reports. The Form D output reported the participation in services by the family (number of home visits), adult education (total contact hours), parenting education (total contact hours), and early childhood education (number of months participated). Form E output reports were more specific; for example, the adult education and parenting education contact hours were reported for each month of the year. This information was important in this evaluation to establish the actual treatment period because not all families started receiving services in the same month. Similarly, while the Form D output reported the total number of months each child participated in early childhood education, the Form E output reported the days scheduled and the days attended for each month of the year. And, like the adult contact hours data, this information served to establish the treatment months for each child in the subsample. Evaluators were careful to record these hours for the one child in the family that was in the subsample. Local project staff printed out Form D and E reports for each family and each program year in this evaluation (1994-1995 to 1996-1997).

The first step in analyzing the treatment hours and days for each adult and child in the family was to summarize the data in the Form D and E reports. Working with each year's reports, the number of treatment months was determined from the monthly data and this was noted on the top page of the printout. Next, the number of hours or days for each treatment year month were listed on the top page and then summed. This step yielded four values to enter in the database: (1) number of adult education contact hours, (2) number of parenting education contact hours, (3) number of early childhood education days scheduled, and (4) number of early childhood education days attended. Then, the ratios of adult education and parenting education contact hours divided by the number of treatment months were computed and recorded. Similarly, the ratios for the two early childhood education day columns were computed and recorded. Also, the child's attendance rate (days attended divided by days scheduled) was computed and reported. Last, from the family's entry date already in the database, the two actual treatment years, of the three years' reports, were noted on the top page of each printout.

The second step toward analyzing the treatment hours and days was to transfer the new information on the paper copies of the Forms D and E output reports into the computer program. This was done in the Data Entry II module of the SPSS-PC+ statistical software package. Data entry was completed by the evaluator with constant double-checking to assure accuracy.

The third step was the actual analysis of the database. Here, the evaluator requested basic descriptive statistics on the (1) home visits, (2) adult education contact hours, (3) parenting education contact hours, and (4) early childhood education days attended. These descriptive statistics were obtained for (1) each of the three program years, (2) the three program years combined, and (3) the two treatment years. For each of the four main service delivery variables analyzed, tables were constructed from the statistical printouts. Each of these tables contained the number, minimum value, maximum value, mode, median, mean, standard deviation, standard error of the mean, and coefficient of variation. This latter value is the standard deviation divided by the mean (Arney, 1990, pp. 73-74).

Treatment Effects

Treatment effects data for every adult and child in the Monongalia County Even Start subsample were taken from the ESIS, Form E output reports. These reports contained both the administration dates and the obtained scores for each outcome measure. For adults, these outcome measure scores were provided for (1) the CASAS Reading Literacy, (2) the CASAS Mathematics Literacy, and (3) the Home Screening Questionnaire. For each child, the outcome measure scores were provided for (1) the PreSchool Inventory, (2) the PLS-3 Auditory Comprehension, (3) the PLS-3 Expressive Communication, and (4) the PLS-3 Total Language Score.

The first step in analyzing the treatment effects on the adult and child subjects in the sample was to transfer their output scores into the software program. There was provision for three scores for each subject on each outcome measure—pretest, posttest, and follow-up. However, not all subjects had all three scores on each outcome measure, as is normal for most projects of this type.

The second step was to compute the number of months in each treatment year for each outcome measure. This was necessary because the families enrolled and were pretested in different months and, also, because not all posttests were completed in the same month at the end of the program year. That is, posttests were started in April and all were completed by June. Too, some families exited in other months and, whenever possible, Family Educators attempted to administer posttests before they left. Two details about these treatment month computations are important. One, the pretest and posttest administration dates were used to compute the actual treatment months for the items/points per month calculations. This actual treatment time between testing administrations differed from the service delivery time described above, by not more than one or two months on average. Two, in computing these testing treatment months, partial months were rounded to the nearest whole month with the date of the 15th as the dividing point.

The third and fourth steps in the treatment effects analysis were completed in tandem for each outcome measure. In the third step, the computation of the t-tests for dependent means by the matched pairs technique was completed. Three separate t-tests were computed for each outcome measure: (1) pretest to posttest, (2) posttest to follow-up, and (3) pretest to follow-up over the two treatment years. The coefficient of variation values were computed by hand from the means and standard deviations in the computer printouts. For the fourth step, the effect sizes were calculated for each t-test by subtracting the second administration score from the first administration score and dividing the remainder by the standard deviation of the first administration (Cohen, 1977). Next, a table for each outcome measure was designed that displayed typical t-test results and the coefficient of variation and effect size values.

Fifth, the items/points per month were computed for each outcome measure. Using information in the database, the items or scale score points gained (or lost) were calculated for each of the same three testing periods as before. Next, these figures were divided by the number of actual treatment months between tests and that new value was saved as a variable in the database. Then, descriptive statistics on each outcome measure by each of the three testing time intervals were produced in the SPSS-PC+ program and the printouts were saved. Next, as above, a table for each outcome measure was constructed displaying the average number of treatment months, number of adults, mean items/points per month, the standard deviation, the coefficient of variation, and the standard error of the mean. Since the mean number of items/points gained (or lost) per month was a ratio less than one point, when it was divided by the standard deviation value, several of the resultant coefficient of variation values were larger than 1.00.

Sixth, and last, graphic displays of the outcome measure scores were prepared. Here, each adult's and child's score at pretest, posttest, and follow-up for each outcome measure was located on a chart, using the measure's possible points as the vertical axis. Then, the scores for each individual were connected with a unique symbol and line. These displays dramatically and effectively displayed the starting points, the growth (or decline) over time, the central tendencies in the scores, and the distribution of those scores. Also, for those scores with national norms, the position of the scores relative to the national norm was clear.

The findings from these analyses of the qualitative and quantitative data in this evaluation are presented in the next two sections—starting with the ICM and the graphic depiction of the program, then moving to the adult educational and parenting outcome findings and the child development outcome findings.

FINDINGS - PROGRAM COMPONENTS

This section presents the evaluation findings for the essential components of the Monongalia County Even Start Program. These findings are presented in the form of the completed Innovation Configuration Matrix, followed by a graphic depiction of those essential components. Next, evaluation findings for each of the eight components in the matrix and graphic depiction are presented. Most of these findings were derived from interviews and observations, but some were from program records.

Program's Innovation Configuration Matrix

The completed Innovation Configuration Matrix (ICM) for the Monongalia County Even Start Program appears as Appendix D. The eight essential components of the completed ICM are (1) identification and recruitment, (2) home visits, (3) early childhood education, (4) parenting education, (5) adult education/literacy, (6) collaboration, (7) evaluation, and (8) staff development. For each of these eight components, there are four possible variations of implementation. Thus, the Monongalia County Even Start Program ICM is a 24-cell matrix of eight essential components by four variations each. In the end, the retention component was judged not to be as essential to the program as the other components were, so it wasn't included in this final ICM. This is not to say that program staff are not concerned about participants staying in the program, only that these concerns are not on the same level or scale as the eight essential components are.

Two vertical lines divide the four variations for the eight components in the ICM (see Appendix D). A dotted line appears between variations A and B for each component and a solid line appears between variations C and D in the ICM. These two lines define the differences among the possible variations of implementation for each component. The implementation variations to the left of the dotted vertical line reflect ideal practices. These are the eight "A" variations in the Even Start ICM. These "A" variations of implementation represent the evaluation findings of the Monongalia County program and are described below, after the graphic depiction of the model. The implementation variations to the left of the solid vertical line reflect acceptable practices. These are all the "A," "B," and "C" variations of the eight components. And, the "D" variations represent unacceptable practices.

Monongalia County Even Start Program Model

Figure 1 is a graphic depiction of the Monongalia County Even Start Program. Here, the eight essential components are shown in their relationship to each other, which represents the model of the Monongalia County program. Some of the component names have been adjusted slightly from the ICM, which was completed prior to the graphic.

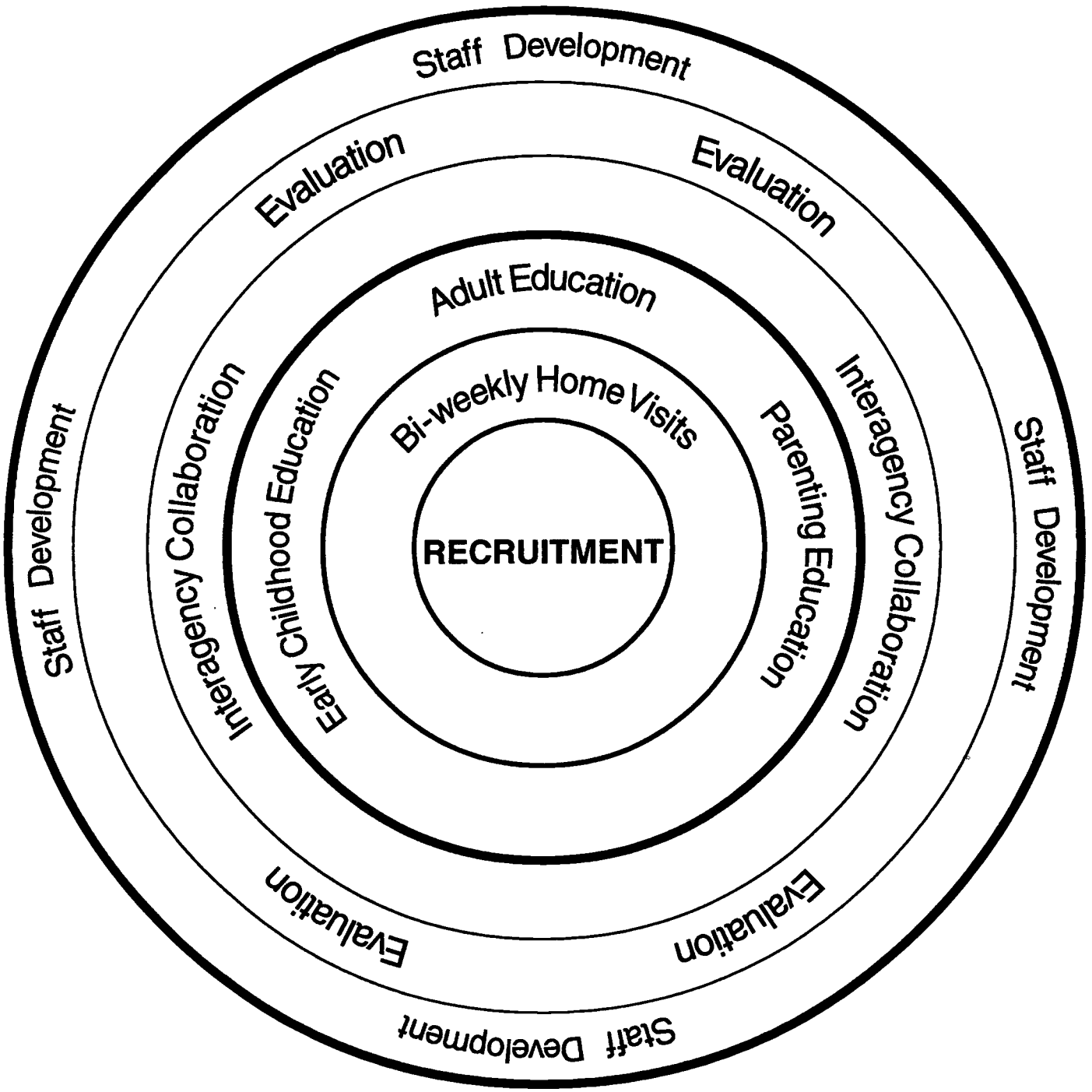


Figure 1

Graphic Depiction of the Monongalia County Even Start Program

The graphic depiction of the Monongalia County Even Start Program consists of six concentric circles containing the eight essential components. At the center of Figure 1 is the Recruitment component representing the identification and recruitment component in the ICM. The next circle in the model is the Bi-Weekly Home Visits, representing the home visit component in the ICM. The third circle in the graphic contains the Early Childhood Education, Adult Education, and Parenting Education components. The next circle surrounding the three core components is Interagency Collaboration, the fifth circle is for the Evaluation component, and the sixth circle is for the Staff Development component.

The second and third circles in the graphic depiction of the program are shaded red because they represent the heart of the Monongalia County Even Start model. Basically, the Monongalia County Even Start model is a rural, home-based program, as opposed to a suburban or urban, center-based program. The red shading of the second and third circles denotes the importance of the linkage of the bi-weekly home visits to the delivery of three core components of early childhood education, adult education, and parenting education.

Below are the evaluation findings for each of the eight components in Figure 1, starting with the component in the center.

Recruitment

The Monongalia County Even Start staff consistently employ a wide variety of identification and recruitment practices to bring eligible families into their program. First, they established clear eligibility requirements for families to receive Even Start services. These eligibility requirements are based on the requirements in the law and include residing in the specific school attendance area. These eligibility requirements are written and communicated to all involved parties, including collaborators with the Even Start program. These communications with collaborators take the form of one-on-one discussions, small group awareness sessions, and large group awareness sessions.

Identification and recruitment practices include soliciting and receiving referrals from Head Start and Title 1 teachers, regular classroom teachers, and collaborating agencies. Staff reported that referrals from Head Start and Title I teachers usually are more accurate than those from regular classroom teachers, because the former are much more familiar with requirements in the law for program participation and because they work with children of similar ages. Also, staff reported that referrals from other children and youth service agencies with whom they collaborate are often good leads for recruiting Even Start families. Contacts and referrals from school guidance counselors and from the adult education courses were reported.

Staff reported that individual identification and recruitment practices also are successful for locating eligible Even Start families. These individual efforts include drives into the neighborhoods or areas with young families in the attendance zone. Here, staff look for toys, swing sets, and other items denoting that the family includes one or more children between 0 and 7 years old. In these

cases, the Even Start staff will knock on the door of the household and conduct an on-the-spot interview to determine eligibility and interest levels. Closely allied with staff's visits to families in the attendance area is the identification of potential Even Start families by present participants. In these situations, an Even Start mother may suggest her sister's family or a neighbor. Staff then follow up on these nominations.

Last, staff held awareness and recruitment sessions with other neighborhood-based programs to locate eligible Even Start families. Mentioned specifically by staff were "Happy Schools" operated by rural church neighborhood houses. Another awareness session was held with the "Bright Beginnings" program conducted by the health department in the county. These awareness sessions did result in direct calls to the Even Start office.

Bi-Weekly Home Visits

Bi-weekly home visits are the heart of the Monongalia County Even Start Program. These planned, scheduled, and reported home visits are the major service delivery strategy in this program. Each enrolled family is assigned to one Family Educator for their services. The Family Educator usually schedules these home visits at the end of each prior visit.

After the home visit is scheduled, the Family Educator plans the actual activities to complete, often in conjunction with the parents' concerns and goals and the needs of the children and parents. This is done at the Family Educator's office. Typically, the home visit is scheduled to last one and a half hours. Although the specifics may vary by whom is planning the visit, all home visits strive to focus on literacy as the primary goal. Also, each visit is to include an activity in the areas of early childhood education and parenting education. If the parent is not enrolled in a center-based education course or class, then some adult education instruction usually is included in each home visit.

Actual bi-weekly home visits varied a little by the Family Educator and by the actual family, but most were planned to follow a regular routine or order. First, the Family Educator reviewed the last visit or asked the child to retell his or her favorite book from the five that were loaned for the two weeks. Second, the Family Educator either read a book to the child or had the parent do it. This was followed by an activity for the child alone, or with the parent. This activity usually was closely associated with the topic or story in the book. Third, the planned parenting education activity was completed, either from the Parent As Teacher curriculum or based on the just-completed activity with the child. Fourth, the adult education component activity was conducted, its extensiveness and subject determined by the amount of formal adult education courses or classes in which the parent was engaged. Fifth, the whole visit was summarized. Books and materials loaned previously were collected and the child made five new selections from a group of available books. While this was being done, the next visit was scheduled and the Family Educator discussed possible topics of interest with the parent.

When the home visit was completed, the Family Educator returned to the Even Start office, returned the borrowed books and materials to the central library, and completed the Even Start Home Visit Report Form (Appendix B). The completed Home Visit Report form was then placed in the family's folder. Any follow-up calls or referrals to other agencies or organizations were made next. With the date for the next home visit scheduled, the Family Educator then started thinking and planning for that visit.

Table 1 displays the descriptive statistics for the bi-weekly home visits to the sample family homes during the three program years. The three individual years' of data are presented in the first three rows, followed by the home visit data for the three years combined, then followed by the data for just the two identified treatment years. The minimum number of visits ranged from 6 (in 3 rows) to 20, while the maximum ranged from 21 to 53. The mode number of visits ranged from 6 to 42 and the medians ranged from 15.50 to 43.00 visits. The means were close to the median visits, ranging from 15.91 to 39.83. The mean number of home visits in the two treatment years was 37.82 with a standard deviation of 11.57. The four other standard deviations were from 4.16 to 13.10. The coefficients of variations were from .261 to .481 and the standard errors of the means ranged from 1.25 to 3.78.

Table 1

Descriptive Statistics for Visits to Sample Family Homes
During the Three Program Years

Home Visits Time Period	Number Families	Minimum Visits	Maximum Visits	Mode Visits	Median Visits	Mean Visits	Standard Deviation	Coeff. of Variation	Std. Error of Mean
1994-1995 Program Year	8	6	28	6	15.50	16.38	6.93	.423	2.45
1995-1996 Program Year	16	6	32	32	16.50	19.44	9.35	.481	2.34
1996-1997 Program Year	11	7	21	15	16.00	15.91	4.16	.261	1.25
Three Years Combined	12	6	53	42	43.00	39.83	13.10	.329	3.78
Two Treatment Years	11	20	53	35	35.00	37.82	11.57	.306	3.49

Asked how they evaluate these bi-weekly home visits and what makes for the best ones, the responses varied. "When families would recall an author or some characters [from the book], that was success," one Family Educator replied. Another replied that an indicator of success was when the child showed interest in books over toys in making their selections. Yet another Family Educator said successful home visits were noted by observable gains in literacy skills by either the child or parent. The "best home visits are ones in which the parent and the child are engaged in a good activity together," another stated. Making and using play dough was an example provided. Finally, heavy involvement in the home visit by the mother was a part of the best visits, according to another Even Start staff member.

Early Childhood Education

This evaluation determined that, in the area of early childhood education, the Monongalia County program staff regularly plan and complete interactive learning activities with parent and child that are based on validated curricula, address the developmental areas of the child, are based on the needs of the child, are interesting and enjoyable to the child, and are models for parents to follow. Also, as appropriate, staff encourage parents to enroll their children in early education classes.

The Parents As Teachers (PAT) program was chosen for this Even Start program, specifically because the curriculum materials were seen as an excellent match with the needs of the local families. PAT materials address the social, emotional, motor, language, and intellectual development areas of children ages 0 to 5 years. A locally developed literacy model is used with 6- and 7-year-old children. This locally-developed literacy model emphasizes children's literature in fostering cognitive and emotional growth. Both of these programs are based on the needs of the child, determined in prior home visits, observations, and interviews with the parent. Too, these early childhood education activities are designed to be both interesting and enjoyable to the child.

The early childhood education curriculum materials utilized in the home visits are designed to serve as models for parents to follow. These activities offer effective demonstrations of how parents can discipline, engage in constructive play, communicate better, etc. For many of the Even Start parents, these demonstrations provided by the Family Educator are new to them, as the parents did not experience them in their own upbringing.

In addition to early childhood experiences completed as part of the bi-weekly home visits, when the children reach enrollment age for early education classes, the Family Educator works with the parents to sign up the child to attend these classes. Examples of such early education classes include preschool, Head Start, and kindergarten. Eleven of the subsample families enrolled children in these classes. Data were kept on all children in the family, but one child in each family was selected for more intensive recordkeeping and testing. Program records reported the number of days scheduled and attended by these children in the two treatment years. Their attendance percentage varied from a low of 60% to a high of 100% (for two children).

Table 2 displays the descriptive statistics for the sample's child early education days attended during the three program years. The three individual years of data are presented in the first three rows, followed by the data for those three years combined, then followed by the data for just the two treatment years. The number of families to which the children belonged varied from 7 in the first program year up to 16 when the three program years were combined. The minimum number of days attended varied widely from 0 in the 1995-1996 program year to 71 in the two treatment years. The maximum days attended ranged from a low of 123 for the 1995-1996 program year to 230 days for the three years combined.

Table 2

Descriptive Statistics for Sample's Child Early Education
Class Days Attended During Three Program Years

Child ECE Days Time Period	Number Families	Minimum Days	Maximum Days	Mode Days	Median Days	Mean Days	Standard Deviation	Coeff. of Variation	Std. Error of Mean
1994-1995 Program Year	7	18	139	23	71.00	66.71	47.18	.707	17.83
1995-1996 Program Year	14	0	123	0	17.50	37.64	46.69	1.214	12.48
1996-1997 Program Year	11	2	135	18	20.00	66.09	59.82	.905	18.04
Three Years Combined	16	5	230	132	130.50	107.56	63.35	.589	15.84
Two Treatment Years	11	71	228	132	135.00	138.09	40.38	.292	12.18

Interestingly, the maximum days attended for the two treatment years was just two days less than the figure for the three years combined. The mode days attended ranged from 0 for the 1993-1996 program year to 132 for the two multiple-year rows in the table. The median days attended value ranged from 17.50 for the 1995-1996 program year to 135.00 for the two treatment years. The mean days attended ranged from 37.64 to 138.09 for the same years as the median range values. The standard deviations had a narrower range of values than their means at 46.69 (1995-1996 program year) to 63.35 (three years combined). But the coefficient of variation values ranged widely from .292 for the two treatment years to 1.214 for the 1995-1996 program year. The standard error of the mean had a narrower range of values from 12.18 for the two treatment years to 18.04 for the 1996-1997 program year.

Even Start program staff were asked what type of learning experiences for children they judged to be most effective and why so. Their responses varied from content to process orientations. One staff member replied that the most effective learning experiences are those that foster or involve creativity, especially when the product of the creative experience can be shared with and appreciated by the child's parent. Another staff member said the most effective learning experiences were those where the child was actively involved. A third staff member responded by naming the two developmental areas of language skills-type activities first, followed by fine-motor skills. Learning activities that the child has had the least experiences with in the past were the most effective ones, another Even Start staff member replied, followed by learning activities started and completed right in the child's home. Last, another staff member replied that the use of the laptop computers, taken into the homes by the Family Educator, by children three years and older is very successful, plus they have the added advantage of appealing to the parent—they like to complete computer-based activities also.

Adult Education

The Even Start staff regularly plan, meet, and discuss with parents a wide variety of possible adult education outcomes ranging from basic literacy skills, through the attainment of the GED, and up to and including entry-level job skills such as acquired through vocational/technical education classes. Based on these planned discussions with parents, individualized and realistic long-range and short-term educational goals are established. Included in these discussions with parents is the topic of what barriers might hinder the attainment of these educational goals. Examples of such common barriers include transportation problems to formal education classes or facilities and child care while the parent is in those classes. When such barriers to attaining educational goals are identified, the Family Educator seeks to help reduce, resolve, or eliminate them. Solutions might include finding transportation alternatives or suitable child care services.

This evaluation found that the Monongalia County Even Start staff made concerted efforts to help parents adjust to, and be successful in, adult education programs and classes. Staff utilize a variety of techniques to assist parents in adjusting to adult education courses or instruction such as preparing parents about what to expect in those classes and actually being present with the parent for the first several classes. Too, after the first classes are completed and the parent is on his/her own, the Family Educator monitors parent attendance in the planned sessions. If attendance by the parent is a problem, the staff attempt to discover why and solve it, if possible. The aforementioned transportation barrier often surfaces as a cause for low attendance. Also, the teachers at the adult education center are more sensitive and encouraging of Even Start parents, but the content in the lessons does not change for them. In fact, the adult education center staff developed a special room for Even Start parents to use for completing lessons.

Even Start staff reported that the adult education goals for members of their families varied widely from basic literacy skills attainment to specific job training skills for employment. In the main, the attainment of the GED was not a realistic adult education goal for many families, especially if

another member of the family possessed either it or a high school diploma. For most of the mothers in the Monongalia County project, more down-to-earth and realistic goals were the main interest. For example, the adult education goals for some mothers were closely related to obtaining part-time or full-time employment. For others, their goals were even more basic, such as being able to read books to their children. Other adults were interested in obtaining computer literacy skills for themselves and family members. Staff reported that the pivotal barrier to establishing and attaining adult education goals was whether or not the parent was willing to leave the home for instruction. If the parent were not willing to leave the home, then the goal was to provide that instruction in the home visits. This reduced the amount of time that could be devoted to adult education instruction, i.e., only in the home visit. At the other extreme, this writer talked to one Even Start mother at the annual collaborators breakfast who was nearing the completion of a 2,100-hour vocational education course, was going to use that training to obtain a job, and then planned to continue her education by being trained as a surgical assistant.

Table 3 displays the descriptive statistics for the sample's adult education contact hours during the three program years, individually first, combined second, then for just the two treatment years. The number of families ranged from 7 to 16, but the contact hours data were just for the mothers in these families. The number of adult education contact hours received by the mothers was extremely varied. This is shown initially by very large differences in the data in the third through seventh columns. For example, the minimum hours ranged from 1 to 18 while the maximum hours ranged from 455 to 1,542. The mode hours ranged from 2 to 18, and the median hours ranged from 12.00 to 70.00. These figures represent two very different groups of mothers who received adult education contact hours according to other data in the file. The larger group received adult instruction mainly during the home visits, while the second group received adult instruction mostly in more formal, center-based classes or courses. The mean number of adult education contact hours ranged from 79.60 in the 1996-1997 program year to 257.55 hours in the two treatment years. The extremely large variability in adult education contact hours is reflected in the standard deviation values, which ranged from 151.27 to 451.25—all of which were larger than their means. This caused all of the coefficient of variance values to be over 1.00. These values ranged from 1.261 for the 1994-1995 program year to 2.190 for the next program year. As expected, the standard errors of the mean were extremely large also, ranging from 47.84 for the 1996-1997 program year to 121.08 for the two treatment years.

The Even Start Family Educators assess their parents' educational needs at the intake stage and help parents choose and determine short-term and long-range goals. Then, during the course of the program year, they help parents understand and evaluate their learning progress through periodic monitoring and assessments. Adjustments and/or changes in parents' educational plans are made as necessary during the year. Even Start staff were asked to name indicators of success for the adult education components and explain their choices. One staff member replied that a success indicator would be that a parent obtained a full- or part-time job she wanted or "employment with satisfaction." This same person said that another indicator would be that parents "read something about 'working' with children, they tried it out, and it worked." Another staff member responded with possible indicators of number of GEDs received, increases in reading test scores, increased interest in and

Table 3

Descriptive Statistics for Sample's Adult Education Contact Hours During Three Program Years

Adult Ed Contact Hours Time Period	Number Families	Minimum Hours	Maximum Hours	Mode Hours	Median Hours	Mean Hours	Standard Deviation	Coeff. of Variation	Std. Error of Mean
1994-1995 Program Year	7	2	516	2	70.00	147.51	185.97	1.261	70.29
1995-1996 Program Year	16	1	900	5	12.00	110.44	241.84	2.190	60.46
1996-1997 Program Year	10	7	455	9	12.00	79.60	151.27	1.900	47.84
Three Years Combined	16	1	1,542	18	23.00	224.75	451.25	2.008	112.81
Two Treatment Years	11	18	1,087	18	24.00	257.55	401.57	1.560	121.08

enjoyment of reading, and more interest in reading books. A third staff person echoed some of what others said but added "improvement in basic literacy skills that provides new levels of independence for adults." The fourth staff member agreed with others that attaining GEDs is one indicator and that obtaining a job through training is another.

Parenting Education

The Even Start staff members plan and complete interactive learning experiences in the area of parenting education for the adults in their assigned families. These parenting education experiences are designed to be an integral part of each home visit, thus they are planned regularly and conducted throughout the program year.

Planning for the parenting education experiences in the home visits begins with the Family Educator's assessment of the particular needs of the family. That is, if child discipline is viewed as a major need of the family, that topic is addressed first. On the other hand, if knowledge of child development stages is needed immediately, then such knowledge is taught first. Also, the Even Start staff are very aware of and sensitive to each family's situation and how the context of parenting education might be positively or negatively influenced by those contextual issues. For example, although the Family Educators are parents themselves, they are taught not to judge their Even Start

families on their personal values and belief systems, but rather to accept their clients where they are. Another way that the needs of the families are addressed is that the Even Start program has their parenting materials prepared at two different reading levels to obtain a match with the parent's reading skill level.

The Even Start staff recognize that parenting and child raising topics are sensitive, yet crucial, elements in their program. Typically, program staff try to wait until the parents asks child raising-type questions before they plan any such learning experiences in the home visits. So, issues in parenting or child rearing asked by parents serve to provide lead-ins to the PAT or other curriculum. This process involves building and maintaining a trusting relationship with each family. Other parenting curricula available to the Family Educators include the Portage curriculum materials and the Megaskills curriculum. Combined with the PAT program, there are enough curriculum materials available to address a wide variety of parenting issues. Still, staff do see ways to adapt, adjust, or develop new curriculum materials that meet the needs of their families.

After the parenting education experiences are planned and taught initially, they are discussed with the parents in terms of their applications. Most staff members reported that they attempted to model the parent behaviors during their instruction in the home visits. The purpose was to teach, demonstrate, and model new parenting skills for the adult(s) in the home during the visit. As one Family Educator put it, "The goal was to make the mom feel good at being a mom." As appropriate, the Even Start staff members might discuss applications of lessons taught by weaving examples from his/her own family—ever sensitive to not be "preachy" or to appear to have a "better family," but to illustrate an application from a personal experience.

Applications of parenting education learning experiences also are planned to take place in group settings. The Monongalia County program staff occasionally organizes group meetings for parents in the program. These group sessions allow for purposeful networking among the parents. But, another important purpose is to teach parenting skills in a large-group mode to supplement the individual instruction in the homes. For example, one such group session focused on how to make homemade toys for their children, while another group session was held at the public library where every participant received an individual library card, lessons on how to use it, and information on the wide variety of resources available in the library, in addition to books for themselves and their children.

Table 4 displays the descriptive statistics for the sample's adult parenting education contact hours for the three program years individually, combined, then for just two treatment years. Like the prior table, the number of families ranged from 7 to 16, but the contact hours data were just for the mothers in these families. The number of parenting education contact hours varied, but not nearly as much as adult education contact hours in the prior table. The minimum hours ranged from 3 in the 1996-1997 program year to 29 for the two treatment years. The maximum contact hours ranged from 40 to 186 for the same respective years as the minimums. The mode hours ranged from 7 for the three years combined up to 48 hours for the 1995-1996 program year. The median hours ranged from 19.50 for the 1996-1997 program year to 54.00 for the three years combined. The mean hours

Table 4

Descriptive Statistics for Sample's Adult Parenting Education Contact Hours During Three Program Years

Parenting Ed Contact Hours Time Period	Number Families	Minimum Hours	Maximum Hours	Mode Hours	Median Hours	Mean Hours	Standard Deviation	Coeff. of Variation	Std. Error of Mean
1994-1995 Program Year	7	8	62	23	23.00	27.43	16.98	.619	6.42
1995-1996 Program Year	6	7	48	48	23.00	26.06	13.75	.528	3.44
1996-1997 Program Year	10	3	40	21	19.50	18.60	9.56	.514	3.02
Three Years Combined	16	7	92	7	54.00	49.69	22.68	.456	5.67
Two Treatment Years	11	29	186	29	53.00	61.73	44.50	.721	13.42

ranged from 18.60 for the 1996-1997 year up to 61.73 for the two treatment years. The standard deviations ranged from 9.56 for 1996-1997 to the large 44.50 for the two treatment years. This caused the coefficient of variation values for the two treatment years to be the largest at .721 and the smallest was for the three years combined at .456. The standard errors of the means ranged more so, from 3.02 for the 1996-1997 year to 13.42 for the two treatment years.

Even Start staff were asked which parenting education activities were most effective and why. One staff person said that the most successful activities were ones in which the "parents learned to open up in their assessments and aspirations for their children." The staff member also said that if she could get a parent to try one new thing with the child during the visit or during the two weeks between visits, that was a real success to her. Another staff member replied that the successful activities resulted in the parent modeling the behaviors taught or the ways of talking to the children. As an example, how to elicit "more than one word answers to questions or commands" was a success. A third supplied two examples: (1) "Helping parents set limits for children," and (2) "Helping children get ready for school." Discussions with parents on specific topics was the response of a fourth Even Start staff member. Finally, one person gave a little longer response:

> The effective ones are those that are most relevant to each family. For example, getting the family to actually see the Family Educator use a new or different way to work with a child, such as holding a disruptive child in her lap and working with this child calmly. Then, this modeling is discussed with the parent to help explain this "new" way of parenting.

Interagency Collaboration

The Monongalia County Even Start Program staff regularly work with a wide variety of different service agencies in the county and region regarding the multiple needs of the client families. Program staff have met with all such other family and child service agencies in the region to explain the Even Start program and the families enrolled in it. At these meetings, the Even Start staff obtained clarifications on what those other service agencies could provide and to whom. These other service agencies represent health services, mental health services, welfare department, employment agencies, substance abuse organizations, literacy organizations, preschools, Title I, Head Start, and Early Head Start.

From these initial meetings with collaborators in other agencies, working relationships have developed with the Even Start staff. These interagency relationships result in sharing appropriate information about family members for the purpose of improving the delivery and receipt of services to them. This sharing of appropriate information about Even Start family members is accomplished by phone calls, face-to-face situations, and in special meetings. Through these various interagency connections, Even Start staff are able to make referrals for their family members. Examples include vision and hearing testing, job connections, substance abuse recommendations, special education placements, mental health referrals, family member abuse, nutritional education training, and other health-related referrals.

In addition to collaboration with other family and child service agencies for specific client family-related issues, the Even Start program staff organizes other interagency meetings to seek input to the design and delivery of the Even Start program. One regularly-scheduled interagency collaborators meeting is the annual breakfast meeting. Appendix E is three items from the 1996 annual collaborators breakfast held on December 6, 1996. The first item in Appendix E is the colorful announcement/invitation sheet (copied in black and white in this report). This announcement/ invitation clearly states that the purposes of the meeting are to (1) present current information on the Even Start grant and (2) seek recommendations in serving the Even Start parents and children in the community. The second item in Appendix E is a list of all those agencies and persons in those agencies who responded affirmatively to the invitation and were signed up to attend the meeting. The third item is a reduced copy of the larger discussion group worksheet that was completed by the cross-agency groups after the Even Start program update was given. The worksheet asked each group to ask and discuss four issue-oriented questions by the three essential components of early childhood, adult, and parenting education. For example, the collaborative discussions groups' first question was, "Upon entrance to your program, what compelling literacy needs do you often find in low income families?" The completed worksheets were collected by Even Start staff for later analysis and summarization.

One outcome of the meeting was that, by working together to hear about and then discuss Even Start program components and their families, staff in collaborating agencies can identify overlapping areas of services to families. The goal, then, would be to develop common purposes across the agencies and, ultimately, share in decisions regarding the directions of resources for services to families. Although this goal may take some time to accomplish, certainly the annual collaborators breakfast demonstrated the Even Start program's willingness to address and advance the issue.

The Even Start program staff work with local businesses to help in achieving program goals. For example, with respect to family literacy and resources for the same in the homes, program staff cooperates with a local bookstore in sponsoring an "angel tree" each Christmas/Holiday season. In this project, the local bookstore constructs and displays a Christmas/Holiday Tree in the store with various paper angels representing Even Start children as ornaments. Customers are encouraged to purchase a book or books targeted to a specific "angel" and the store sees that those books are given to Even Start staff to deliver as Christmas/Holiday presents to client children. As reported at the collaborators meeting, 149 holiday gifts of books were presented to Even Start families the previous holiday (see Appendix A). Other local businesses that program staff work with for services include dairy mart stores, convenience stores, shoe repair shops, and various food stores.

Staff were asked to evaluate the interagency collaboration aspect of their Even Start program by identifying what activities with other agencies have been most beneficial to client families and why. One staff member responded in general by stating that the Even Start program has no particular "turf" to protect regarding its families, so it is not seen as a competing agency for adult and child services with other agencies. Thus, "we have helped families connect with new or different agencies in the county." Another opined that "the personal visits and face-to-face meetings [with other agencies] were best." A third staff member said that "referrals, in both directions, work best for extra services or how to connect to them." A fourth staff member said that the Women, Infants, and Children (WIC) collaboration was "especially effective" and the adult basic education connection also was effective.

A Head Start collaborator agreed that Even Start program has aided in her program. She said that some of the parents might not have sent their children to Head Start if they did not have the Even Start experiences. Communications across two agencies is "very effective," she continued, citing how the Even Start staff conduct previews of Head Start classes for parents to encourage participation and, also, visits to the classes after enrollment. This interviewee liked how the Even Start staff were in the homes of their families on a regular basis, because that amount of regularity is more than Head Start provides. Given the opportunity to make a suggestion for an improvement in the collaboration across the two agencies, this Head Start staff person suggested that a formal time to share program and family information would be helpful.

Evaluation

Evaluation was found to be a very important component in the Monongalia County Even Start Program. This evaluation of the Even Start program found three distinct, operative levels of evaluation data collection and reporting. These three levels of evaluation of the local program are designed to assure ongoing program effectiveness by providing different types of program information for decision-making purposes. The three evaluation levels are family-level data, local-level evaluations, and third-party evaluation. A discussion of each follows.

The Even Start program staff consistently complete and maintain a variety of reports and logs for each family, which are compiled in a "running" family folder maintained by the assigned Family Educator. Included in each family's folder are the original intake forms and initial assessment profiles. Also, information and referrals forms from other collaborating agencies are included in the family "jackets." Next, the completed home visit report forms and contact logs are included, along with the family book list for each year. Local program staff designed one form to record all the relevant identifying information and scores from evaluation instruments completed for the family members (adults and children) as part of being in the national sample. This one convenient form has spaces for the pretest, posttest, and follow-up scores for each subject in the family. Relatedly, the annual local and national family summary reports (including the number of home visits, contact hours, and early childhood education days) are maintained electronically and in paper form, when needed.

The program staff consistently monitor, review, and evaluate their Even Start program's processes and effectiveness through several means. One of those self-evaluation means is by periodic staff meetings and, also, less formal discussion sessions. For example, staff constantly strive to look for, access, and discuss new evaluation instruments that might help in the diagnosis and assessment of the progress they are making with their families. These local evaluations were elevated in program year 1996-1997 when staff chose to conduct self-evaluation in a systematic manner by meeting periodically to review, discuss, and self-assess their program across quality considerations developed specifically for Even Start programs.

Table 5 displays the Monongalia County Program staff's self-evaluation ratings on Even Start program quality indicators in ten different areas. The ten areas and the specific quality indicators were published as a separate appendix to a publication by Dwyer (1995), based on her national study of Even Start programs. The ten areas in Dwyer's 1995 publication were (1) integration of components, (2) collaboration, (3) recruitment, (4) parent-child interactions and parenting, (5) home visiting, (6) adult education, (7) early childhood education, (8) retention, (9) transitions, and (10) staff development. As shown in Table 5, the number of specific indicators in each area varied from 9 (transitions) to 19 (parent-child interaction and parenting). Staff discussed and rated their program on each specific indicator in each area. These indicator ratings were averaged and the resultant means ranged from 1.02 (Integration of Components) to 2.15 (Collaboration) on the 5-point Likert scale (1 being Descriptive and 5 being Not At All Descriptive). The standard deviations for these mean ratings ranged from 0.43 (Staff Development) to 0.89 (Recruitment). The last column in the table shows the coefficients of variation to range from 0.321 (Collaboration) to 0.533 (Recruitment).

The outcome of staff's self-evaluation ratings in the ten areas was a summary report of their own discussions. Staff summarized the year-long self-study and produced a report that appears as Appendix F in this report. Here, staff developed action statements to complete and reported those accomplished and those yet to do—all in the ten indicator areas.

Table 5

Staff's Self-Evaluation Ratings on Even Start Program Quality Indicators in Ten Areas

Quality Program Indicator Area[a]	Number of Items in Indicator	Mean Rating[b]	Standard Deviation	Coefficient of Variation
Integration of Components	10	1.02	0.63	.525
Collaboration	13	2.15	0.69	.321
Recruitment	12	1.67	0.89	.533
Parent-Child Interaction & Parenting	19	1.37	0.60	.438
Home Visiting	15[c]	1.47	0.64	.435
Adult Education	18	1.78	0.88	.494
Early Childhood Education	14	1.36	0.63	.463
Retention	13	1.77	0.73	.414
Transitions	9	1.22	0.44	.361
Staff Development	14	1.21	0.43	.355

[a]Source: Dwyer, C. M. (1995). Quality Considerations from *Guide to Quality for Even Start Family Literacy Programs*. Portsmouth, NH: RMC Research Corporation.

[b]Rating scale was from "1, Very Descriptive" to "5, Not At All Descriptive."

[c]Because the Monongalia County Even Start Program is entirely home-based, four items dealing with center-based programs were judged Not Appropriate and, thus, were not rated by the staff.

At the third level, Even Start staff periodically contracts with a third-party evaluator to conduct evaluations of the program. Prior to this evaluation, two formative evaluations of the program were completed by advanced graduate students at West Virginia University. These formative evaluations helped staff understand how the processes of their programs were operating and how they contributed to the attainment of the program's objectives and goals. This evaluation is a summative study of the program to describe its essential components and how those components have resulted in certain educational and developmental outcomes through the administration of several instruments at different times. Staff's interest in independent, outside evaluations is evidenced by their agreement that this evaluation meet nationally-recognized *The Program Evaluation Standards* (Joint Committee, 1994).

Staff Development

Staff development is an important, essential component in the Monongalia County Even Start Program. Even Start staff have completed extensive staff development training since the program started, in a variety of areas and in topics related to the national and local Even Start program. These areas and topics of training sessions completed include program management, adult education, parenting education, early childhood education, literacy training, change leadership, evaluation, Internet training, and other topics.

Table 6 displays the hours of the Even Start staff development activities completed over four years listed by major program areas and the specific session names. Listed are the dates, hours per activity, number of staff that attended each activity, the total hours completed by all staff, and the percent of total hours for each activity and session. The six major program areas, the number of hours of training completed in the area, and the percent of those hours are of the total are program administration, 151 hours (10.80%); program evaluation, 76 hours (5.44%); early childhood/ parenting curriculum, 435 (31.12%); parenting skills, 215 hours (15.38%); literacy, 462 hours (33.05%); and adult education, 59 hours (4.22%). More than 60% (64.17) of the staff development sessions were in the areas of early childhood/parenting education (31.12%) and literacy training (33.05%).

The total of 1,398 staff development hours completed over four years equals 349.50 hours per year or 87.38 hours per staff member per year. This works out to be an average of 11.65 days of staff development per person per year. The national evaluation report for the 1994-1995 program year reported that the average number of days of in-service training per staff member was nine for administrators and eight for family specialists (Tao, Swartz. St. Pierre, and Tarr, 1997, p. 198). Thus, the Monongalia County staff's average number of days of staff development exceeded that of the national Sample Study by about three and a half days for the family educators and two and a half days for the administrators.

The Even Start staff have all completed extensive training in the Parents As Teachers (PAT) program, as shown in Table 6, and all staff are certificated in that program. Also, the staff have received extensive training and certification in the STEPS program and some in the Portage Program. Program staff also have participated in adult education training, including Peer Assistance Training and Monongalia County staff development sessions. Data in the table show that Even Start staff have participated in national and state Even Start meetings and conferences, international and national reading conferences, and national family literacy conferences. Program staff do share information and other materials that they receive in staff development sessions with their peers.

Even Start staff have input into the design and selection of their staff development sessions. In addition to the selection of national-level conferences and training described above, staff help design their local staff development sessions. For example, staff helped to select the completions of their self-assessment sessions in weekly meetings completed during the 1996-1997 program year,

Table 6

Hours of Staff Development Activities Completed Over Four Years Listed by Major Program Areas and Specific Sessions

Program Area Staff Development Activity Session Name	Date	Activity Hours	No Staff Attended	Total Hours	Percent of Total Hours*
Program Administration					
Even Start Technical Assistance Conf.	Oct. 1993	10	1	10	0.72
Assoc. for Supervision & Curriculum Devel.	March 1995	10	2	20	1.43
Internet Training	Nov. 1995	18	2	36	2.58
Making Real Change Happen	Nov. 1995	15	1	15	1.07
Even Start State Meeting	July 1996	3	1	3	0.21
Leadership for Change	Aug. 1996	25	1	25	1.79
Internet Training	Dec. 1996	4	3	12	0.86
National Title 1 Conference	Jan. 1997	15	2	30	2.15
Subtotal		**100**		**151**	**10.81**
Program Evaluation					
Even Start National Evaluation Conference	May 1993	10	2	20	1.43
CREATE National Evaluation Conference	July 1994	15	1	15	1.07
Sample Study Evaluation Conference	Oct. 1994	10	2	20	1.43
Even Start National Evaluation Conference	Feb. 1997	15	1	15	1.07
Denver II Training	Feb. 1997	2	3	6	0.43
Subtotal		**52**		**76**	**5.44**
Early Childhood/Parenting Curriculum					
Parents as Teachers Training (0-3)	June 1993	36	2	72	5.15
Parents as Teachers Training (3-5)	Oct. 1993	16	3	48	3.43
Parents as Teachers Training (0-3)	Oct. 1994	36	1	36	2.58
Natl. Assoc. for Education of Young Children	Nov. 1994	20	1	20	1.43
Parents as Teachers Training (2-5)	Feb. 1995	36	1	36	2.58
Parents as Teachers Training (0-3)	March 1995	36	1	36	2.58
STEPS Workshop on Preschool Science	May 1995	3	1	20	0.21
Natl. Assoc. for Education of Young Children	Dec. 1995	20	3	60	4.29
The Work Sampling System	Aug. 1996	12	2	24	1.72
Parents as Teachers Training (0-3)	Sept. 1996	36	2	36	5.15
Parents as Teachers Training (3-5)	Sept. 1996	16	1	16	1.14
Polaroid Portfolio Workshop	Nov. 1996	3	4	12	0.86
Subtotal		**270**		**435**	**31.12**

*Percentages have been rounded.

Table 6 (continued)

Program Area Staff Development Activity Session Name	Date	Activity Hours	No Staff Attended	Total Hours	Percent of Total Hours*
Parenting Skills					
Mega Skills Training	Fall 1993	15	2	30	2.15
Working With Families	Oct. 1994	6	4	24	1.72
STEPS Wkshp. on Nutritional Needs of Infants	May 1995	4	2	8	0.57
STEPS Wkshp. on Nut. Needs of Children	May 1995	4	2	8	0.57
Portage Project Training (0-3)	Oct. 1995	15	2	30	2.15
Child Abuse Prevention	Oct. 1996	6	4	24	1.72
Developing Capable People Workshops	Fall 1996	9	1	9	0.64
Parenting for 0-2	Nov. 1996	2	1	2	0.14
Partnership with Parents	10/96-5/97	20	4	80	5.72
Subtotal		**81**		**215**	**15.38**
Literacy					
Rethinking Literacy	Aug. 1994	25	4	100	7.15
Reading Miscue Analysis	10/94-5/95	36	4	144	10.30
National Center for Family Literacy	April 1995	15	2	30	2.15
Literacy Learning Summer Institute	Aug. 1995	25	4	100	7.15
Whole Language Umbrella Conference	July 1995	15	1	15	1.07
International Reading Association	May 1997	25	1	25	1.79
National Center for Family Literacy	April 1997	16	3	48	3.43
Subtotal		**157**		**462**	**33.05**
Adult Education					
Peer Assistance Training	Spring 1994	6	2	12	0.86
Peer Assistance Training	Fall 1994	6	1	6	0.43
Peer Assistance Training	Dec. 1994	6	1	6	0.43
Peer Assistance Training	April 1995	6	1	6	0.43
Developing Critical Thinking Skills	Nov. 1995	6	1	6	0.43
Welfare Reform Training	Dec. 1995	2	1	2	0.14
Monongalia County Staff Development Mtgs.	Aug. 1995	3	3	9	0.64
Monongalia County Staff Development Mtgs.	Aug. 1996	3	4	12	0.86
Subtotal		**38**		**59**	**4.22**
Grand Total		**698**		**1,398**	**100.00**

*Percentages have been rounded.

which was reported above in the evaluation component. Another example of staff selection of their own training and development is in their seeking and attaining training in the use of new adult and child assessment instruments, such as the Denver II assessment instrument and the Stages of Concern questionnaire and Levels of Use interview technique in the Leadership for Change training. As part of the selection of staff development sessions, the budget available each year is discussed by the full group. Based on staff's self-reported needs, the group helps decide who will attend what sessions to spread the budget as far as it can go, and plans are made for those selected to go to report back to the full group in sharing sessions.

In evaluating the staff development component, the interviewees were unanimous in their appreciation to the Even Start administrator to be allowed to meet, discuss, select, and share their training activities. All reported satisfaction with their acquiring fundamental skills to be good Even Start Family Educators, plus obtaining new skills to become better at their positions. Three of the staff cited the PAT training as being the most helpful and useful on the job, although one person tempered the response with the opinion that the program is "not full enough, plus it is very teacher-oriented—too much so." "But," she continued, "the PAT home structure was very good." A fourth staff person said the Portage Program materials were most meaningful and useful to her, and these same Portage materials were named second by another person. The first person said "These Portage workshops had a different slant/perspective to families that we didn't have till then. Plus, they had some new and different parenting materials to use." Last, staff cited cross-training and learning with collaborators with the Even Start program as being quite helpful.

FINDINGS - PARTICIPANTS' EDUCATIONAL AND DEVELOPMENTAL OUTCOMES

This section presents the evaluation findings for the educational and developmental outcomes for participants in the Monongalia County Even Start Program subsample. First, the educational and parenting outcomes for the adults in the subsample are presented. Next, the developmental and language development outcomes for the children in the subsample are presented. The third part of this section presents a summary of all the adult and child outcome findings, aided by a convenient summary table of all the outcome results.

The findings for each outcome measure in this section are presented in tables, text, and graphic displays. There is a systematic pattern in the presentation of the results for each outcome measure. A table displaying the t-test results and the effect sizes is displayed first, along with narrative copy highlighting the data in the table, with the t-test data explained first followed by the effect size results. These data are displayed and discussed in the regular order of (1) pretest to posttest, (2) posttest to follow-up and (3) pretest to follow-up. Next, a table of the scale score or raw score points gained per treatment period is displayed and discussed. Comparisons of the local evaluation findings to the corresponding findings in the national Even Start Sample Study are made next. Then, a graphic display of the score for each participant over the two treatment years is presented and discussed. To summarize, the order of the results for each adult and child outcome measure is (1) t-test results, (2) effect size results, (3) points/items per month results, (4) comparisons to national Sample Study results, and (5) graphic display of scores.

For this evaluation, data on the educational and developmental outcomes were collected from 16 Even Start families in the Monongalia County subsample. These families met several criteria to be included in the randomly-drawn national Sample Study. Outcome measures were administered to 16 mothers and 16 children (one each per family) in the subsample three times over a two-year treatment time period. Half of the subsample families started Even Start in the 1994-1995 program year and the other half started one year later. All of the outcome measures were administered, scored, and reported by local Even Start program staff, who were trained to do so. This evaluation was designed to capitalize on those data collected by local staff and compare it to national results, as appropriate. Thus, this evaluation represents two treatment years for each participant, although there was some attrition in the subsample.

Adult Educational and Parenting Outcomes

Two instruments were used to measure adult educational and parenting outcomes in this evaluation. The CASAS Life Skills assessment system was used to assess adults' reading literacy and mathematics literacy. The results of both of these CASAS achievement tests were reported in the same standard scale score points, which range from a low of 150 points to a high of 260 points. The

Home Screening Questionnaire (HSQ) was used to assess adults' parenting skills. The possible scores on the HSQ range from 0 points to 56 points for the 3- to 6-year old child version and 0 to 43 points for the 0- to 3-year-old child version.

Reading Literacy Outcomes

T-Tests and effect sizes. Table 7 displays the CASAS Reading Literacy t-test and effect size results for the adults over two treatment years. All six means in the third column are above the national high school level of 225 scale score points. The pretest and posttest results are displayed in the first set of rows. For these 13 adults, the mean score increased from 229.46 to 231.00, an increase of 1.54 scale score points. At the same time, the standard deviation increased by more than 3 points, causing the coefficient of variation to be larger at the posttest. Similarly, the posttest standard error of the mean was larger than the pretest values, by 1.004. The resulting t-value was less than 1 point, which is not close to being statistically significant at the .05 level. The posttest to follow-up results are displayed in the second set of rows. For the eight adults with both test scores, the mean scores increased from 235.38 to 238.75, an increase of 3.38 scale score points. However, the standard deviation increased only .65 of a point, which caused the coefficients of variance to be almost the same. Similarly, the standard errors of the mean were very close. The posttest to follow-up t-value was 1.32, but it was not statistically significant at the .05 level. The pretest to follow-up results are displayed in the third set of rows. The mean scores increased from 229.75 to 238.75, an increase of 9.00 scale score points over the two treatment years. The standard deviations were very close with that for the follow-up being .34 larger at 11.34. The coefficient of variance values were almost identical and the standard errors of the mean were close, although the value for the follow-up was just over 4.000, consistent with all the second values in the other rows. The pretest to follow-up t-value was 5.11, which was statistically significant at the .001 level. In summary, all three treatment periods' scores were above the national high school level (225 scale score points) and all three showed increases over their test administration times but only the pretest to follow-up's 9-point increase was statistically significant (.001 level).

Table 7 also displays, in the last column, the effect sizes for the CASAS Reading Literacy test administration over the three time periods. You will recall that the effect size is computed by subtracting the first mean score from the second mean score and then dividing the difference by the standard deviation of the first score. Tao, Swartz, St. Pierre, and Tarr (1997, p. 139) state this is another way to express the size of the gain. That is, in terms of standard deviation units, which yields a standardized gain. As shown in the last column, the pretest and posttest effect size was 0.111, which is labeled "small" by Cohen's (1977) definitions. The posttest to follow-up effect size was 0.315, again, "small" by Cohen's terms. In contrast to the first two effect sizes, the pretest to follow-up effect size was 0.818, which is "large" by Cohen's definitions. In summary, the effect sizes for each of the individual treatment years was "small," but the effect size for Reading Literacy over the two treatment years was "large."

Table 7

CASAS Reading Literacy T-Test and Effect Size Results
Over Two Treatment Years

Test Administration	Number	Mean[a]	Standard Deviation	Coefficient of Variation	Standard Error of Mean	t-Value	Probability	Mean Score Difference	Effect Size
Pretest	13	229.46	13.84	.060	3.837	0.64	.536	+1.54	0.111
Posttest	13	231.00	17.46	.076	4.841				
Posttest	8	235.38	10.72	.046	3.789	1.32	.229	+3.38	0.315
Follow-up	8	238.75	11.37	.048	4.008				
Pretest	8	229.75	11.00	.048	3.890	5.11	.001	+9.00	0.818
Follow-up	8	238.75	11.34	.047	4.008				

[a]Reported in scale score points, which range from 150 to 260.

Points per month. Table 8 displays the CASAS Reading Literacy scale score points per month results for the adults over two treatment years. The pretest and posttest results are displayed in the first row. The average number of treatment months for the 13 adults was 6.86. The mean scale score points gained per month was 0.336 with a standard deviation of 1.43. Because the points per months ratio was less than 1.00 and the standard deviation was more than 1.00, the coefficient of variation was a large 4.256. The standard error of the mean was 0.395. The posttest to follow-up results in the second row show the average number of treatment months to be 7.82 for the eight adults. The mean scale score points gained per month was 0.379 with a standard deviation of 0.84. The coefficient of variation was 2.216, while the standard error of the mean was 0.294. The two-year pretest to follow-up period averaged 14.64 months for the eight adults in the subsample. The mean scale score points gained over those months was 0.617, with a standard deviation of just 0.34. This much smaller deviation and larger points gained per month caused the coefficient of variation to be 0.551, the smallest in that column. And, at 0.121, the standard error of the mean for the pretest to follow-up treatment period also was the smallest. To summarize, the eight adults in the subsample gained almost two-thirds of a scale score point per month on the CASAS Reading Literacy test over the nearly 15 months in the treatment record.

Table 8

CASAS Reading Literacy Scale Score Points Per Month
Over Two Treatment Years

Treatment Period	Average No. Treatment Months	Number of Adults	Mean S.S. Points/ Month[a]	Standard Deviation	Coefficient of Variation	Standard Error of Mean
Pretest to Posttest	6.86	13	+0.336	1.43	4.256	0.395
Posttest to Follow-up	7.82	8	+0.379	0.84	2.216	0.294
Pretest to Follow-up	14.64	8	+0.617	0.34	0.551	0.121

[a]CASAS standard scale score points range from 150 to 260.

Comparison to national results. We turn now to comparing the Monongalia County subsample adults' outcomes on the CASAS Reading Literacy measure to the outcomes of the national Sample Study. These national Sample Study results appear in Chapter 7 of the 1995 interim report (Tao et al., 1997) and are for the pretest to posttest treatment period (one year). The national Sample Study reading literacy mean score increased from 227.7 to 232.6 scale score points, or a 9.9 point increase in the first treatment year, statistically significant at the .05 level. The effect size for this gain was 0.28, which is "small" by Cohen's definition. The Monongalia County subsample's gain for the same treatment period was 1.54 scale score points, which was not statistically significant at the .05 level (see Table 7). The effect size for the local subsample was 0.111, "small" by Cohen's terms and similar to that of the national Sample Study. Thus, the size of the local subsample's gain on the Reading Literacy measure was smaller (and not statistically significant) than that of the national Sample Study, but the effect size or standardized gain was similar. The pretest to follow-up data for the national Sample Study were not published at the time of this writing to compare to the local subsample, but we do know from Table 7 that the local gain was exactly 9 points over the two treatment years, which was statistically significant at the .001 level and, at 0.818, the effect size was "large." The scale score points gained per month on the Reading Literacy measure were not reported for the national Sample Study, so direct comparisons with the Monongalia County subsample's results are not possible.

Graphic display. Figure 2 displays each subsample adult's CASAS Reading Literacy scores over the two treatment years. Most of the adults had three scores, but four had only two scores and, interestingly, three of those declined rather sharply from pretest to posttest. The general pattern for those with three scores was for them to increase gradually from pretest to posttest then again from

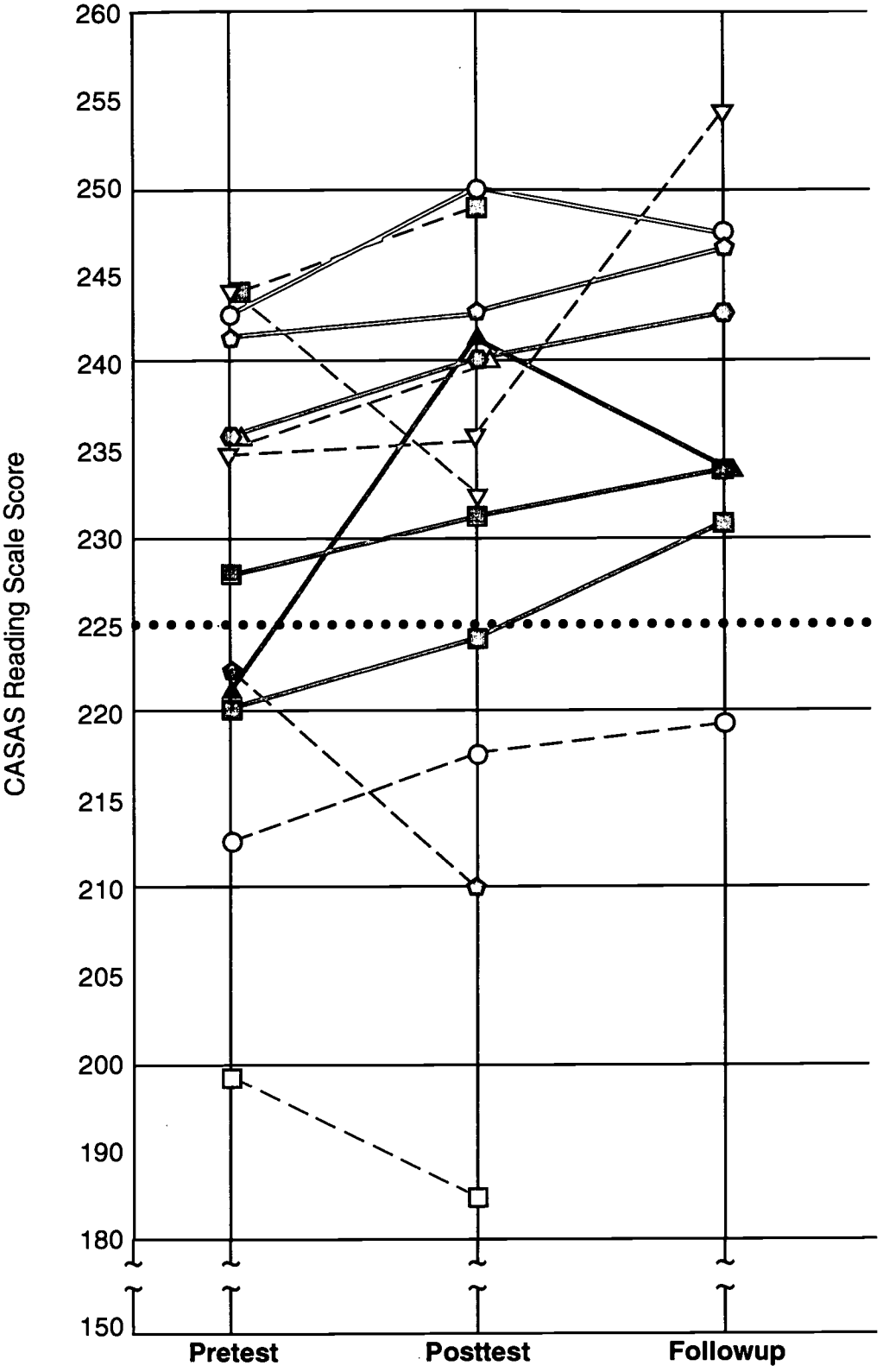


Figure 2

Graphic Display of Each Adult's CASAS Reading Literacy Scores
Over Two Treatment Years

posttest to follow-up. The few exceptions to this gradual increase are evident by their very steep increase lines, especially for one adult from pretest to posttest and another adult from posttest to follow-up (to the highest score obtained). Two adults' scores declined slightly from posttest to follow-up and one of those was the person who had a sharp increase at pretest to posttest. Figure 2 clearly shows that the vast majority of the adults' reading scores were above the national high school level of 225 scale score points and, at follow-up, only one adult in the Monongalia County subsample had a score under the national high school level. The figure also shows the rather wide and steady dispersion of scores at all three administrations.

Mathematics Literacy Outcomes

T-Tests and effect sizes. Table 9 displays the CASAS Mathematics Literacy t-test and effect size results for the adults over two treatment years. Half of the group means in the third column are above the national high school level of 225 scale score points and half of the means are below the national high school level. The third set of rows shows that the group mean moved from almost 8 points below the national high school level at pretest to 4.50 scale score points above it at follow-up. For the 12 adults in the pretest to posttest row, the mean score rose from 215.50 to 222.25, an increase of 6.75 scale score points. The standard deviation rose from 11.17 to 13.96, causing the coefficient of variation to be slightly larger at posttest. Also, the standard error of the mean was larger at posttest. The resulting t-value was 4.23, which was statistically significant at the .001 level. The mean score for the eight adults in the posttest to follow-up rows increased from 225.25 to 229.50 for a gain of 4.25 scale score points. The standard deviation dropped 3.69 points, causing the coefficient of variation to drop from .061 to .044. Similarly, the standard error of the mean decreased by 1.304 points. At 1.82, the t-value was not statistically significant at the .05 level. The pretest to follow-up mean scores increased from 217.13 to 229.50 for a gain of 12.38 scale score points. The standard deviation dropped 1.65 points and the coefficient of variation values were only .010 apart. The standard error of the mean also declined from 4.142 to 3.561. The pretest to follow-up t-value was 5.15, which was statistically significant at the .001 level. In summary, the mean score increased over 12 scale score points from pretest to follow-up, from below the national high school level to above it, and this increase was statistically significant at the .001 level, similar to the more than 6-point pretest to posttest increase.

Table 9 also displays, in the last column, the effect sizes for the CASAS Mathematics Literacy test administrations over the three time periods. Effect sizes are another way of expressing the size of the gain, that is, in terms of standard deviation units, which yields an index of standardized gain. The pretest to posttest effect size was 0.604, which is "medium" in Cohen's terms. The posttest to follow-up effect size was 0.309, which is "small" in Cohen's scheme. This second treatment year effect size is about half of the first year effect size. In contrast to these two effect sizes, the pretest to follow-up effect size was 1.056, which is "large" in Cohen's scheme. Thus, the more than 12-point pretest to follow-up increase in the adult's Mathematics Literacy score was "large" over the two treatment years, while the gains for the first and second treatment years individually were "medium" and "small," respectively.

Table 9

CASAS Mathematics Literacy T-Test and Effect Size
Results Over Two Treatment Years

Test Administration	Number	Mean[a]	Standard Deviation	Coefficient of Variation	Standard Error of Mean	t-Value	Probability	Mean Score Difference	Effect Size
Pretest	12	215.50	11.17	.052	3.225	4.23	.001	+6.75	0.604
Posttest	12	222.25	13.96	.063	4.029				
Posttest	8	225.25	13.76	.061	4.865	1.82	.112	+4.25	0.309
Follow-up	8	229.50	10.07	.044	3.561				
Pretest	8	217.13	11.72	.054	4.142	5.15	.001	+12.38	1.056
Follow-up	8	229.50	10.07	.044	3.561				

[a]Reported as scale score points, which range from 150 to 260.

Points per month. Table 10 displays the CASAS Mathematics Literacy scale score points per month results for the adults over two treatment years. The average number of treatment months for the 12 adults in the pretest to posttest year was 6.86. The mean scale score points gained per month was 0.987 or nearly one point per treatment month. The standard deviation was 0.760 and, at 0.770, the coefficient of variation was very close to it. The standard error of the mean was 0.219. The posttest to follow-up results in the second row reveal the average number of treatment months to be 7.82 for the eight adults. The mean scale score points gained per month was 0.244, considerably less than for the first treatment year. Also, at 1.225, the standard deviation was much larger, causing the coefficient of variance value to be a large 5.020. Too, the standard error of the mean was 0.433, almost double that of the first treatment year. The two-year pretest to follow-up period averaged 14.64 months for the eight adults. The mean scale score points gained over those months was 0.797, with a standard deviation of 0.443 points. This was the smallest standard deviation in the table, causing the coefficient of variance value also to be the smallest at 0.556. Too, the standard error of the mean, at 0.157, was the smallest of the three. In summarizing the data in Table 10, the eight adults in the subsample gained nearly eight-tenths of a scale score point per month on the CASAS Mathematics Literacy test over the nearly 15 months in the treatment period. This gain was nearly two-tenths of a point larger than the gain experienced in Reading Literacy for the same treatment period.

Table 10

CASAS Mathematics Literacy Scale Score Points Per Month
Over Two Treatment Years

Treatment Period	Average No. Treatment Months	Number of Adults	Mean S.S. Points/ Month[a]	Standard Deviation	Coefficient of Variation	Standard Error of Mean
Pretest to Posttest	6.86	12	+0.987	0.760	0.770	0.219
Posttest to Follow-up	7.82	8	+0.244	1.225	5.020	0.433
Pretest to Follow-up	14.64	8	+0.797	0.443	0.556	0.157

[a]CASAS standard scale score points range from 150 to 260.

Comparison to national results. Now we move to comparing the Monongalia County subsample adults' outcomes on the CASAS Mathematics Literacy measure to the outcomes of the national Sample Study (Tao et al., 1997, Chapter 7) for the pretest to posttest period only. The national Sample Study Mathematics Literacy mean score for that period increased from 217.8 to 224.5 scale score points, or a 6.9 point increase, which was statistically significant at the .05 level. The effect size for this gain was 0.49, which is the highest possible "small" value. The Monongalia County subsample's gain for the same treatment period was 6.75 scale score points, which was statistically significant at the .001 level (see Table 9). The effect size for the local subsample was 0.604 or "medium" by Cohen's terms and slightly larger than that of the national Sample Study. Thus, the size of the local subsample's pretest to posttest gain on the Mathematics Literacy measure was almost exactly the same as that of the national Sample Study and both were statistically significant (.05 level for national and .001 level for local). Also, the effect sizes, or standardized gains, differed in that the national sample's was nearly "medium," while the local subsample's was almost halfway between "medium" and "large." The pretest to follow-up data for the national Sample Study were not published at the time of this writing to compare to the local subsample, but we know from Table 9 that the local gain was 12.38 scale score points over the two treatment years, which was statistically significant at the .001 level and, at 1.056, the effect size was "large." The scale score points gained per month on the Mathematics Literacy measure were not reported for the national Sample Study, so direct comparisons with the Monongalia County subsample are not possible.

Graphic display. Figure 3 displays each subsample adult's CASAS Mathematics Literacy scores over the two treatment years. (Note: There is no correspondence among the colors of the symbols and lines used in this graphic display to those used in all the other displays: Each is independent of the other.) Most of the adults had three scores in the graphic, but four had only two scores. Interestingly, in contrast to the Reading Literacy graphic display, three of the adults with only two scores had increases from pretest to posttest, while one score remained the same. The general pattern for the adults' scores in the figure was to increase rather evenly from pretest to posttest and then increase a little less so, or level off, or even decline slightly from posttest to follow-up. But, there were a few exceptions. For example, two adults' scores increased very dramatically from pretest to posttest (231 to 246 and 219 to 237). Another example, one adult's score declined slightly from pretest to posttest and two adults' scores increased rather dramatically from posttest to follow-up (209 to 226 and 207 to 217). Figure 3 clearly shows that all but two adults' scores were under the national high school level at pretest, but all but five were above that level at posttest and, further, all but two were above the national high school level at follow-up. The figure also shows the rather wide and steady dispersion of scores at all three administrations, very similar to that of the Reading Literacy spread of scores.

Parenting Outcomes

T-Tests and effect sizes. Table 11 displays the Home Screening Questionnaire (HSQ) for 3 to 6-year-old child t-test and effect size results for the adults over two treatment years. For the 11 adults in the pretest to posttest rows, the mean score increased from 36.91 to 39.73, a gain of 2.82 points on the 56-point maximum score. The standard deviation decreased slightly from 7.61 to 7.43, causing the coefficient of variation to decline from .206 to .187. Also, the posttest standard error of the mean (2.240) was a little less than that of the pretest. The t-value for the pretest to posttest gain was 4.82, which was statistically significant at the .001 level. The mean score in the posttest to follow-up rows increased from 43.57 to 45.00 for the seven adults, a gain of 1.43 points. The standard deviation dropped slightly from 4.28 to 3.79 points, and the coefficient of variation also dropped slightly from .098 to .084. Similarly, the standard error of the mean decreased slightly to 1.431 at follow-up. At 1.43, the posttest to follow-up t-value was not statistically significant at the .05 level. The HSQ pretest to follow-up mean scores increased from 40.86 to 45.00 for the seven adults, a gain of 4.14 points. At the same time, the standard deviation dropped 1.36 points and the coefficient of variation value declined .042, which tied for the smallest such value in the table. The standard error of the mean also declined about half a point to 1.431. The pretest to follow-up HSQ t-value was 5.00, which was statistically significant at the .002 level. Thus, HSQ parenting outcomes t-test data in Table 11 show mean score increases for all three treatment periods, and the 4.14 point increase over the two treatment years was statistically significant at the .002 level.

Table 11 also displays, in the last column, the effect sizes for the HSQ parenting outcomes measure over the three time periods. The pretest to posttest effect size was 0.370, which is "small" in Cohen's terms. The posttest to follow-up effect size was similar at 0.334, which also is "small." In contrast to the first two effect sizes, the pretest to follow-up effect size was 0.805, which just barely fell in the "large" category, according to Cohen. So, while both the HSQ parenting outcome gains were "small" for the two treatment years individually, the effect size for the 4.143 gain on the 56-point scale over the two treatment years was "large."

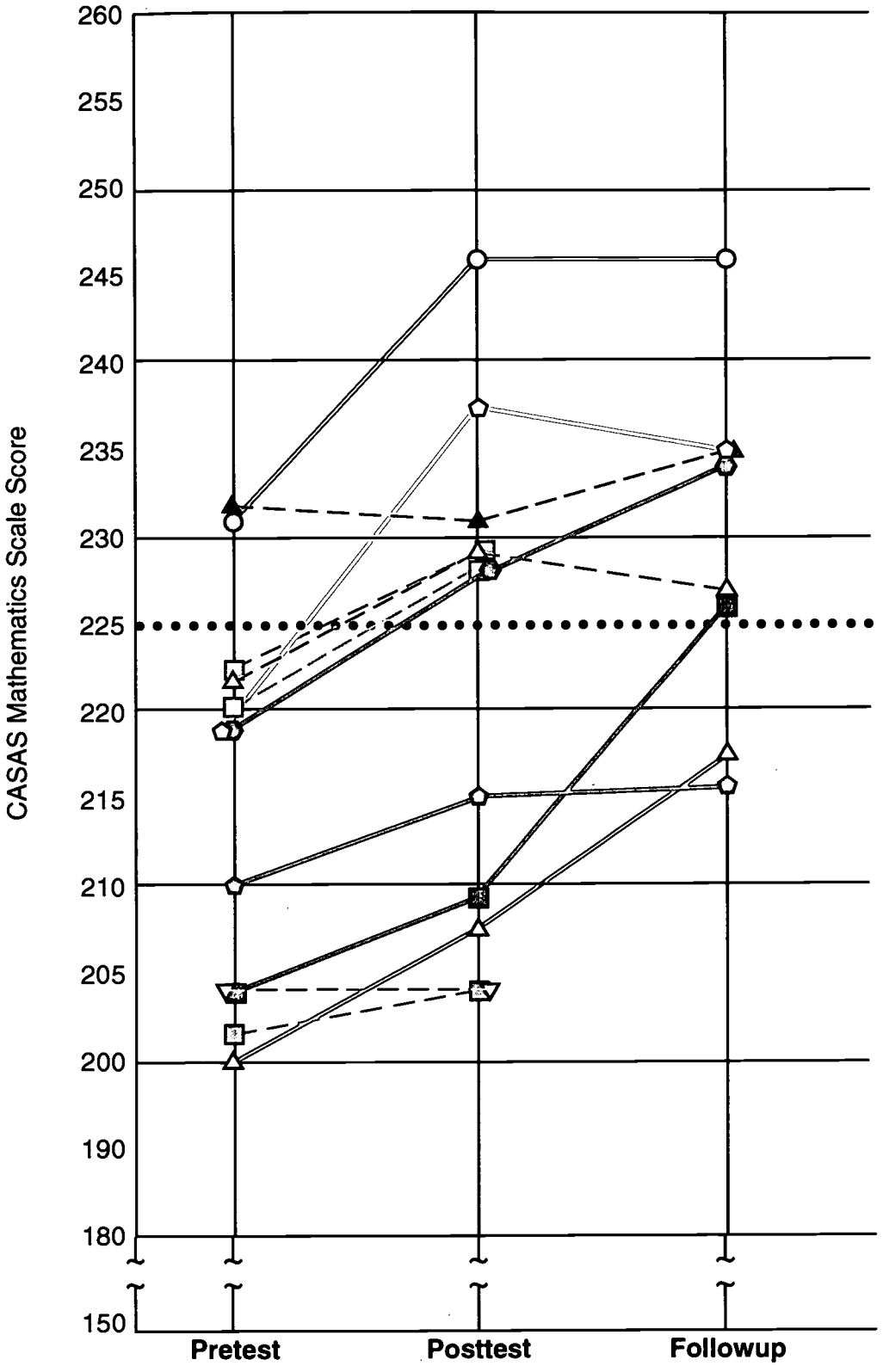


Figure 3

Graphic Display of Each Adult's CASAS Mathematics Literacy Scores Over Two Treatment Years

Table 11

Home Screening Questionnaire for 3- to 6-Year-Old Children
T-Tests and Effect Size Results Over Two Treatment Years

Test Administration	Number	Mean[a]	Standard Deviation	Coefficient of Variation	Standard Error of Mean	t-Value	Probability	Mean Score Difference	Effect Size
Pretest	11	36.91	7.61	.206	2.294	4.82	.001	+2.82	0.370
Posttest	11	39.73	7.43	.187	2.240				
Posttest	7	43.57	4.28	.098	1.616	1.43	.202	+1.43	0.334
Follow-up	7	45.00	3.79	.084	1.431				
Pretest	7	40.86	5.15	.126	1.945	5.00	.002	+4.14	0.805
Follow-up	7	45.00	3.79	.084	1.431				

[a]Reported as total raw score items, which range from 0 to 56 points.

Items per month. Table 12 displays the HSQ (3-6) items per month results for the adults over two treatment years. The average number of treatment months for the 11 adults in the pretest to posttest year was 6.86. The mean items gained per month was 0.406, or just under half of an item per treatment month. The standard deviation was 0.247 and the coefficient of variation was 0.608. The standard error of the mean was 0.075. The average number of months in the posttest to follow-up year was 7.82 for the seven adults. The mean items gained per month in that second treatment year was 0.175 and the standard deviation was 0.345. Because the standard deviation was larger than the mean, the coefficient of variation was over 1.000 at 1.971. Also, the standard error of the mean was larger at 0.131. The two-year pretest to follow-up period averaged 14.64 months for the seven adults. The mean HSQ items per month gained was 0.298, with a standard deviation of 0.137. This was the smallest standard deviation of the three, making the coefficient of variation value also the smallest at 0.460. Likewise, the standard error of the mean continued the trend at 0.052. Thus, the seven adults gained nearly three-tenths of an HSQ item per month over the nearly 15 months in the treatment period. This HSQ item per month gain over the two treatment years was about midway between the larger item per month gain in the first treatment year and the smaller gain in the second treatment year.

Table 12

Home Screening Questionnaire for 3 to 6-Year-Old Children
Items Per Month Over Two Treatment Years

Treatment Period	Average No. Treatment Months	Number of Adults	Mean S.S. Points/ Month[a]	Standard Deviation	Coefficient of Variation	Standard Error of Mean
Pretest to Posttest	6.86	11	+0.406	0.247	0.608	0.075
Posttest to Follow-up	7.82	7	+0.175	0.345	1.971	0.131
Pretest to Follow-up	14.64	7	+0.298	0.137	0.460	0.052

[a]Home Screening Questionnaire raw score items range from 0 to 56.

Comparison to national results. Here we compare the Monongalia County subsample adults' scores on the HSQ (3-6) children parenting outcomes measure to the scores of the national Sample Study (Tao et al., 1997, Chapter 7) for the pretest to posttest period only. The national Sample Study HSQ (3-6) mean score for that period increased from 34.6 to 38.0, a 3.4 item gain, which was statistically significant at the .05 level. The effect size was .48, which is "small" but just .02 below the "medium" category. The Monongalia County subsample's gain for the same treatment period was 2.818 items, which was statistically significant at the .001 level. The effect size for the local subsample was 0.370 or "small" by Cohen's terms. Thus, although the size of the local subsample's pretest to posttest HSQ gain was slightly less than that of the national Sample Study, both were statistically significant (.05 level for national and .001 level for local). Also, both samples' HSQ pretest and posttest effect sizes were "small," but the national sample's was very close to being "medium." Another comparison between the national and local samples can be made on their actual HSQ scores, not just their difference scores. For both administrations, the local sample's mean score was slightly higher than national sample's; 2.31 items higher at pretest and 1.73 items higher at posttest. The pretest to follow-up data for the national Sample Study were not published at this time to compare to the local subsample, but we do know from Table 11 that the local gain was 4.143 items on the 56-point scale over the two treatment years, which was statistically significant at the .002 level. Also, at 0.805, the effect size was "large." The HSQ items gained per month were not reported for the national Sample Study, so direct comparisons with the Monongalia County subsample are not possible.

Graphic display. Figure 4 displays each subsample adult's HSQ scores over the two treatment years. Here, in contrast to the two prior HSQ tables, both the 0- to 3-year-old and the 3- to 6-year-old children's HSQ scores are presented for the reader's convenience. Recall, though, that the maximum points differed in that 43 was the maximum for 0-3 and 56 was the maximum for 3-6. The majority of the adults had three HSQ scores, but five had only two scores—all pretest to posttest scores. The general pattern for the HSQ scores was to increase rather gradually and evenly from pretest to posttest, then increase a little less so from posttest to follow-up. There were a few exceptions to this general pattern, but only for two adults with all three scores. Both of these scores declined very slightly (two and one items only) from posttest to follow-up. The 2-item decline was for the one adult with all three scores above 50 on the 56-point maximum. In contrast to the prior CASAS Reading Literacy figure, those adults with just the two scores experienced increases from pretest to posttest, not sharp declines. Figure 4 shows the wider dispersion, or spread, of HSQ scores at pretest when compared to the scores at posttest and follow-up. Also, the figure clearly shows those adults' scores who clustered around the mean score and the four clear outliers (one high and three low). Overall, the rather steady progression of increasing HSQ scores is displayed vividly in the figure.

Child Development Outcomes

Two instruments were used to measure child development outcomes in this evaluation. The PreSchool inventory was used to measure children's readiness for school outcomes and it had a range of possible scores from 0 to 32 points. The PreSchool Language Scales-3 (PLS-3) were used to measure children's language development. Three PLS-3 scales were employed in this evaluation: (1) Auditory Comprehension, (2) Expressive Communication, and (3) Total Language score. For the first two PLS-3 scales, the raw scores are converted into standard scores based on the age of the child. These standard scores ranged from 50 to 150 points, with a normed mean of 100 points and a standard deviation of 15 points. The Total Language PLS-3 score is derived by adding the two other scales' raw scores, then converting the sum to the same standard score range of 50 to 150 points, with the same mean of 100 points and the same standard deviation of 15 points.

Readiness For School Outcomes

T-Tests and effect sizes. Table 13 displays the PreSchool Inventory (PSI) t-test and effect size results for the subsample children over the two treatment years. The mean PSI scores for the 11 children in the pretest to posttest rows increased from 18.09 to 26.00, an extremely large gain of 7.91 points on the 32-point maximum score. The standard deviation dropped from 5.61 to 3.46, a 2.15 decline in one treatment year. Also, the coefficient of variation decreased from .310 to .133, and the standard error of the mean decreased 0.648 of a point to 1.044. The t-value for the extremely large gain was 6.89, which was statistically significant at the .0001 level. The mean scores in the post-test to follow-up rows increased from 25.00 to 28.27 for the 11 children, a gain of 3.27 points.

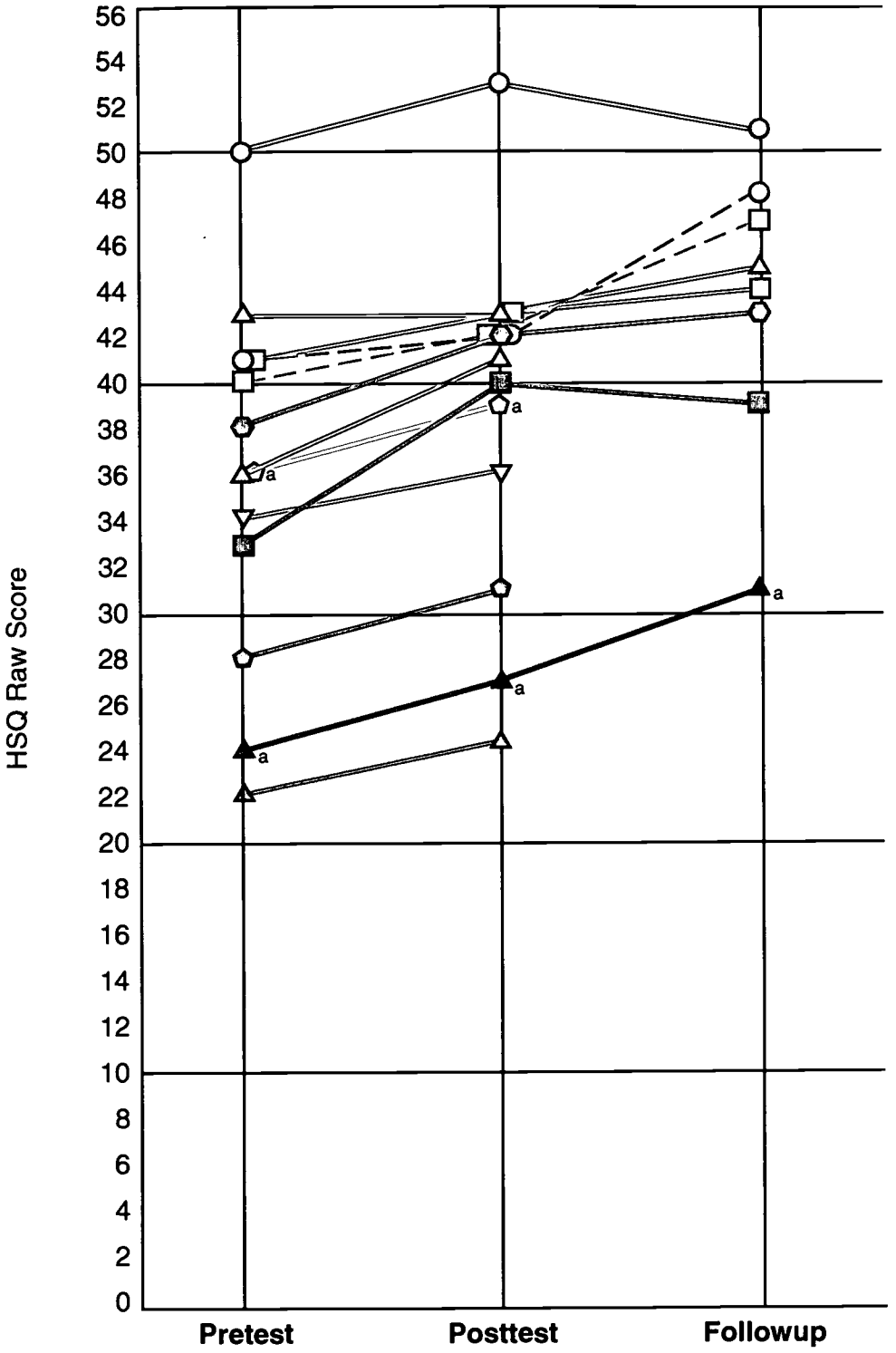


Figure 4

Graphic Display of Each Adult's Home Screening Questionnaire Scores Over Two Treatment Years

Table 13

PreSchool Inventory T-Test and Effect Size Results
Over Two Treatment Years

Test Administration	Number	Mean[a]	Standard Deviation	Coefficient of Variation	Standard Error of Mean	t-Value	Probability	Mean Score Difference	Effect Size
Pretest	11	18.09	5.61	.310	1.692	6.89	.0001	+7.91	1.409
Posttest	11	26.00	3.46	.133	1.044				
Posttest	11	25.00	5.24	.210	1.578	4.10	.002	+3.27	0.625
Follow-up	11	28.27	3.85	.080	1.161				
Pretest	10	18.40	5.82	.316	1.839	9.76	.0001	+10.70	1.840
Follow-up	10	29.10	2.85	.098	0.900				

[a]Reported as total raw score points, which range from 0 to 32 points.

The standard deviation dropped from 5.24 to 3.85, a little less than the previous decline, and the .210 to .080 decrease in the coefficient of variation value was a smaller drop than of the previous row. The standard error of the mean decreased slightly to 1.161 at follow-up. At 4.10, the posttest to follow-up t-value was statistically significant at the .002 level. The PSI pretest to follow-up mean increased from 18.40 to 29.10 for the 10 subsample children. This is an extremely large gain (10.7) on the PSI, given that its maximum is 32 points. Also, the standard deviation declined from 5.82 to 2.85, a 2.97 point drop. The coefficient of variation also dropped from .316 to .098, such that the latter was less than one-third of the former. At 0.900, the follow-up standard error of the mean was less than half of the pretest value. The pretest to follow-up t-value was an extremely large 9.76, which was statistically significant at the .0001 level. In summary, the PSI school readiness outcomes t-test data in Table 13 show very large mean score gains for all three treatment periods (two were extremley large). Further, all three t-tests were statistically significant, with those for the pretest to posttest and posttest to follow-up being statistically significant at the .0001 level.

Table 13 also displays the effect sizes (last column) for the PSI school readiness outcomes measure over the three time periods. The pretest to posttest effect size is 1.409. In Cohen's (1997) categories for t-tests, the last category defined is "large", starting at .80. His "medium" category

started at .50, so there is a .30 difference between Cohen's "medium" and "large." Given this information, we choose to label the pretest to posttest size of 1.409 as "very large." The posttest to follow-up effect size was 0.625, which is "medium" in Cohen's terms. The pretest to follow-up effect size was 1.840, which is a full 1.040 above the ***start*** of Cohen's "large" category. In fact, this difference between the PSI follow-up effect size and the beginning of the Cohen's "large" category (1.040) is 3.5 times the difference between his "medium and "large" categories (.30). Given this information, we choose to label this pretest to follow-up effect size as "extremely large." Thus, all three PSI school readiness outcome effect sizes were "medium" or more with that label given to only the posttest to follow-up treatment period. The effect size was "very large" for the pretest to posttest period and was "extremely large" for the pretest to follow-up period.

Items per month. Table 14 displays the PSI items per month results for the subsample children over two treatment years. The average number of treatment months for the 11 children in the pretest to posttest year was 6.54. The mean PSI items gained per month was 1.189, or more than one item per month on the 32-point instrument. The standard deviation was 0.540 and the coefficient of variation was 0.454. The standard error of the mean was 0.163. The average number of months for the 11 children in the posttest to follow-up year was 7.42. The mean PSI items gained per month was 0.470—less than half that of the previous year—and the standard deviation was 0.347. The coefficient of variation was 0.738, much larger than the previous year's. The standard error of the mean, though, was smaller at 0.105. The two-year treatment period was exactly 14.00 months for the 10 children. The mean PSI items per month gained was 0.764, with a standard deviation of 0.243 items. This was the smallest standard deviation of the three in the table, helping to make the coefficient of variation value also the smallest at 0.318. Similarly, the standard error of the mean was the smallest at 0.077. Summarizing, Table 14 shows that the subsample children gained more than one item per month in their first treatment year, almost half of an item per month in their second treatment year, and nearly eight-tenths of an item over the 14 months of the two-year treatment period.

Comparison to national results. Now we compare the Monongalia County Subsample children's scores on the PSI school readiness outcome measure to (1) the scores of the national Sample Study and (2) the developmental age norms, both reported in Tao et al. (1997, Chapter 7) for the pretest to posttest period only. The national Sample Study PSI mean score for that period increased from 12.7 to 18.5, a 5.8 item gain, which was statistically significant at the .05 level. The effect size was 0.90, which is "large." The Monongalia County subsample's gain for the same period was 7.909 items, which was statistically significant at the .0001 level. The effect size for the local subsample was 1.409, which was labeled "very large." Thus, while both gains over the period were rather large on the 32-point instrument, the Monongalia County subsample's statistical significance and effect size both were greater than that of the national sample. Another comparison of national and local samples can be made on the actual PSI scores, not just their differences. The national sample's pretest mean score was about five and a half points under that of the local subsample. Indeed, the national sample's posttest mean score was just .41 more than the local sample's ***pretest***

Table 14

PreSchool Inventory Items Per Month Over
Two Treatment Years

Treatment Period	Average No. Treatment Months	Number of Children	Mean Items/ Month[a]	Standard Deviation	Coefficient of Variation	Standard Error of Mean
Pretest to Posttest	6.54	11	+1.189	0.540	0.454	0.163
Posttest to Follow-up	7.42	11	+0.470	0.347	0.738	0.105
Pretest to Follow-up	14.00	10	+0.764	0.243	0.318	0.077

[a]PreSchool Inventory raw score items range from 0 to 32.

mean score. Too, the local subsample's mean posttest score was 7.50 points higher than the national sample's. The pretest to follow-up PSI data for the national sample were not published as of this writing to compare to the local subsample, but data in Table 13 reveal that the local gain was an extremely large 10.700 items on the 32-point maximum score, which was statistically significant at the .0001 level. Also, at 1.840, the effect size was labeled "extremely large." The PSI items gained per month and the developmental age norm comparisons between national and local samples are discussed in the next paragraph.

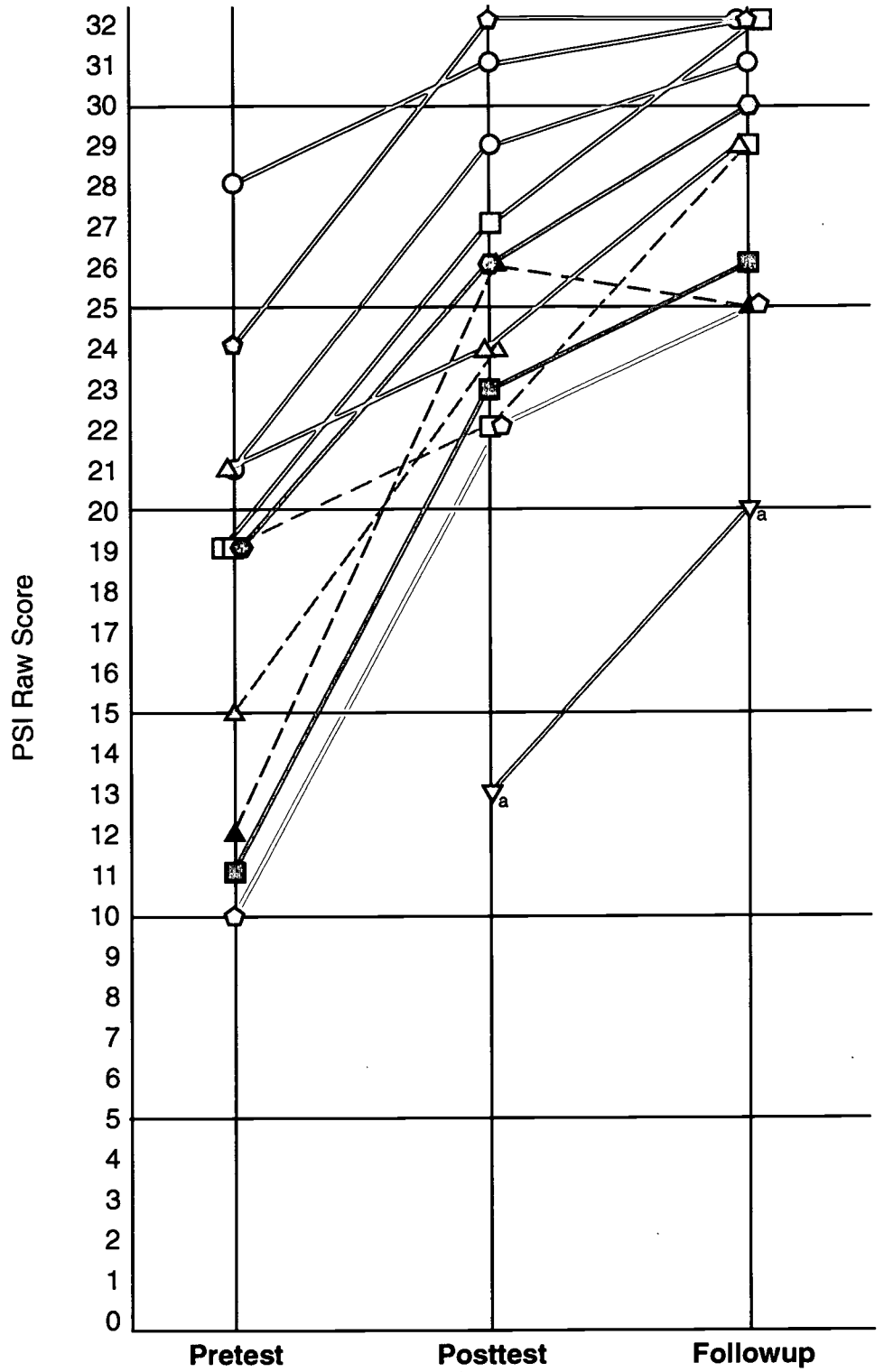
From an analysis of Even Start children's scores on the PSI, Tao et al. (1997, pp. 139 & 140) report that children are expected to gain an average of .40 of an item per month, based solely on normal development. Any gains above .40 PSI items per month are more than expected. In the national Sample Study, the average gain per month on the PSI was .94 items. Table 14 reveals that the comparable gain per month for the Monongalia County subsample for the pretest to posttest year was 1.189 PSI items per month, or 0.249 an item more than the national sample. The authors of the national sample study state their results very elegantly, "Thus, the observed gain of .94 items per month can be expressed as a combination of .40 items due to normal development and .54 items per month due to Even Start. This means that participation in Even Start more than doubles the expected rates of learning on the PSI." Borrowing freely and appreciatively from those authors, with respect to the local subsample PSI scores over the same treatment period, we find that the observed gain of 1.189 items per month can be expressed as a combination of .400 items due to normal development and .789 items per month due to Even Start. **This means that participation in the Monongalia County Even Start Program very nearly *triples* the children's expected rate of learning on the school readiness outcome measure in the first year.**

The posttest to follow-up and the pretest to follow-up PSI items per month data for the national Sample Study were not published at this time to compare to the local subsample. However, these data for the local sample are displayed in Table 14. For the posttest to follow-up period, the local subsample gained an average of .470 PSI items per month. This is considerably less of a gain on the PSI than experienced in the first treatment year, but it is slightly above what is expected due to normal development. For the pretest to follow-up period, the local subsample gained an average of .764 PSI items per month. This is more than the second year's gain but less than first year's. Over the two-year treatment period, we find that the observed gain of .764 can be expressed as a combination of .400 items due to normal development and .364 items per month due to the local Even Start program. This means that participation in the Monongalia County Even Start Program very nearly ***doubles*** the children's expected rate of learning on the school readiness outcome measure in the first two treatment years.

Graphic display. Figure 5 displays each subsample child's PSI scores over the two treatment years. Most of the children had three PSI scores, but one only had the first two scores and one child attempted the pretest, but it was not scorable. The first trend noted in the figure is the very dramatic increases in all of the children's PSI scores from the pretest to the posttest. These steep increases represent the very large items per month gains discussed above. The second trend is the increase in the posttest to follow-up scores, although these increases are not as dramatic as in the first year. Only one child's score declined in the second year, and it was only a single PSI item. The third trend evident is the rather steep increases over the two year period for all children. A fourth pattern is the narrowing of the spread of the PSI scores from pretest to posttest then again from posttest to follow-up. This narrowing of the spread of scores graphically illustrates the progressively smaller standard deviations presented in Table 13 above. Last, the figure shows that three children had reached the maximum score of 32 items by the end of the second year. In fact, one child reached the ceiling score at the end of the first year and maintained it after the second year.

Auditory Comprehension Outcomes

T-Test and effect sizes. Table 15 displays the PLS-3 Auditory Comprehension t-test and effect size results for the subsample children over the two treatment years. Of the six mean scores in the table, only the two follow-up means were above the national norm of 100 standard score points. For the 13 children in the pretest and posttest rows, their mean scores were 93.31 and 96.77, respectively, a gain of 3.46 standard score points. The standard deviation dropped slightly to 14.72 and the coefficient of variation values were very close. Similarly, the standard errors of the mean were close with the posttest's being smallest at 4.081. The t-value for the small gain was .079, which was far from being significant. The mean scores for the posttest to follow-up administration increased from 98.00 to 104.00, a 6.00 gain in that treatment year for those 11 subsample children. The standard deviation dropped .80 to 14.57 and the coefficient of variation declined .17 to .140 at posttest. The standard error of the mean decreased slightly to 4.392 at follow-up. At 1.30, the t-value was higher than the first year, but it was not significant either. The pretest to follow-up mean scores increased from 93.91 to 104.00, a gain of 10.09 points. The standard deviation dropped over 2 points from



a = Pretest administration was attempted, but it was not scorable.

Figure 5

Graphic Display of Each Child's PreSchool Inventory Scores
Over Two Treatment Years

16.77 to 14.57, while the coefficient of variance declined from .179 to .140, the largest difference in those values in the table. Similarly, the standard error of the mean dropped the most of the three rows to 4.392. The resultant t-value was 1.59, the largest of the three, but still not significant at the .05 level. Summarizing the findings in Table 15, we see that the mean gained about 3.5 points the first year, 6 points the second year, and 10.1 points over the two year on the 50-150 scale. However, none of these three gains was significant at the .05 level.

Table 15

PLS-3 Auditory Comprehension T-Test and Effect Size
Results Over Two Treatment Years

Test Administration	Number	Mean[a]	Standard Deviation	Coefficient of Variation	Standard Error of Mean	t-Value	Probability	Mean Score Difference	Effect Size
Pretest	13	93.31	15.40	.165	4.272	0.79	.446	+3.46	0.225
Posttest	13	96.77	14.72	.152	4.081				
Posttest	11	98.00	15.37	.157	4.634	1.30	.223	+6.00	0.390
Follow-up	11	104.00	14.57	.140	4.392				
Pretest	11	93.91	16.77	.179	5.057	1.59	.143	+10.09	0.602
Follow-up	11	104.00	14.57	.140	4.392				

[a]Reported as standard score points, which range from 50 to 150.

Table 15 also displays the effect sizes (last column) for the PLS-3 Auditory Comprehension scores over the three time periods. The pretest to posttest effect size was 0.225, which is "small" in Cohen's terms. The posttest to follow-up effect size was 0.390, which again is "small." However, the pretest to follow-up effect size was 0.602, which is "medium" in Cohen's terms. It is almost triple the effect size for the first year. Thus, while none of the t-tests in Table 15 was statistically significant, the effect size for the two-year treatment period was "medium." The effect size expresses the size of the gain in standard deviation units, producing a standardized gain. Stated differently, while the 10-point gain in Auditory Comprehension scores over the two years was not statistically significant, the gain had practical importance.

Points per month. Table 16 displays the PLS-3 Auditory Comprehension standard score points per month results for the subsample children over the two treatment years. For the 13 children in the pretest to posttest year, their average number of months was 6.54. The mean standard score points per month gained was 0.537 or a little more than half a standard score point per month on the 50-150 scale. The standard deviation was 2.977 and, because it was much larger than the mean, the coefficient of variation was very large at 5.544—the largest of the three values in that column. The pretest to posttest standard error of the mean was 0.826, also the largest such value in that column. The mean standard score points per month gained from posttest to follow-up for the 11 children was 0.831 or about .30 more than the first treatment year. The standard deviation was 2.136 and the coefficient of variation was 2.570, less than half that of the first treatment year. The standard error of the mean was 0.644, slightly less than the first year. The two-year pretest to follow-up period was 14.00 months for the 11 children in the subsample. The mean standard score points per month gained was 0.710 over the two treatment years. The standard deviation was 1.495 points, the smallest such value in the column. The coefficient of variation was 2.106, also the smallest in that column. Likewise, at 0.451, the standard error of the mean was the smallest of the three. Thus, Table 16 shows that subsample children gained a little more than seven-tenths of a standard score point in the 14 months of the two-year treatment period, on an instrument with a range of 100 points.

Table 16

PLS-3 Auditory Comprehension Standard Score Points Per Month Over Two Treatment Years

Treatment Period	Average No. Treatment Months	Number of Children	Mean S.S. Points/ Month[a]	Standard Deviation	Coefficient of Variation	Standard Error of Mean
Pretest to Posttest	6.54	13	+0.537	2.977	5.544	0.826
Posttest to Follow-up	7.42	11	+0.831	2.136	2.570	0.644
Pretest to Follow-up	14.00	11	+0.710	1.495	2.106	0.451

[a]PLS-3 standard score points range from 50 to 150.

Comparison to national results. Now we move into comparing the Monongalia County subsample children's scores on the PLS-3 Auditory Comprehension outcome measure with those of the national Sample Study. This comparison can be made only for the first treatment year, as that is all that is available in the national evaluation, at this writing (Tao et al., 1997, p. 143). The national

Sample Study mean score for that period increased from 88.6 to 93.3 standard score points, a gain of 4.7 points, which was statistically significant at the .05 level. The effect size was .32, which is "small." The Monongalia County subsample's gain for the same period was 3.462 standard score points, which was not statistically significant. The effect size for the local subsample was 0.225, which was "small." The local subsample's gain was from 93.31 to 96.77, both above the national norm. Thus, while the national sample's gain on Auditory Comprehension over the first treatment year was larger and statistically significant compared to the local subsample's, we note that both of the local subsample's scores exceeded the national sample's. In fact, the local subsample's pretest score was slightly higher than the national sample's posttest score. Both samples' effect sizes were "small." The national pretest to follow-up data on the PLS-3 Auditory Comprehension measure were not available at this writing to compare to the local subsample, but data in Table 15 reveal that the local gain was just over 10 standard score points on the 100-point range. And, although this 10-point gain was not statistically significant, the effect size of 0.602 was labeled "medium." Finally, the average gain per month for the national sample study was .77 points in the first treatment year, while it was .37 points for the local subsample.

Graphic display. Figure 6 displays each subsample child's PLS-3 Auditory Comprehension scores over the two treatment years. Most of the children had three scores, but two only had the first two scores. The most obvious pattern evident in the figure is that the scores are all over the place. That is, some are up, some are down, some show large differences, and some show small differences. For example, one child increased from 82 points at pretest to 125 at posttest, a 43-point gain in one year! At the same time, another child's score declined from 111 points to 89 points, a 22-point drop in the same time period. And yet another score was 109 points at both administrations. Another trend evident is the majority of the scores were below the national norm, especially at pretest and posttest administrations, reflected in their means being less than the 100 standard score points. Another trend in the figure is the movement in the scores to above the national norm (except for three) at follow-up, again reflected in the mean at 104.00 standard score points. A fourth point evident in the graphic display of scores is the posttest mode score of 88 points. Several declines in scores, some of them being relatively large declines, are portrayed in the figure. At the same time, the figure shows some relatively large increases for some subsample children, especially from the posttest to follow-up administrations. Last, Figure 6 shows the rather steady spread of Auditory Comprehension scores at all three administrations, which is reflected in the standard deviation scores in Table 15, above.

Expressive Communication Outcomes

T-Test and effect sizes. Table 17 displays the PLS-3 Expressive Communication t-test and effect size results for the subsample children over the two treatment years. All but the pretest mean score in the third column were above the national norm of 100 standard score points and, it was just .77 of a point below the national norm. The pretest mean was 99.23 standard score points and the posttest mean was 106.85 for the 13 children in the local subsample, resulting in a gain of 7.62 standard score points over the first treatment year. In the same period, the standard deviation

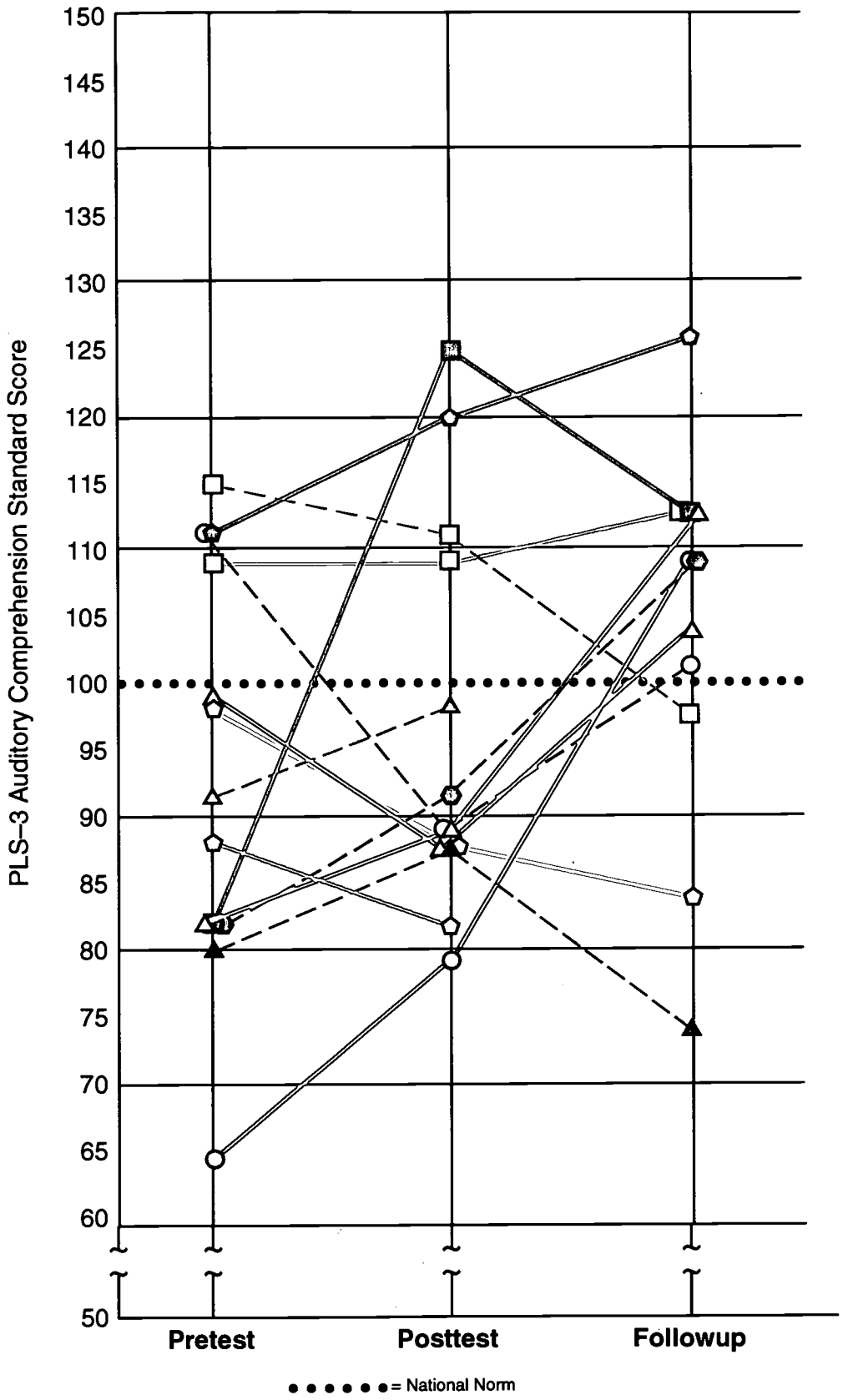


Figure 6

Graphic Display of Each Child's PLS–3 Auditory Comprehension Scores Over Two Treatment Years

declined from 16.37 to 13.01, and the coefficient of variance also declined from .165 to .122. The standard error of the mean declined almost a point to 3.607 at posttest. The t-value for the 7.62 gain was 1.82, which was not statistically significant at .05, but was approaching that level. The mean scores for the 11 children in the posttest to follow-up administrations decreased from 110.36 to 107.00, a 3.36 decline. This is the first and only decline in scores experienced in this evaluation. The standard deviation for these scores increased from 9.62 to 13.79, causing the coefficient of variance to increase from .087 to .129. The standard error of the mean also increased to 4.158. The t-value, at 0.96, was about half of the first one and, with a probability .360, was not close to being statistically significant. For the pretest to follow-up scores, the mean increased 4.0 standard score points from 103.00 to 107.00. At the same time, the standard deviation dropped slightly from 14.33 to 13.79 and the coefficient of variance declined slightly from .139 to .129. Similarly, the standard error of the mean declined slightly to 4.158 at follow-up. The pretest to follow-up t-value was 1.24, which was not close to statistical significance. Summarizing the Expressive Communication data in Table 17, we see that the largest gain of more than 7 points was in the first treatment year, but it was not statistically significant, and the mean score actually declined slightly more than 3 points in the second treatment year. The net result of this was a gain of exactly 4 standard score points over the two treatment years and, while the final score was 7 points above the national norm and the pretest score, it was not statistically significant.

Table 17

PLS-3 Expressive Communication T-Test and Effect Size
Results Over Two Treatment Years

Test Administration	Number	Mean[a]	Standard Deviation	Coefficient of Variation	Standard Error of Mean	t-Value	Probability	Mean Score Difference	Effect Size
Pretest	13	99.23	16.37	.165	4.539	1.82	.093	+7.62	0.465
Posttest	13	106.85	13.01	.122	3.607				
Posttest	11	110.36	9.62	.087	2.899	0.96	.360	-3.36	-0.350
Follow-up	11	107.00	13.79	.129	4.158				
Pretest	11	103.00	14.33	.139	4.321	1.24	.243	+4.00	0.279
Follow-up	11	107.00	13.79	.129	4.158				

[a]Reported as standard score points, which range from 50 to 150.

Table 17 also displays, in the last column, the effect sizes for the PLS-3 Expressive Communication scores over the three administrations. The pretest to posttest effect size was 0.465, which is "small" in Cohen's terms but very close to the "medium" category (starts at 0.50). The posttest to follow-up effect size was -0.350, which is "small" but negative—the only negative effect size discovered in this evaluation. The pretest to follow-up effect size was 0.279, again "small", but positive. So, while none of the effect sizes in the table was "medium" or "large," they do show movement from "positive small," to "negative small," and then back to "positive small" over the two treatment years.

Points per month. Table 18 displays the Expressive Communication standard score points per month results for the subsample children over the two treatment years. For the 13 children in the pretest to posttest year, their average number of months was 6.54. The mean standard score points per month gained was 1.175, or a little more than one point per treatment month on the 50-150 scale. The standard deviation was 2.203 and, because it was larger than the mean, the coefficient of variation was large at 1.875. At 0.611, the standard error of the mean was the largest in that column. The average number of months in the posttest to follow-up year was 7.42 for those 11 children. The mean standard score points declined in the posttest to follow-up year by an average of 0.422 points per month with a standard deviation of 1.645 points. The coefficient of variation was -0.578 and the standard error of the mean was 0.496. The pretest to follow-up treatment period was 14.00 months and the mean standard score points gain per month was 0.345, the smallest gain or loss reported in that column. At 0.732, the standard deviation was more than double the gain, resulting in the coefficient of variation being 2.121, the largest of the three. The standard error of the mean was 0.221, the smallest value in that column. So, Table 18 reveals that the subsample gained more than one point per month in the first treatment year, but over the 14 months in the two-year period, they gained an average of about three-tenths a standard score point per month on the Expressive Communication measure (100-point range).

Comparison to national results. Here we compare the Monongalia County subsample children's scores on the PLS-3 Expressive Communication outcome measure with those of the national Sample Study. This comparison is only for the first treatment year, as that is what is available for the national sample (Tao et al., 1997, p. 143). The national Sample Study scores for that period increased from 89.6 to 93.9 standard score points, a gain of 4.3 points on the 100-point range of scores. This 4.3 gain was statistically significant at the .05 level. The standard deviation increased slightly from 14.9 to 15.2. The effect size for the national sample was 0.28, which is "small." The Monongalia County subsample's scores increased from 99.23 to 106.85, a 7.62 gain that was not statistically significant, but it was approaching significance. The effect size for the local subsample was 0.465, technically "small" but very close to the "medium" value of .50. Comparing the Expressive Communication scores to the national norm of 100 standard score points, we see the local subsample's pretest score to be just under the norm and the posttest score to be almost 7 points above it. Both of the national sample's scores were under the national norm. In fact, the local subsample's ***pretest*** score (99.23) was about 5 points above the national sample's ***posttest*** score. Both samples' effect sizes were "small," but the local subsample's was very close to being "medium." The pretest to follow-up data were not available at this writing to compare to the local subsample, but Table 17

Table 18

PLS-3 Expressive Communication Standard Score Points
Per Month Over Two Treatment Years

Treatment Period	Average No. Treatment Months	Number of Children	Mean S.S. Points/ Month[a]	Standard Deviation	Coefficient of Variation	Standard Error of Mean
Pretest to Posttest	6.54	13	+1.175	2.203	1.875	0.611
Posttest to Follow-up	7.42	11	-0.422	1.645	-0.578	0.496
Pretest to Follow-up	14.00	11	+0.345	0.732	2.121	0.221

[a]PLS-3 standard score points range from 50 to 150.

shows the local gain to be exactly 4 points on the 100-point range. This 4-point gain was not statistically significant and its effect size was "small." And, both the local subsample's posttest and follow-up means were above the national norm. Finally, the average gain per month for the national sample study was .78 in the first treatment year, while it was 1.175 for the local subsample, almost four-tenths a point more.

Graphic display. Figure 7 displays each subsample child's PLS-3 Expressive Communication scores over the two treatment years. Most of the children had three scores, but two had only the first two scores. Inspection of the scores in this figure show mixed trends when compared to the three adult outcome measure displays, but not as mixed as the Auditory Comprehension scores displayed in the previous figure. For example, the general trend for the majority of the pretest to posttest scores was to increase, although four scores did decline. Another trend was that the largest increases were produced in the first year. Yet another was that two modes appeared at the posttest administration—one just under the national norm and one 15 points above it. Interestingly, the figure shows the majority of the posttest to follow-up scores to decline, but most are above the national norm, as reflected in the mean score being 107.00 (Table 17). Last, Figure 7 displays the rather steady spread of Expressive Communication scores at all three administrations, reflective of the standard deviations in Table 17.

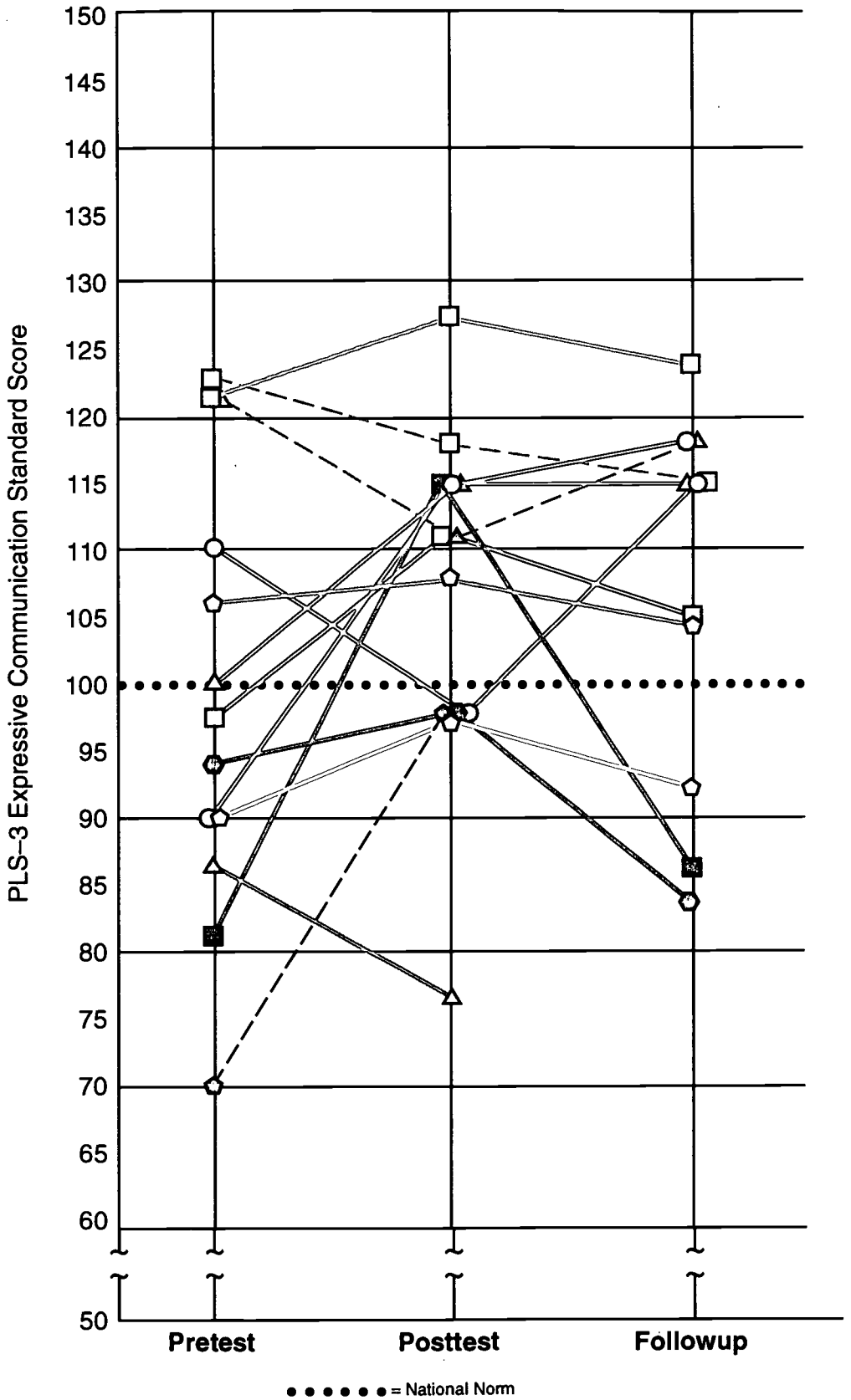


Figure 7

Graphic Display of Each Child's PLS–3 Expressive Communication Scores Over Two Treatment Years

Total Language Development Outcomes

T-Test and effect sizes. Table 19 displays the PLS-3 Total Language t-test and effect size results for the subsample children over the two treatment years. Three of the mean scores in the third column were above the national norm. For the 13 children in the first treatment year, their mean score increased from 95.85 to 99.77, or a 3.92 standard score point gain. The standard deviation dropped 4.45 points to 12.11. Likewise, the coefficient of variation dropped to .121, as did the standard error of the mean to 3.359, the smallest value in that column. The t-value for the nearly 4-point gain over the first treatment year was 1.09, which was not close to being statistically significant. The mean score for the 10 children in the posttest to follow-up administrations increased 5.60 points to 108.60. The standard deviation increased slightly to 12.69, the coefficient of variation increased slightly to .119, and the standard error of the mean also increased to 4.012. The t-value of 1.26 was larger than the first year's, but it was not close to being statistically significant. The pretest to follow-up mean scores increased from 98.50 to 108.60, a 10.10 standard score point gain over the two years. The standard deviation dropped more than 5 points to 12.69, and the coefficient of variation also dropped from .180 to .117. Too, the standard error of the mean declined from 5.616 to 4.012. The pretest to follow-up t-value was 2.03, which was not statistically significant but, at .073, was approaching significance. In summarizing the Total Language mean scores in Table 19, we see that the means increased steadily over the three time periods from almost 4 points, in the first year, to more than 5 points in the second year, and to a little more than 10 points in the two treatment years combined. Too, the mean score increased from about 4 points below the mean for the larger subsample of 13 children to more than 8 points above the national norm. And while this more than 10-point gain was not statistically significant at the .05 level, it was close to it. Also, the 10.10 gain was just over 10% of the total possible gain in points.

Table 19 also displays the effect sizes for the PLS-3 Total Language scores over the three administrations (last column). The pretest to posttest effect size was 0.237, which is "small." The posttest to follow-up effect size was 0.513, which is "medium" in Cohen's terms. Also, it is more than double the pretest to posttest effect size. The pretest to follow-up effect size was 0.569, again "medium." This two-year effect size is more than double the first year's and slightly larger than the second year's. Thus, the second-year effect size was much larger than the first year's, which contributed much to the two year's "medium" effect size.

Points per month. Table 20 displays the PLS-3 Total Language standard score points per month results for subsample children over the two treatment years. The average number of months for the 13 children in the pretest to posttest year was 6.54. The mean score points per month gained was 0.426 or just less than half a point per treatment month on the 50-150 point scale. The standard deviation was 2.004 and because it was much larger than the mean, the coefficient of variation was very large at 4.704, the largest in the table. However, at 0.556, the standard error of the mean was in the middle of the three values. The average number of months in the posttest to follow-up year was 7.42 for those 10 children. The mean points increased in that second year to 0.591 standard score points per month. With a standard deviation of 1.868, the coefficient of variance declined to 3.161, still relatively large. The standard error of the mean increased slightly to 0.591. The pretest

Table 19

PLS-3 Total Language T-Test and Effect Size Results
Over Two Treatment Years

Test Administration	Number	Mean[a]	Standard Deviation	Coefficient of Variation	Standard Error of Mean	t-Value	Probability	Mean Score Difference	Effect Size
Pretest	13	95.85	16.56	.173	4.592	1.09	.296	+3.92	0.237
Posttest	13	99.77	12.11	.121	3.359				
Posttest	10	103.00	10.92	.106	3.454	1.26	.238	+5.60	0.513
Follow-up	10	108.60	12.69	.119	4.012				
Pretest	10	98.50	17.76	.180	5.616	2.03	.073	+10.10	0.569
Follow-up	10	108.60	12.69	.117	4.012				

[a]Reported in standard score points, which range from 50 to 150.

to follow-up treatment period was 14.00 months and the mean standard score points gained per month was 0.715, the largest gain per month in the table. The standard deviation, at 1.545, was the smallest of the three. Likewise, at 0.350, the standard error of the mean was the smallest value in that column. Summarizing these data, the subsample gained slightly more points per month in the second year than the first year and, combined over the 14-month treatment period, they gained a little more than seven-tenths of a standard score point per month on the Total Language score, which had a range of 100 points.

Comparison to national results. We turn now to comparing the Monongalia County subsample children's scores on the PLS-3 Total Language outcome measure with those of the national Sample Study. As above, these direct comparisons can be made on the first treatment year only. The national Sample Study scores for that period increased from 87.8 to 93.0 standard score points, a gain of 5.2, which was statistically significant at the .05 level. The standard deviation increased by .20 to 14.60. The effect size for the national Sample Study was 0.35, which is "small." The Monongalia County subsample scores increased from 95.85 to 99.77, a 3.92 point increase over the first year. This nearly 4-point gain for the local subsample was not statistically significant. The effect size for the local subsample was 0.237, again "small", but less than the national Sample Study. Comparing the Total Language scores obtained to the national norm of 100 points, all four scores

Table 20

PLS-3 Total Language Standard Score Points Per Month Over Two Treatment Years

Treatment Period	Average No. Treatment Months	Number of Children	Mean S.S. Points/ Month[a]	Standard Deviation	Coefficient of Variation	Standard Error of Mean
Pretest to Posttest	6.54	13	+0.426	2.004	4.704	0.556
Posttest to Follow-up	7.42	10	+0.591	1.868	3.161	0.591
Pretest to Follow-up	14.00	10	+0.715	1.105	1.545	0.350

[a]PLS-3 standard score points range from 50 to 150.

were under the norm, but the local subsample's posttest score (99.77) was very close to the national norm. The pretest to follow-up data were not available at this writing to compare to the local subsample, but Table 19 reveals the local gain to be over 10 points on the 100-point range. This gain was not statistically significant (.073), but it was close to it. Further, at .0569, the effect size was "medium." At 108.60, the local subsample's follow-up score was more than 8 points above the national norm. Last, the average gain per treatment month was .89 points for the national Sample Study and was .426 for the local subsample, about half as much.

Graphic display. Figure 8 displays each subsample child's PLS-3 Total Language scores over the two treatment years. Most of the children had three scores, but three had only the first two. The overall pattern of the scores in the figure is that they are somewhat mixed across the two treatment years, but not nearly as mixed as for either of the two other PLS-3 scales. For example, seven scores increased from pretest to posttest, while six decreased in that same period. And, more of the increases were larger than the decreases. The general trend of the scores in the posttest to follow-up period is that they increased, except that two scores declined—one dramatically (24 points). Another pattern evident is that the posttest scores mainly clustered in two distinct areas, one 15 to 22 points above the national norm and a second area from 102 to 92 points. Yet another pattern is that every follow-up score except one is above the national norm. This is reflected in the mean scores being 108.60 (Table 19). Last, Figure 8 vividly displays dramatic differences in the spread of scores between the pretest administration and the two subsequent administrations. This difference in the dispersion of scores is evident in the about 5-point decline in the standard deviation from the pretest to both the posttest and follow-up administrations.

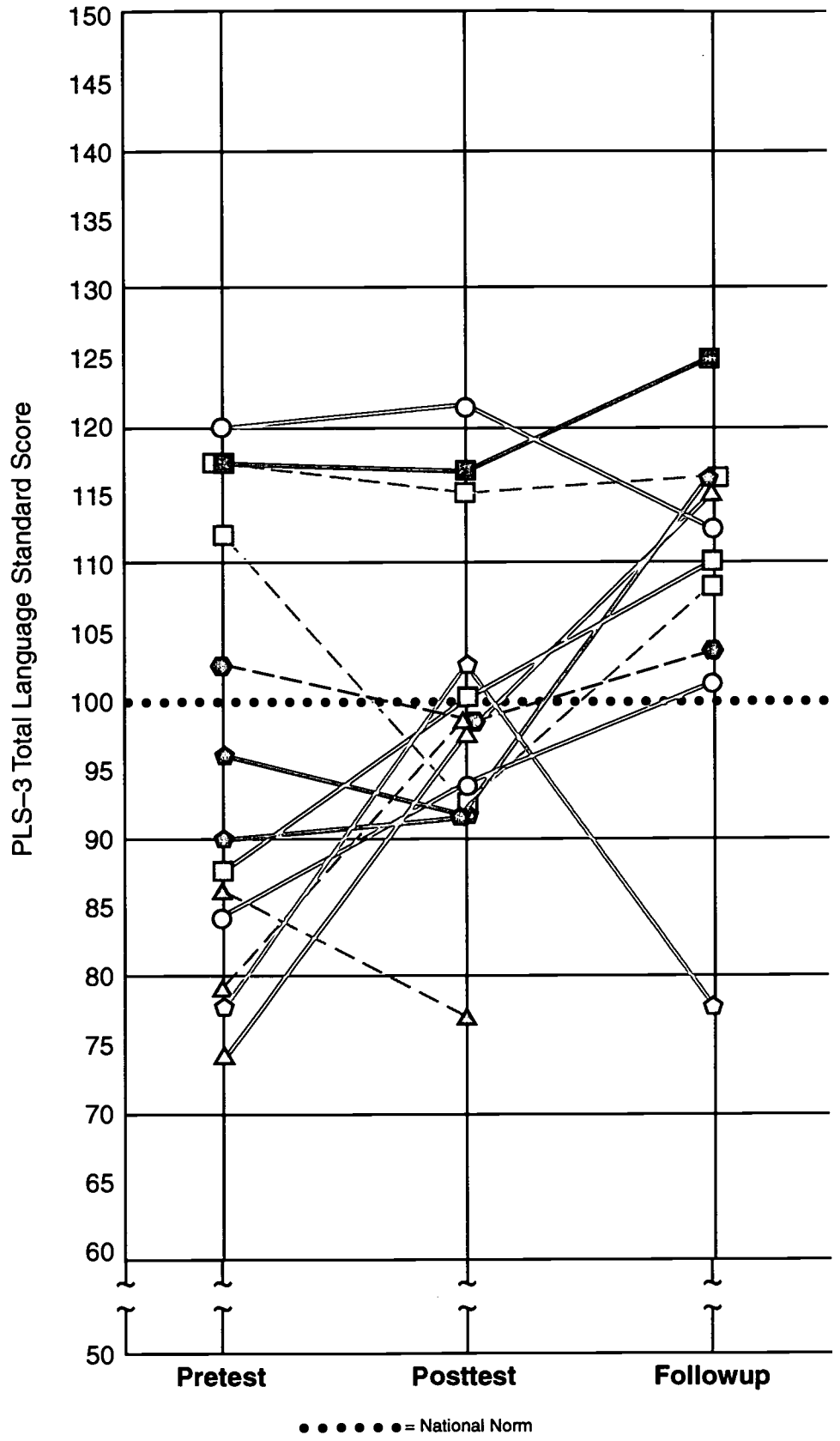


Figure 8

Graphic Display of Each Child's PLS–3 Total Language Scores Over Two Treatment Years

Summary of All Adult and Child Outcomes

Table 21 summarizes all of the evaluation findings for the adult and child outcome measures. For each outcome measure, this table displays (1) their observed score differences and their significance levels, (2) their effect sizes and descriptors, and (3) their items/scale score points per month and the direction of that value. These data are displayed by the three adult outcome measures and the four child outcome measures. For each measure, three rows of data are presented: (1) pretest to posttest, (2) posttest to follow-up, and (3) pretest to follow-up. These summary evaluation findings are discussed in the following subsections, organized by the three columns of (1) observed score differences, (2) effect sizes, (3) items/scale score points per month, and (4) differences by individual outcome measures.

Observed Score Differences

The observed score differences for all seven outcome measures by three administration periods are displayed in the third column in Table 21. There are 21 observed score differences in the column, and 20 of them are positive. Only the PLS-3 Expressive Communication posttest to follow-up value is negative, indicating a decline in scores. Of the observed score differences, eight are statistically significant at the .01 level or more; two at .0001, four at .001, and two at .01. In terms of timing, four of the statistically significant differences were at the pretest to follow-up administrations, three were at the pretest to posttest administrations, and one was at the posttest to follow-up administration. In terms of outcome measures, all three administration periods for the PreSchool Inventory experienced statistically significant increases in scores, and two administrations of both the CASAS Mathematics Literacy and the Home Screening Questionnaire experienced only statistically significant increases. The CASAS Reading Literacy measure experienced only one statistically significant increase in observed scores (pretest to follow-up administration).

Not all of the outcome measure scales reported in Table 21 are the same, but the two CASAS and the three PLS-3 measures are. For the three adult measures, the two CASAS observed score differences are directly comparable, while the HSQ scores are unique. Looking at the observed score differences for CASAS Reading Literacy (260-point maximum), we see that they increase steadily to the 9-point increase over the two-year treatment period, which was statistically significant at the .001 level. The CASAS Mathematics Literacy scores increased more in the first year (and that increase was statistically significant at .001) than in the second year, and then they increased even more, at more than 12 points, over the two years, which was again statistically significant at the .001 level. The third adult outcome measure was the 56-point HSQ. The nearly 3-point increase in the first year was statistically significant at the .001 level, the nearly 1.5-point increase in the second year was not statistically significant, and the more than 4-point increase over the two years was statistically significant at the .0001 level.

Table 21

Summary of Observed Score Differences and Their Significance Levels, Effect Sizes, and Items/Points Per Month Gained/Lost for All Seven Outcome Measures Over Two Treatment Years

Outcome Measure Scale Name	Test Administration Times	Observed Score Difference	Effect Size of Difference and Descriptor[a]	Items/Scale Score Points Per Month
Adult Subjects				
CASAS Reading Literacy	Pretest-Posttest	+1.539	0.111-Small	+0.336
	Posttest-Follow-up	+3.375	0.315-Small	+0.379
	Pretest-Follow-up	+9.000***	0.818-Large	+0.617
CASAS Mathematics Literacy	Pretest-Posttest	+6.750***	0.604-Medium	+0.987
	Posttest-Follow-up	+4.250	0.309-Small	+0.244
	Pretest-Follow-up	+12.375***	1.056-Large	+0.797
Home Screening Questionnaire	Pretest-Posttest	+2.818***	0.370-Small	+0.406
	Posttest-Follow-up	+1.429	0.334-Small	+0.175
	Pretest-Follow-up	+4.143**	0.805-Large	+0.298
Child Subjects				
PreSchool Inventory	Pretest-Posttest	+7.909****	1.409-Very Large	+1.189
	Posttest-Follow-up	+3.273**	0.625-Medium	+0.470
	Pretest-Follow-up	+10.700****	1.840-Extremely Large	+0.764
PLS-3 Auditory Comprehension	Pretest-Posttest	+3.462****	0.225-Small	+0.537
	Posttest-Follow-up	+6.000**	0.390-Small	+0.831
	Pretest-Follow-up	+10.091****	0.602-Medium	+0.710
PLS-3 Expressive Communication	Pretest-Posttest	+7.615	0.465-Small	+1.175
	Posttest-Follow-up	-3.364	-0.350-Small	-0.422
	Pretest-Follow-up	+4.000	0.279-Small	+0.345
PLS-3 Total Language				
	Pretest-Posttest	+3.923	0.237-Small	+0.426
	Posttest-Follow-up	+5.600	0.513-Medium	+0.591
	Pretest-Follow-up	+10.100	0.569-Medium	+0.715

**Significant at .01.
***Significant at .001.
****Significant at .0001.

[a]From Cohen (1977), t-test descriptors in Chapter 2 (supplemented in two cases)

For the four child outcome measures in Table 21, the three PLS-3 observed score differences are directly comparable, while the PreSchool Inventory scores are unique. Looking first at the PSI observed scores (32-point maximum), we see that they were close to 8 points in the first treatment year, which was statistically significant at the .0001 level. In the second treatment year, the PSI scores increased a little more than 3 points, which was statistically significant at the .01 level. Next, the two year treatment period increase was 10.70, and this extremely large difference was statistically significant at the .0001 level. Regarding the three PLS-3 scales (150-point maximum), we know from above that none of the increases (nor the single decline) was statistically significant across any of the nine administration times. However, since the scores are comparable, it is useful to do so. The Auditory Comprehension scores increased steadily over the three administration periods to more than 10 points on the 100-point range. The Expressive Communication scores gained more than 7 points in the first treatment year, then dropped more than 3 points in the second treatment year, and then showed a 4-point gain over the two years. The Total Language scores were almost a replication of the Auditory Comprehension scores in that they increased nearly 4 points in the first year, then increased more than 5-points in the second year, and ended up with more than a 10-point gain over the two years.

Effect Sizes and Descriptors

The effect sizes and their descriptors for all seven outcome measures by three administration periods are displayed in the fourth column in Table 21. Recall effect size is a standardized gain index, thus, all these data are directly comparable. Also, all but two of the effect size descriptors are taken from Cohen (1977, Chapter 2), and the remaining two effect sizes were large enough to merit new descriptors. Of the 21 effect sizes, 5 are "large" or more, 5 are "medium," and 11 are "small." In terms of timing, 4 of the 5 "large" or more effect sizes were for the pretest to follow-up administrations and 1 was for a pretest to posttest administration; 1 of the "medium" descriptors was for pretest to posttest, 2 were for posttest to follow-up, and 2 were for pretest to follow-up. Of the 11 "small" descriptors, 5 were for the pretest to posttest administrations, 5 were for posttest to follow-up, and 1 was for pretest to follow-up.

We turn now to the effect sizes and their descriptors for the three adult outcome measures by the three administration periods in the fourth column of Table 21. For the CASAS Reading Literacy scores, their effect sizes were "small," "small," and "large" over the three administrations. The CASAS Mathematics Literacy score effect sizes were "medium," "small," and "large." We see that the final Mathematics Literacy effect size value was more than two-tenths larger than the Reading Literacy value, even though both were described as "large." For the HSQ scores, their effect sizes were "small," "small," and "large." And, although the final effect size value was the smallest of the three pretest to posttest values at 0.805, it is just slightly less than the CASAS Reading Literacy value.

Next, we discuss the effect sizes and their descriptors for the four child outcome measures by the three administration periods. The effect sizes for the PSI were "very large" (1.409), "medium" (0.625), and "extremely large" (1.840). The PSI pretest to follow-up effect size was the largest in the table. The effect size descriptors for PLS-3 Auditory Comprehension were "small," "small," and "medium," while effect size descriptors for PLS-3 Expressive Communication were all "small. But, for the PLS-3 Total Language, they were "small," "medium," and "medium."

Items/Points Per Month

The items or scale score points per month gained or lost for the three administration periods for the seven outcome measures are displayed in the last column in Table 21. Only the two CASAS outcome measures and the three PLS-3 outcome measures are directly comparable. Of the 21 items/points per month, 20 are positive, with the one negative value being the posttest to follow-up decline in scale score points per month for the PLS-3 Expressive Communication measure. In terms of actual items/points per month gained or lost over the administration periods, only the pretest to posttest values for the PSI and PLS-3 Expressive Communication measures were more than one full point.

The item/points gained per month for the three adults outcome measures are interesting. For example, on the CASAS Reading Literacy measure, the scale score points gained per month in the first two years were nearly equal at more than one-third of a point, and over the two treatment periods it was a 0.617 point gain. In contrast, the CASAS Mathematics Literacy measure gained almost a full point per month in the first year, then dropped to .244 per month in the second year, and ended the two-year period with .797 per month gain, the fifth largest value in the column. The HSQ parenting outcome measure followed a very similar pattern to the Mathematics Literacy points per month. On the 56-point scale, the points gained per month were 0.406, 0.175, and 0.298, respectively.

The items/points gained or lost for the four child outcome measures also are interesting. As an example, the items gained per month on the 32-point PSI measure were larger by far than those on any other outcome measure. The 1.189 item per month gained on the PSI from pretest to posttest is the largest value reported in that column. In the posttest to follow-up period, the items per month gained was 0.470, somewhat of a drop, but the value rose again to 0.764 items per month gained over the pretest to follow-up period, the sixth largest value in the column. The scale score points gained per month on the Auditory Comprehension measures were 0.537, 0.831, and 0.710. These values for Expressive Communication were much different. At pretest to posttest, it was 1.175 points gained per month, the second largest value in the column. As mentioned above, for the posttest to follow-up period, the points per month declined 0.422, the only decline in the column. Then, in the pretest to follow-up period, it evened out to 0.345 points per month gained—about half of the value for the Auditory Comprehension measure. Finally, for Total Language, the points per month gained were 0.426, 0.591, and 0.715.

Differences by Individual Outcome Measures

Inspecting the results in Table 21 ***across*** the three columns for each outcome measure and administration is another way to discern important patterns in the data. Regarding adults on the CASAS Reading Literacy measure, we see the gains were small in the individual treatment years but, when combined, the two years yielded a 9-point gain, which was statistically significant; a 0.818 effect size, which is "large"; plus a point per month gain of .617. The CASAS Mathematics Literacy results were larger, but not as smooth as Reading Literacy. For example, the gains were much larger in the first treatment year (6.750 points, statistically significant at .001; 0.604 effect size, "medium;" and 0.987 points per month gained) than in the second year. Over the two treatment years, the 12.375 gain was statistically significant, the 1.056 effect size was "large," and the 0.797 point gained per month was the second largest such value for the adult outcome measures. Similar to the previous description, the HSQ parenting outcome measure had a larger (and statistically significant) gain in the first treatment year than the second year; but, the two-year period had the largest gain at 4.143 points, was statistically significant at .01, had a "large" effect size at 0.805, and an item per month gain of almost three-tenths of a item.

Relative to the four child outcome measures, the PSI scores showed the largest gains for all the child measures and the adult measures. The PSI differences were large (7.909 points, 1.409 effect size, and 1.189 items per month gained) in the first treatment year. These differences declined a little in the second treatment, but the 3.273 difference still was statistically significant. Then, over the two treatment years, the 10.700 gain was statistically significant at the .0001 level, the 1.840 effect size was "extremely large," and the items per month gained was 0.764 (sixth highest in the column). The Auditory Comprehension gained smoothly and steadily over the two treatment years to the point where the 10-point gain yielded a "medium" effect size and the eighth largest point per month gain in the table (0.710). The Expressive Communication results were the oddest in the table. In the first year, the 7.615 point gain was "small," but the second largest points per month gained in the table. Then, unexpectedly, all three values declined in the second year, to yield a 4-point, "small" increase in this two-year period. Finally, the Total Language scores were very much like those of the first PLS-3 measure. That is, the values increased steadily from the first year to the second year, then again to the two treatment years. Similar to Auditory Comprehension, the two-year Total Language gain of just over 10 points yielded a "medium" effect size with the seventh highest (0.715) point per month gain.

In summarizing the data in Table 21, we note that there is a total of 63 indicators (3 rows of 21 indicators each) of outcome measure differences. Of those 63 outcome measure indicators, 60 are positive and just 3 (all in the same row) are negative. By columns, of the 21 observed score differences, 8 are statistically significant: 2 at .0001, 4 at .001, and 2 at .01. Of the 21 effect sizes in the next column, 5 are "large" or higher, 5 are "medium," and the remaining 11 are "small." Of the 21 items/points gained per month in the last column, 20 are positive and only 1 is negative, plus 8 are more than seven-tenths of an item/point gained per month.

Finally, another way to summarize the adult and child outcomes presented in this section is to average the two-year (pretest to follow-up) effect sizes. The average two-year effect size for the three adult outcome measures is 0.893, which is "large" in Cohen's terms. The average two-year effect size for the four child outcome measures is .823, which also is "large" in Cohen's terms.

CONCLUSIONS

Two broad goals guided this evaluation of the Monongalia County Even Start Program. The first goal was to describe and evaluate the unique rural and home-based Even Start program. The second goal was to evaluate the program on the local families in the national sample in terms of their educational and developmental outcomes. The evaluation findings for these two goals were presented in the two prior sections. This section presents the conclusions drawn from those evaluation findings. These conclusions are presented by the major topic areas assessed in the evaluation plus additional areas from the methodology and instrumentation.

Innovation Configuration Matrix

The completed Innovation Configuration Matrix (ICM) for the Monongalia County Even Start Program (see Appendix D) identifies the eight essential components in the program. Those eight essential components are (1) identification and recruitment, (2) home visits, (3) early childhood education, (4) parenting education, (5) adult education/literacy, (6) collaboration, (7) evaluation, and (8) staff development. These eight components of the local Even Start program are not presented in order of importance, rather, they are presented in the order they appear in the graphic depiction, below. One Even Start program component not seen as essential as a result of this evaluation is "retention" (of families), although it might be an essential component in many center-based Even Start programs.

The Monongalia County Even Start Program is operating at high levels on all eight essential components. This is evident in the ICM by the "A" or first variation in each component's row. These eight "A" variations describe the manner in which the local program was operating when this evaluation was conducted. On the completed ICM in Appendix D, these "A" variations are separated from the other variations by the vertical dotted line. As the ICM key indicates, variations to the left of the dotted line denote ideal practices. Thus, the Monongalia County Even Start Program is operating at the ideal practice level on all eight essential components.

The ICM for this rural, home-based Even Start program provides useful descriptions of less than ideal levels of operations for the eight essential components. That is, variations "B," "C," and "D" for each component describe successively lower levels of functioning below the ideal level. Variations "B" and "C" in the ICM represent acceptable practice while variation "D" represents unacceptable practices. The solid vertical line in the ICM divides the acceptable practice from the unacceptable practice for each component.

Thus, the completed ICM from this evaluation of the Monongalia County Even Start Program could be employed not only to ***describe*** *itself to others but, moreover, to* ***differentiate*** *it from others on the basis of where other programs are operating with respect to the variations for each component.*

Even Start Program Model

The eight essential components in the Monongalia County Even Start Program can be arranged in concentric circles to depict the program in a convenient graphic manner shown earlier as Figure 1.

In the graphic depiction of this rural, home-based Even Start program, recruitment is in the center circle. Bi-weekly home visits are in the next circle. This is followed by a third circle containing the three Even Start components of (1) early childhood education, (2) adult education, and (3) parenting education. Interagency collaboration is the fourth circle, while evaluation is fifth. Finally, the outer circle is the staff development component.

The red shading of the second and third circles in the figure is purposeful because this denotes the heart of the Monongalia County Even Start model. This coloring highlights the importance of the bi-weekly home visits to the delivery of the early childhood education, adult education, and parenting education components.

Thus, we conclude that Figure 1 is an accurate graphic portrayal of the Monongalia County Even Start Program.

Adults' Reading Literacy

The reading test in the CASAS Life Skills assessment system was used to evaluate adults' Reading Literacy over the two-year treatment period. All of the Reading Literacy mean scores over the three administrations were above the national high school level. Even so, the Monongalia County subsample adults gained one and a half standard score points in the first treatment year and almost three and a half points in the second treatment year. More dramatically, over the two treatment years combined, the mean score gained 9 points. This large gain was statistically significant at the .001 level. Also, at 0.818, the effect size was "large" over the two treatment years, while it was "small" for each year individually.

Therefore, from these findings, we conclude that the Monongalia County Even Start Program had a large impact on adults' Reading Literacy, especially over the two-year treatment period.

Adults' Mathematics Literacy

The mathematics tests in the CASAS Life Skills assessment system was used to evaluate adults' Mathematics Literacy over the two-year treatment period. At the pretest administration, the subsample's mean score was about 10 points below the national high school level. Over the first year, that mean score rose 6.75 points, a statistically significant gain at the .001 level. There was a 4.25 point gain in the second year, which was not statistically significant. But, over the two-year treatment period, there was a 12.38 point increase, which was statistically significant at the .001 level. The effect size of three gains were "medium," "small," and "large."

Therefore, from these findings, we conclude that the Monongalia County Even Start Program had a large impact on adults' Mathematics Literacy, especially in the first year in comparison to the second year, and when the two years were combined.

Adults' Parenting Outcomes

The Home Screening Questionnaire (HSQ) was used to assess the adults' parenting outcomes. There is no national norm for this 56-point instrument for 3- to 6-year-old children (nor for the two adults with 0-3 year old children). Over the first year, the mean HSQ score increased almost 3 points and was statistically significant at the .001 level. There was a 1.43 point gain in the second year, which was not statistically significant. However, over the two-year treatment period, there was a 4.14 point increase, which was statistically significant at the .002 level. The effect size for each of the treatment years individually was "small," but for these two years combined, it was 0.805, which is "large."

Therefore, from these findings, we conclude that the Monongalia County Even Start Program had a large impact on adults' parenting outcomes, especially in the first year in comparison to the second year, and when the two years were combined.

Children's Readiness for School

The PreSchool Inventory (PSI) was used to assess children's readiness for school outcomes. There are no national norms for this 32-point instrument. In the first treatment year, the mean score increased nearly 8 points to just 4 points below the maximum score. This gain was statistically significant at the .0001 level, and its effect size of 1.409 was labeled "very large." There was an average gain of more than one item per treatment month in the first year. The gain in the second year was less than half of the first year, but it was statistically significant at the .02 level and its effect size was "medium." The biggest gain on the PSI was over the two treatment years, where it was almost 11 points, which was statistically significant at the .0001 level. At 1.840, the two-year effect size was labeled "extremely large." Also, the points gained per treatment period (0.764) was almost double that expected for normal growth.

Overall, then, from these findings we conclude that the Monongalia County Even Start Program had tremendous impact on children's readiness for school outcomes, especially in the first year in comparison to the second year, but even more so when the two treatment years were combined.

Children's Auditory Comprehension

The Auditory Comprehension scale of the Preschool Language Scale-3 (PLS-3) was used to assess the receptive language skills of the children in the local subsample. With a national norm of 100 standard score points, the mean score at pretest was under the norm, but at follow-up, it was above the norm by 4 points. In the first treatment year, the mean score increased 3.46 points, still under the national norm. This gain was not statistically significant and its effect size was "small." The 6-point gain in the second treatment year produced a mean score 4 points above the national norm but, again, it was not statistically significant and its effect size also was "small." When combined, though, the two treatment years yielded a 10.1 gain and, although not statistically significant, the effect size of 0.602 was "medium." The two-year gain was .710 points per treatment month.

Thus, taking all of these findings together, we conclude that the Monongalia County Even Start Program had a moderate impact on children's Auditory Comprehension outcomes, especially over the two-year treatment period rather than either year individually.

Children's Expressive Communication

The Expressive Communication scale of the PLS-3 was used to assess the verbal language skills of the subsample children. With the national norm at 100 standard score points, the pretest mean was less than one point under it, but by follow-up, it was 7 points above the national mean. The mean score increased 7.62 points in the first treatment year, which was not statistically significant and had a "small" effect size. Unexpectedly, the mean score dropped 3.36 points in the second year (the only decline across all outcomes by all treatment periods). This decline in Expressive Communication was not statistically significant and its effect size was "small." When combined, the two treatment years yielded a 4-point gain to above the national norm, but it was not statistically significant and its effect size was "small."

Therefore, considering the up and down movement of the mean scores, we conclude that the Monongalia County Even Start Program had some impact on children's Expressive Communication outcomes over the two-year treatment period.

Children's Total Language

The PLS-3 Total Language scale score was used to assess the total language development of the children in the subsample. With the same national norm of 100 standard points as its two scales, the mean was a little more than 4 points under the norm at pretest but moved up to more than 8 points above the norm at follow-up. The mean score increased nearly 4 points in the first year, which was not statistically significant and had a "small" effect size. In the second year, the mean score gained 5.6 points and, although not statistically significant, the 0.513 effect size was "medium."

When combined, the two treatment years produced a 10.1 gain to nearly 9 points above the national norm. Although not statistically significant for this small group, its effect size was 0.569, which is "medium."

Therefore, taking these findings together, we conclude that the Monongalia County Even Start Program had a moderate impact on children's Total Language development outcomes over the two-year treatment period and, further, more so over the second year in comparison to the first year.

Outcomes By Treatment Year

The amount of outcomes data collected for the adults and children in the subsample and the fact that it was analyzed by each treatment year and, then combined, permits this evaluator to study those outcomes by treatment years and draw appropriate conclusions. This means we can inspect the observed score differences, significance levels, effect size values and their descriptions, and the items/points gained or lost per treatment period for each of the three adult and four child outcome measures.

Thus, based on the various types of data generated for the adults and children in the Monongalia County Even Start Program, we conclude that, for the first treatment year, the program impacted adults CASAS Mathematics Literacy and children's readiness for schools (PSI) and PLS-3 Expressive Communication outcomes the most; for the second treatment year, the program impacted children's PLS-3 Auditory Comprehension and Total Language outcomes the most; and for both years combined, the program impacted adults' CASAS Mathematics Literacy and children's readiness for school (PSI), and PLS-3 Total Language outcomes the most.

Instrumentation

Interview protocols, paper-and-pencil achievement tests, and questionnaires were the major instruments utilized in this evaluation. The three interview protocols were developed specifically for this evaluation as major inputs (but not the only inputs) to the construction of the Innovation Configuration Matrix (ICM). The four paper-and-pencil instruments were selected as a requirement for participation in the national Sample Study. These four instruments were the CASAS Life Skills assessment system, the Home Screening Questionnaire (HSQ), the PreSchool Inventory (PSI), and the Preschool Language Scale-3 (PLS-3). The CASAS system included two scales: Reading Literacy and Mathematics Literacy. The PLS-3 had three outcome measures: Auditory Comprehension, Expressive Communication, and Total Language. The evaluator of this effort designed, implemented, and analyzed the data from the three interview protocols. However, the selection, implementation, scoring, and reporting of the four paper-and-pencil instruments were completed by the local Even Start staff. This evaluation utilized the outcome measures data as provided by staff in summary reports. So, we can draw some conclusions about the instruments employed in this evaluation.

Overall, then, we conclude that the instrumentation employed in this evaluation of the Monongalia County Even Start Program was both useful and effective. We found that the three interview protocols were very effective in constructing the ICM. Also, we found the four paper-and-pencil instruments to be very useful and effective in evaluating the adult and child outcomes locally over the two treatment years and, moreover, for making comparisons to the national Sample Study. However, we do note that three of the subsample children had reached the ceiling score on the PSI at the end of their second year in the program.

Overall Conclusion

This year-long evaluation of the Monongalia County family literacy program was an interesting experience. One of the most interesting aspects was that the local program conformed to the law that created and funded it, and participated in the national Sample Study, yet, while most Even Start programs are center-based and ***do not feature*** home visits, the local program operates in a rural environment and ***does feature home visits*** as a critical component. So, the interesting aspect for the evaluator was to discover how this home-visit component was operating in concert with the other essential components. Once it was determined what those essential components of the local Even Start program were, and how they were implemented, the major goal was to discover their impact on Even Start families—as determined by the outcome measures administered as a function of participating in the national Sample Study evaluation. Thus, we can draw on all of these various aspects of the evaluation to produce an overall conclusion.

Therefore, combining many of the findings in this evaluation, we conclude that, overall, the Monongalia County Even Start Program is a comprehensive, rural, home-based family literacy program that produces many impacts on the adults and children studied in this effort. We conclude that the Monongalia County staff have designed and implemented an effective rural, home-based Even Start program for its client families. Further, we conclude that this home-based family literacy program has applications beyond Even Start: The program could be applied in other situations with similar environments.

Fathers' Outcomes

While the reader may feel from the above that the program is perfect, and has no room for improvement, we have reserved one conclusion regarding something we ***did not*** discover in this evaluation. In all the previous sections, adults referred to mothers in the families, not fathers, which is attributable to the data collection plan as part of the national Sample Study. Outcomes data were collected on fathers, but not as complete as for the mothers. Too, adults participating in the national Sample Study had to be participating in the core services of the Even Start program (e.g. adult education and parenting education). Not all of the Monongalia County subsample fathers met that criterion and, thus, were not in the national Sample Study.

Therefore, we conclude that there is insufficient data on the fathers in the Monongalia County Even Start Program to make any statements of program impact on them. There are some service delivery and outcomes data on some of the fathers in the local subsample, but not enough to analyze at the level of the mothers in the families. We conclude that this is an area of needed improvement.

RECOMMENDATIONS

Certain recommendations can be made based on the findings and conclusions of this evaluation of the Monongalia County Family Literacy model. These recommendations are presented under the four major headings of (1) program continuation, (2) program improvements, (3) evaluation utilization, and (4) new evaluations. Although the four major headings are not in any particular order, the specific recommendations under each heading are in priority order.

Program Continuation

First, given that the essential components of the model are functioning well and given that positive outcomes were discovered on all seven measures, we recommend that the Monongalia County Even Start Program be continued in future years. This evaluation uncovered sufficient evidence of positive successes to warrant the continuation of the program.

Second, if the program is continued in future years, we recommend continuing the collection of outcome measure data on adults and children in the family. We also feel that the program should continue to collect service delivery information (contact hours and Early Childhood Education days) from Even Start adults and children.

Third, as the program continues, we recommend that the Even Start staff continue their ongoing self-assessment activities, especially evaluating the eight essential components to determine if they are continuing at the"A" variation level and, if not, to determine why they are not.

Fourth, based on the very positive findings from this evaluation, we recommend that Even Start staff consider expanding the program to other families, if more resources can be found to do so. Further, we recommend that staff use this evaluation to seek additional resources from "old" and "new" funding sources.

Fifth, Even Start staff should continue to identify, secure, and implement new/different diagnostic and outcome measures with adults and children. At the same time, program staff should continue to assess the advantages and possible disadvantages of using these new instruments in their program.

Program Improvements

First, Monongalia County program staff should strive to improve the children's PLS-3 Auditory Comprehension scores. Program staff should inspect their curriculum for receptive language lessons and discuss how and when those lessons are implemented in the home visits. The objective of this attention to receptive language lessons would be to reduce the large variance in children's PLS-3 Auditory Comprehension scores by improving children's performance on that outcome measure.

Second, we recommend that the Even Start program staff should strive to get more fathers in the families involved in the events, sessions, activities, and other components in the program. This is not to say that the program staff do not make such efforts now; rather, this recommendation is to continue those efforts and, perhaps, increase them in upcoming years.

Third, we note that the local program staff have started to develop original curriculum materials in areas they feel are needed. This recommendation calls for the continuation of these curriculum development efforts. Then, these materials should be pilot and field tested with families and implementation data should be collected.

Fourth, since "collaboration" was the lowest-rated program component in the staff's self-assessment, we recommend that staff implement that part of the action plan first. Also, we recommend that staff re-assess themselves on "collaboration" at the end of the 1997-1998 program year.

Fifth, this evaluation found that several of the subsample children had reached the ceiling on the 32-item PreSchool Inventory (PSI). This was unexpected, since it did not happen in the national Sample Study. Therefore, we recommend that program staff consider using the 64-item PSI instrument instead of the 32-item PSI. Since neither version contains any subscales (only one overall scale), the 64-item PSI scores could be halved if the local program participates in the national evaluation again, or simply to compare the local scores to national scores. The anticipated advantage of the longer PSI would be that fewer students would reach the ceiling score.

Evaluation Utilization

First, we recommend that this evaluation be submitted to the national ERIC system for accession so it will be available to others interested in evaluating rural, home-based family literacy programs. Since several ERIC Clearinghouses could be appropriate for this evaluation, we recommend that the ERIC Clearinghouse on Rural Education and Small Schools at AEL be given the charge of placing this evaluation in the correct ERIC Clearinghouse.

Second, in addition to the ERIC system, it is recommended that this evaluation report be disseminated locally, regionally, and nationally by program staff both within and outside the Even Start system. Moreover, we recommend that articles and papers be prepared on the basis of the design, implementation, and findings of this local Even Start evaluation and that they be submitted to appropriate journals and organizations as another means of dissemination.

Third, in addition to the dissemination of the full evaluation, as described above, we feel there is value in letting others know of the unique features of the Monongalia County Even Start Program. Therefore, we recommend the model of the program, as displayed in Figure 1, and the Innovation Configuration Matrix, with its variation "A" descriptions for each of the eight essential components, be disseminated widely by local program staff.

Fourth, above and beyond those normal dissemination activities already discussed, we recommend that local program staff use this entire evaluation report or relevant sections to seek some sort of national recognition for the Monongalia County program. Here, we are talking about a National Diffusion Network (NDN)-type approval. We understand that the NDN is no longer in operation, but our recommendation would be to seek its successor, such as "Proven Practices," and submit an application. A related suggestion would be to send this evaluation report, and perhaps other program reports, to the National Center for Family Literacy in Louisville, KY, seeking national recognition as a rural, home-based family literacy program that works.

New Evaluations

First, with this largely quantitative evaluation completed, we recommend that the program staff take the next year to collect case histories of participants who have been especially successful. We know of several good candidates for extensive, qualitative case histories based on the site visit trip and we suspect that there are others. The major purpose of the recommendation is to augment this strong, positive, mainly-quantitative evaluation with a strong, positive, mainly-qualitative evaluation about successful Even Start participants.

Second, if the opportunity arises, the Monongalia County Even Start Program should participate in the next national evaluation. There are numerous advantages to participating in the national evaluation study, including training in administering evaluation instruments and preparing summary reports of service intensity and outcome measures for each Even Start family and each family member. Having service delivery and outcome data for an outside evaluation is another advantage.

Third, and last, we recommend that this evaluation be repeated in two years with two additional cohorts of Even Start families and the same or comparable instrumentation. Then direct comparisons with data in this evaluation can be made. Comparisons across four cohorts of families would be very useful. Another evaluation design could include all the families in the Monongalia County program, which would increase the yearly sample size, but also would increase the variability on each measure.

REFERENCES

Abt Associates, Inc. (1994). *Preschool inventory: Administration manual: Revised version.* Cambridge, MA: Abt Associates, Inc.

Arney, W. R. (1990). *Understanding statistics in the social sciences.* New York: W. H. Freeman and Company.

Caldwell, B. M., & Bradley, R. H. (1978). *Home observation for measurement of the environment manual.* Little Rock, AR: University of Arkansas.

CASAS. (1993). *CASAS technical manual.* San Diego, CA: Author.

CASAS. (1997). *Comprehensive adult student assessment system: Assessment, training, resources.* San Diego, CA: Author.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* (Rev. ed.). New York: Academic Press.

Coons, C. E., Gay, E. C., Fandal, A. W., Ker, C., & Frankenburg, W. K. (1981). *The home screening questionnaire reference manual.* Denver, CO: University of Colorado, Health Sciences Center, J.F.K. Child Development Center.

CTB/McGraw-Hill. (1970). *Preschool inventory, revised edition 1970: Handbook.* Monterey, CA: Author.

Dwyer, C. M. (1995). *Guide to quality for Even Start family literacy programs.* Portsmouth, NH: RMC Research Corporation.

Even Start Staff. (1995). *Monongalia County Even Start family literacy project.* Morgantown, WV: Monongalia County Schools.

Even Start Staff. (1996). *Even Start Program: 1995-1996.* Morgantown, WV: Monongalia County Schools.

Joint Committee. (1994). *The program evaluation standards* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Quality Education Data. (1996). *QED state school guide: West Virginia, 1996-1997* (17th ed.) Denver, CO: Author.

St. Pierre, R., Swartz, J., Gamse, B., Murray, S., Deck, D., & Nickel, P. (1995). *National evaluation of the Even Start family literacy program*. Cambridge, MA: Abt Associates, Inc.

St. Pierre, R., Swartz, J., Murray, S., Deck, D., & Nickel, P. (1993). *National evaluation of the Even Start family literacy program: Report on effectiveness*. Cambridge, MA: Abt Associates, Inc.

Tao, F., Swartz, J., St. Pierre, R., & Tarr, H. (1997). *National evaluation of the Even Start family literacy program: 1995 interim report*. Washington, DC: U.S. Department of Education, Planning and Evaluation Services.

U.S. Congress. (1994). *Goals 2000: Educate America Act. Report 103-446*. Washington, DC: U.S. Government Printing Office.

U.S. Congress. (1994). *Improving America's Schools Act. Report 103-761*. Washington, DC: U.S. Government Printing Office.

Walker, D., Bane, M. J., & Bryk, A. (1973). *The quality of Head Start planned variation*. Cambridge, MA: Huron Institute.

Zimmerman, I. L., Steiner, V.G., & Pond, R. E. (1992). *Preschool Language Scale-3: Examiner's manual*. San Antonio, TX: The Psychological Corporation, Harcourt Brace Jovanovich, Inc.

APPENDICES

Appendix A:

Report of the Monongalia County Schools Even Start Program at the End of the 1995-1996 Year

Even Start Program
1995-1996

Families

Total Enrolled: 62 families

62 families (100%) participated in at least two program components. (early childhood education & parenting)

54 families (87%) participated in all three program components. (early childhood, parenting & adult education)

99 adults participated in at least one program component
62 mothers, 37 fathers
122 children ages 0-7
17 children ages 8-18

Number of New Families Enrolled
1995-1996

Continued from 1994-95	37
September	5
October	5
November	3
December	2
January	2
February	4
March	3
April	0
May	1
TOTAL	62

Family Composition:

Couple households: 34 families (55%)
Single households: 28 families (45%)

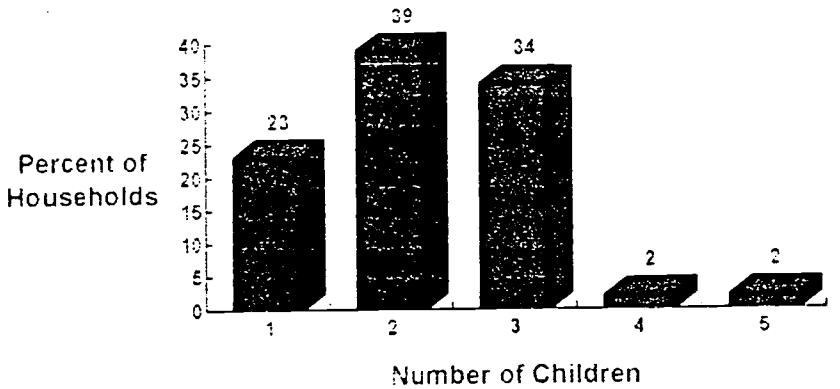
Number of Children:

Mean

All families	2.2
Couple households	2.5
Single households	1.9

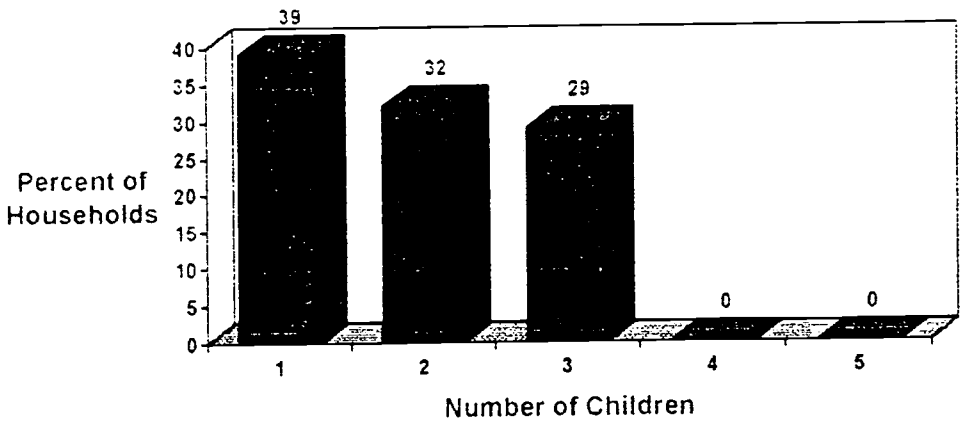
Distribution of Children in Couple Households

34 Families



Distribution of Children in Single Households

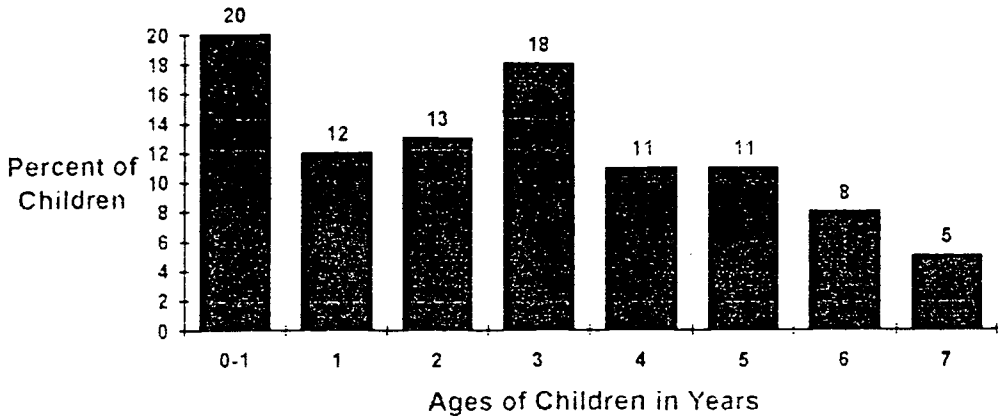
28 Families



Ages of Children:

Distribution of Ages of Children

122 Children (0-7)



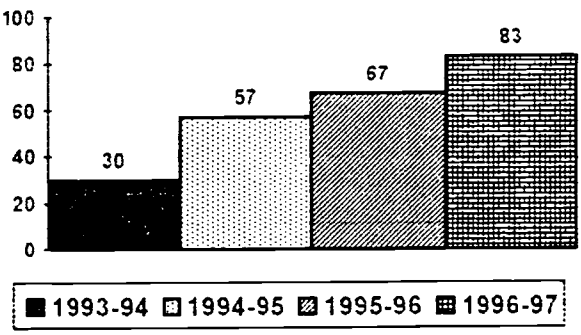
63% of children in the program are ages 0-3 years.

*17 Children 8+ years old also benefited from the program.

Number of Four Year Old Participants

Year	Number
1993-94	10
1994-95	14
1995-96	12
1996-97	17

Percentage of Four Year Old Even Start participants enrolled in Head Start



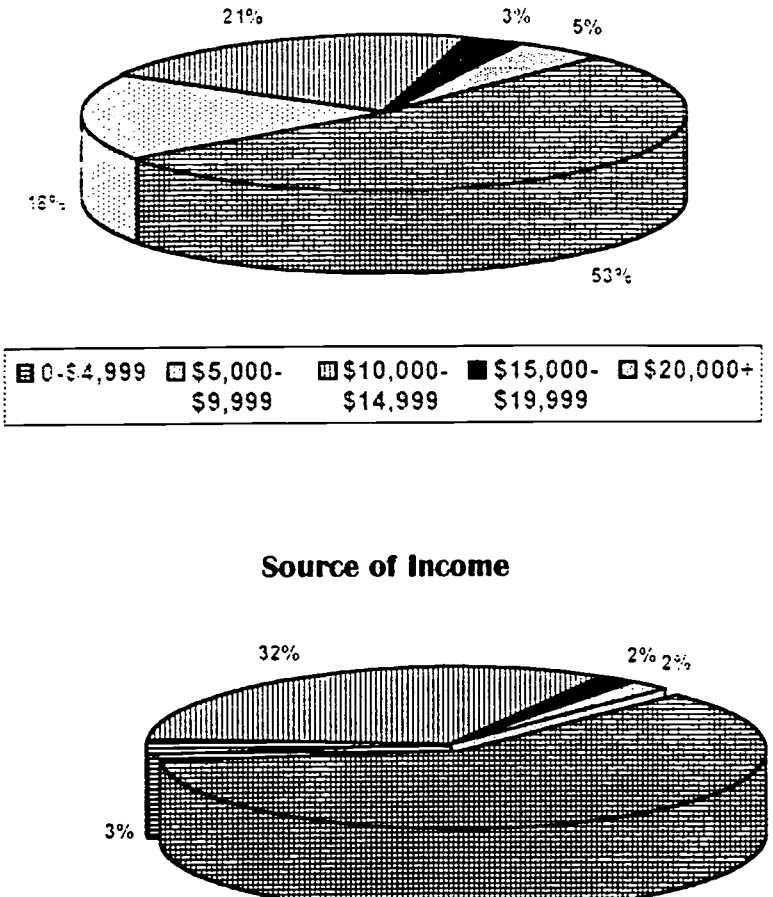
Income

Mean

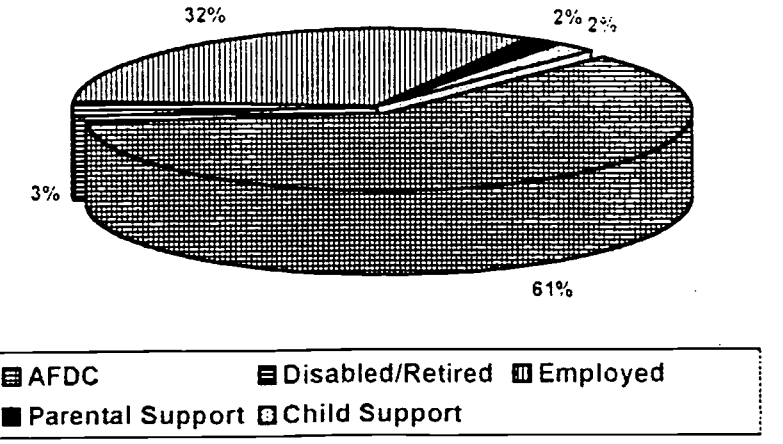
Average: all families	$7,583.50
Couple households	$10,383.00
Single Households	$4,040.00

Income Distribution

Range: $2,206-$28,000



Source of Income



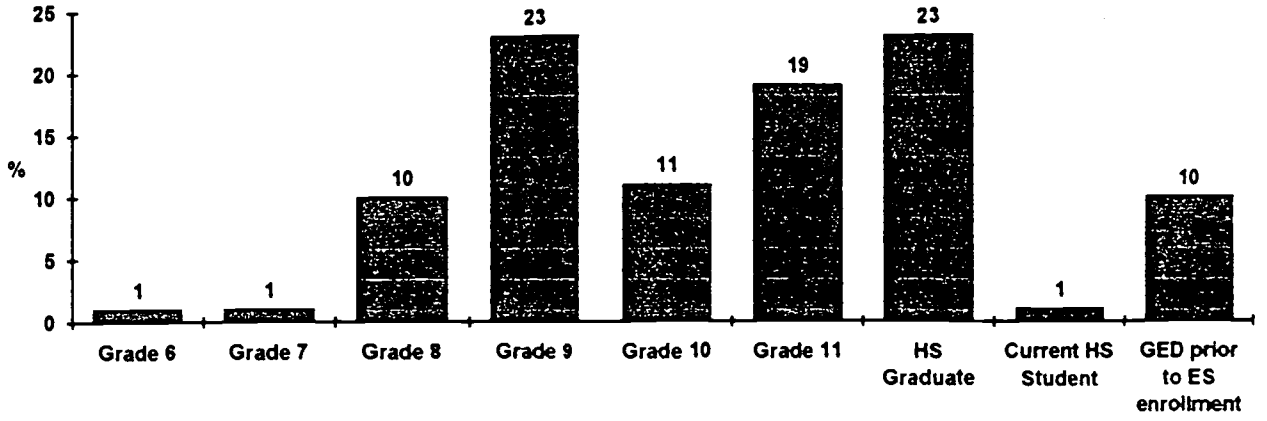
Parent School Years Completed at Enrollment

51% of fathers and 33% of mothers had a GED or High School diploma prior to enrolling in Even Start. Of the remaining mothers, five (12%) received their GED during the 1995-96 program year.

Mothers:

Mothers' School Years Completed

62 Mothers



Mothers' Adult Education Program

Fifty-one Even Start mothers (82%) participated in Adult Education.

# Mothers	Adult Education Program	% Mothers
1	High School Student	1
5	Post HS Continuing Ed.	8
18	Only ES Home based	29
27	Adult Basic Ed. with some ES direct activity	44

Educational Status of Mothers Not in Adult Ed. Program

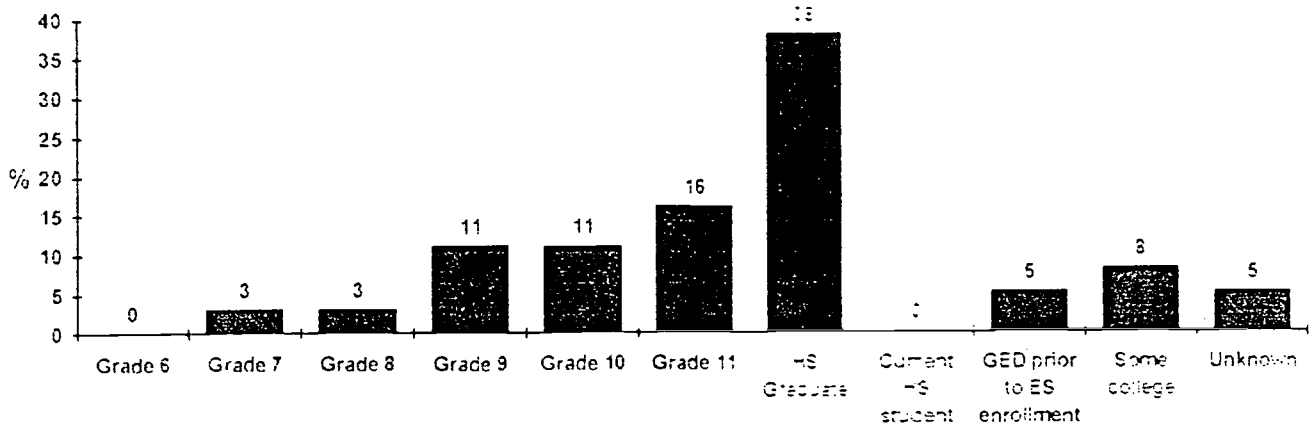
Eleven Even Start mothers (18%) did not participate in Adult Education.

# Mothers	Status	% Mothers
3	HS graduate	27
1	GED prior to ES	9
7	Non-graduate	64

Fathers:

Fathers' School Years Completed at Enrollment

37 Fathers



Fathers' Adult Education Program

Seven Even Start fathers (19%) participated in Adult Education.

# Fathers	Program	% Fathers
5	Adult Basic Ed. with some ES direct activity	71
1	College student	14
1	Literacy volunteer	14

Most fathers participating in Adult Education programs had direct instruction with Even Start staff.

Educational Status of Fathers Not in Adult Education

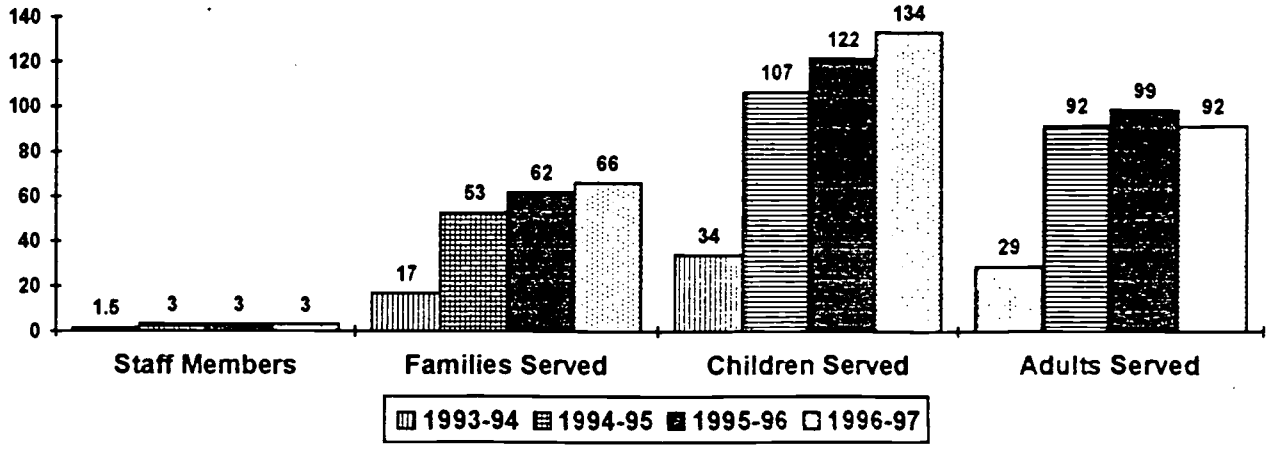
Thirty Even Start fathers (81%) did not participate in Adult Education.

# Fathers	Status	% Fathers
13	HS graduate	43
2	GED prior to ES	7
2	Some college	7
13	Non-graduate	43

Due to rounding some percentages may not equal 100.

Even Start Program Growth

The Monongalia County Even Start program has grown rapidly since its implementation in 1993. Since then the number of staff has doubled and the number of families has nearly quadrupled.



Additional Information for 1995-96

Number of Home Visits	868
Contact Hours	1,128
Children's Books Circulated	3,922

Community Book Donations Through Pinocchio's Book Angel Tree: 149 holiday gifts of books were presented to families

Appendix B:

Even Start Home Visit Report

EVEN START HOME VISIT REPORT

Name:_____ Visit Number:_____

Date:_____ Family Members Present:_____

Visit Theme:_____

Parent's Comments/Concerns:

Handouts/Materials left:

Early Childhood Goals:
*Developmental Focus: _____Language _____Social _____Intellectual _____Fine/Gross Motor

Child Benchmarks/Performance Criteria: (Be specific about what was accomplished & cite examples of learning):

*Materials & Activities:

Adult Goals:
*Focus: _____Parenting _____Adult Life Skills

Adult Benchmarks/Performance Criteria:

Follow-up for family educator:

Follow-up for parent:

Comments:

Appendix C:

Site Visit Schedule for December 1996

Appalachia Educational Laboratory

Planning, Research, and Evaluation Unit

Site Visit Schedule for December 2-6, 1996

to

Monongalia County Even Start Program

Monday

8:30 am - 10:30 am	Orientation with Even Start Staff (n = 4)
10:30 am - 12:00 pm	Interview with Janet Spring
12:00 am - 1:00 pm	Lunch
1:00 pm - 4:00 pm	Work with Even Start Records

Tuesday

8:30 am - 11:00 am	Observe Self-Assessment Session with Even Start Staff
11:00 am - 12:00 pm	Lunch
12:00 am - 1:30 pm	Interview with Angie Swisher
1:30 pm - 2:30 pm	Interview with Susan Hildebrand (sub)
2:30 pm - 4:00 pm	Work with Even Start Records

Wednesday

8:30 am - 10:00 am	Interview with Sandy Walsh
10:00 am - 12:00 pm	Work with Even Start Records
12:00 am - 1:00 pm	Lunch
1:00 pm - 2:15 pm	Work with Even Start records
2:15 pm - 2:45 pm	Travel to Riverside school
2:45 pm - 4:00 pm	Interview with Cindy Griffin

Thursday

8:30 am - 9:30 am	Interview with Bob Snider, Adult Education collaborator
9:30 am - 10:30 am	Interview with Linda Bungard, Adult Education collaborator
10:30 am - 4:00 pm	Work on Even Start Records and Lunch

Friday

7:30 am - 9:30 am	Even Start Breakfast at Erickson Center on WVU Campus - Observe discussion groups
9:30 am - 10:30 am	Informal interviews with various collaborators
10:30 am - 2:00 pm	Records, lunch, and more records
2:00 pm - 3:00 pm	Interview with Ruth Siegel, Head Start teacher

Appendix D:

The Innovation Configuration Matrix

Even Start Program, Monongalia County Schools

Innovation Configuration Matrix

Component #1: Identification and Recruitment

A. Staff consistently employ a wide variety of recruitment practices, including door-to-door recruitment; awareness sessions; community information campaigns; and solicitation and referrals from Head Start and Title I teachers, classroom teachers, other service agencies, and clients in the program. Clear eligibility criteria have been communicated to all parties.

B. Staff employ a variety of recruitment practices, including door-to-door recruitment, and soliciting and receiving referrals from Head Start and Title I teachers, plus any two of the other five practices named in the cell to the left. Clear eligibility criteria have been communicated to all parties.

C. Staff employ several recruitment practices, including at least one form of soliciting of referrals.

D. Staff rely mainly on others to provide referrals of potential client families.

Component #2: Home Visits

A. Staff consistently schedule, plan, and complete biweekly home visits in which literacy is the main focus, and each visit includes an activity in two areas: (1) early childhood education, and (2) parenting. Adult education is also a focus of that visit or of additional contacts.

B. Staff consistently schedule, plan, and complete biweekly home visits in which literacy is a focus, and each visit includes an activity in at least two of the three areas.

C. Staff consistently schedule, plan, and complete monthly home visits in which literacy is a focus, and each visit includes an activity in at least two of the three areas.

D. Staff consistently schedule, plan, and complete monthly or bimonthly home visits that include at least one activity in one of the three areas.

Component #3: Early Childhood Education

A. Staff regularly plan and complete interactive learning activities with parent and child that are based on validated curricula; address the developmental areas of the child (social and emotional, motor, language, and intellectual); are based on the needs of the child; are interesting and enjoyable to the child; and are models for parents to follow.

B. Staff regularly plan and complete interactive learning activities with parent and child that are based on established curricula, address the majority of the developmental areas of the child, are based on the needs of the child, are interesting and enjoyable to the child, and are models for parents to follow.

C. Staff regularly plan and complete learning activities with the child that are based on some curricula, address several of the developmental areas of the child, are based on the needs of the child, and are interesting and enjoyable to the child.

D. Staff complete learning activities with the child that address one area of child development.

KEY: ——— = To left of this line is acceptable practice.

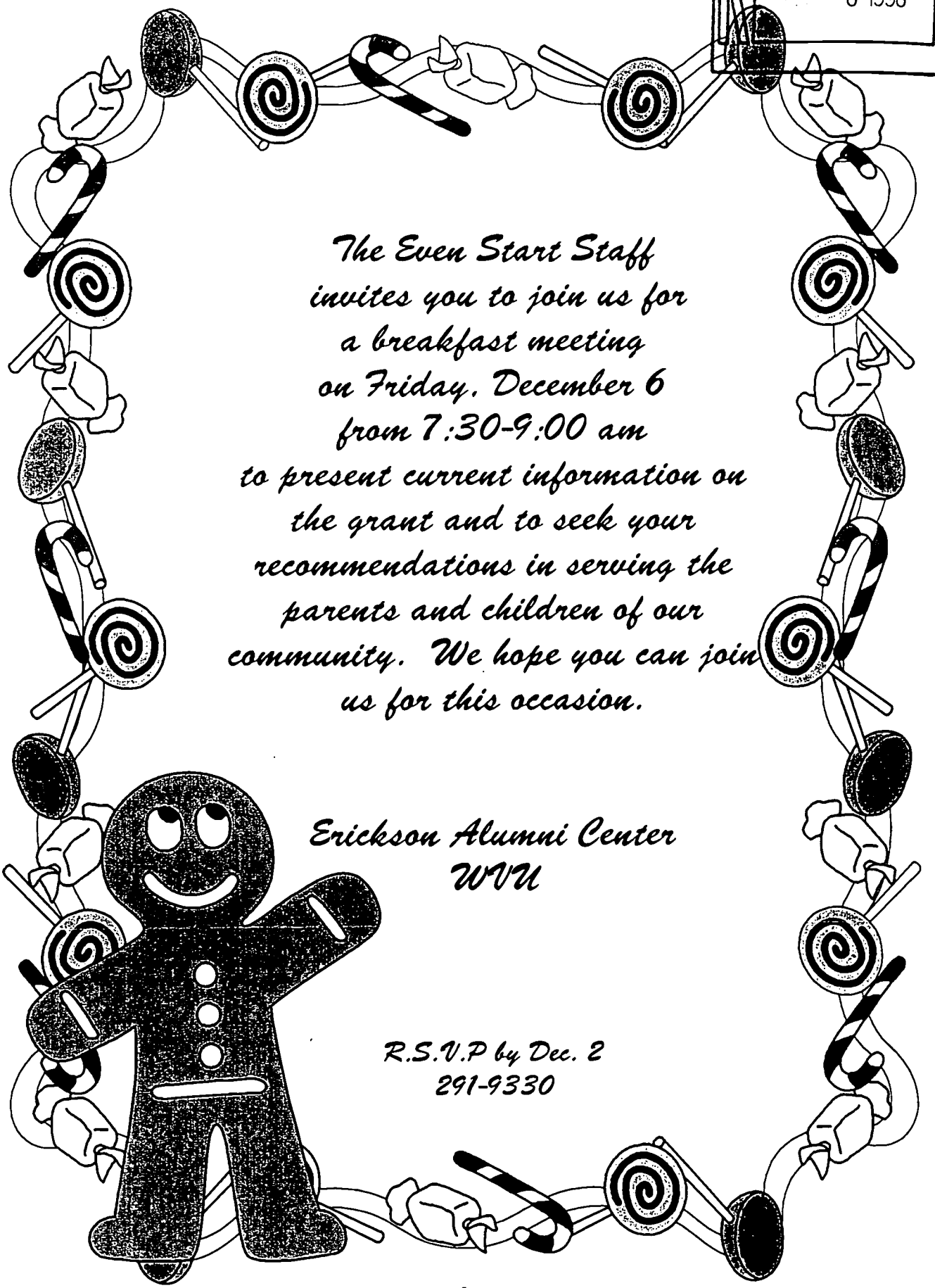
— — — = To left of this line is ideal practice.

Component #4: Parenting Education			
A. Staff regularly plan and complete interactive learning experiences that are based on the needs of the family, are based on validated curricula, utilize parent concerns about child raising, address a wide variety of parenting issues, are sensitive to family situations, are designed to teach the parent about the stages of child development, and are discussed in terms of application.	B. Staff regularly plan and complete interactive learning experiences that are based on the needs of the family, utilize parent concerns about child raising, address a wide variety of parenting issues, and are discussed in terms of application.	C. Staff periodically plan and complete learning experiences that are based on the needs of the family, and address several parenting issues.	D. Staff sometimes complete learning experiences that address a parenting issue.
Component # 5: Adult Education/Literacy			
A. Staff regularly discuss with parents a wide variety of possible outcomes to establish realistic long- and short-term educational goals, and to help eliminate common barriers such as transportation and child care; then help parents adjust into instruction through physical presence, encouragement, and the provision of services, if necessary; and help parents assess their learning progress through periodic monitoring and assessment.	B. Staff regularly discuss with parents a variety of possible outcomes to establish long- and short-term educational goals, and to help eliminate a typical barrier; then help parents adjust into instruction through encouragement and the provision of services, if necessary; and discuss parents' progress toward goals.	C. Staff regularly discuss with parents a variety of possible outcomes to establish short-term educational goals, then help parents adjust into instruction through encouragement, and discuss parents' progress toward goals.	D. Staff discuss with parents possible outcomes; may help set up a short-term goal, and might discuss progress.
Component #6: Collaboration			
A. Staff regularly work with a wide variety of different service agencies regarding multiple needs of client families; share agency/family information, if appropriate; communicate regularly, sometimes in meetings and other face-to-face situations; solicit other agencies' input as to Even Start program delivery; refer client families to proper agencies; develop common purposes and directions of resources; and work with local businesses to achieve goals.	B. Staff regularly work with a variety of service agencies regarding multiple needs of client families; share agency/family information, if appropriate; communicate regularly; refer client families to proper agencies; conduct common meetings; and develop common purposes.	C. Staff work with several service agencies regarding multiple needs of client families; share agency/family information, if appropriate; and refer families to agencies as needed.	D. Staff may work with service agencies for a few client families; share some agency/family information, if appropriate; and refer families to agencies as needed.

Component #7: Evaluation			
A. Staff maintain three levels of evaluation data to assure program effectiveness: (1) staff consistently complete and maintain established family reports that include the intake form, home visit report, book list, contact log, client assessment profile, and annual local and national family summary reports; (2) staff review and explore its processes and effectiveness through local evaluations; and (3) staff periodically contract with a third-party evaluator to conduct formative and summative evaluations resulting in reports that meet the *Program Evaluation Standards.*	**B.** Staff maintain two levels of evaluation data to assure program effectiveness: (1) staff consistently complete and maintain established family reports that include the intake form, home visit report, book list, contact log, client assessment profile, and annual local and national family summary reports; and (2) staff periodically contract an outside evaluator to conduct a local evaluation.	**C.** Staff maintain two levels of evaluation data to assure program effectiveness: (1) staff consistently complete and maintain established family reports that include the intake form, home visit report, contact log, a few client assessments, and annual local and national family summary reports; and (2) staff review and explore its processes and effectiveness through local evaluations.	**D.** Staff complete and maintain one level of evaluation data, the annual national family summary report.
Component #8: Staff Development			
A. Staff have completed extensive training and certification in the program's nationally validated curriculum model that addresses two of the Even Start program components: early childhood education and parenting. Staff have also completed training so that they are recognized as adult educators. Staff have provided input into the design and selection of ongoing and varied staff development sessions. Sessions are both program and individual-based, and participation and involvement is shared/discussed with other program staff. Cross-training and learning with collaborators occurs frequently.	**B.** Staff have completed extensive training and certification in the program's nationally validated curriculum model that addresses two of the Even Start program components: early childhood education and parenting. Staff have also completed training so that they are recognized as adult educators. Staff development sessions are ongoing and varied, and are program-based; staff development information is shared/discussed with other program staff.	**C.** Staff have completed extensive training and certification in the program's nationally validated curriculum model that addresses two of the Even Start program components: early childhood education and parenting. Staff have also completed training so that they are recognized as adult educators.	**D.** Staff attend some training related to the Even Start program; the training may be varied, and it may be shared with others.

Appendix E:

1996-97 Annual Collaborators Breakfast
Meeting Materials



The Even Start Staff
invites you to join us for
a breakfast meeting
on Friday, December 6
from 7:30-9:00 am
to present current information on
the grant and to seek your
recommendations in serving the
parents and children of our
community. We hope you can join
us for this occasion.

Erickson Alumni Center
WVU

R.S.V.P by Dec. 2
291-9330

Even Start Collaborative Program
Fall, 1996 Meeting

Even Start staff
Susan Hilderbrand
Harry Lewis
Angie Swisher
Janet Spring

Title I
Marie Alsop
Sandy Walsh

Head Start
Cheryl Wienke
Ruth Siegel
Marcy Shadle
BeckyDavies
Meredith Kiger
Cindy O'Brien
Joanne Martin
Suzanne Smart

Special Education
Judy Kelly Minor
Jane Cutwright

Adult Learning Center
Bob Snider
Linda Bungard
Edie Jett

Board of Ed
Dr Ed Warnick
Jenifer Snider
Debbie Smith

Central School
Frank Mrazeck
Anita Nedeff
Mary Lou Hutchinson

Cass School
Ken Fetty
Carol Trythall
Betty Hamilton
Carla Mason

Cool Springs School
Helen Davies
Debbie Savage

Valley Mental Health
Kay Nottingham Chaplin

Mon Cnty Home Health
Beth Thewlis, RN

Health & Human Resources
Joan Kramer

Shack Neighborhood House Happy School
Jane Cooley

Rock Forge Neighbhd House Happy School
Debbie Hagedorn

Joy Saab
Clay Pytlik
Jean Hagan, Pinnocchio's
Merle Meehan

Sabraton Reading Class
Cathy White

Mountainview School
Steve King
Patty Benedum

Summers School
Alan Keener
Lynn Sobolov
Ginny Gross

Riverside School
Bob Glock
Christine Wilson
Hannah Lewis

Appendix F:

Program Staff's 1996-97 Self-Assessment
Analysis Report

1996-97 Self Assessment Analysis

Integration of Components

The program components of early childhood education, parenting and adult education are well interconnected in our home based model where each staff person delivers all components to each family.

Self assessment has identified a need to:

Establish regular staff meetings to address integration of components in home visits.
(Increased staff and decreased case loads for next year should facilitate this.)

Offer more parent group meeting and socialization times.
(A social work intern will focus much of her work on this next year.)

Collaboration

Multiple relationships with a variety of agencies has yielded good partnerships which benefit families.

Self Assessment has identified a need to:

Seek suggestions from collaborators
(Accomplished at 12/6/97 Collaborators Breakfast. This is an annual event.)

Improve collaboration with Department of Special Education and Department of Human Services by becoming more familiar with their procedures.
(Accomplished: 3/17/97 Meeting with Special Education Coordinator for inservice on preschool handicapped services. Some staff have attended welfare reform information and inservices.)

Create communication mechanisms:
Develop and disperse local program information
Update agency list with current contact persons (completed 12/97)
Make a personal visit to each contact person

Increase contacts with Title I Schools
(Accomplished in 1996-97 through attendance and participation in Partnerships With Parents and Parenting Meetings at Cool Springs and Summers, presentation at Central 1/20/97, and informal and personal collaboration with Cass and Riverside Schools. This will be a continual process.)

Recruitment

Multiple methods (parent, agency, and school referral as well as door to door) are used by all staff. Recruitment is an ongoing and friendly process with a clear message: Through its three components, Even Start is a program for the whole family.

Self Assessment identified a need to:

Establish selection criteria to eliminate subjectivity in recruitment and to insure that the neediest families are being served. Developing a weighted matrix is a possibility.

Parent Child Interactions and Parenting

Relationships between parents and children are enhanced through:

- **Parents As Teachers (PAT) and related programs**
- **High rate of home based one to one interchanges dealing with parenting issues**
- **Modeling parent/child activities with materials to extend the activity**
- **Identification of collaborators who support or provide parenting activities**

Self Assessment has identified a need to:

Encourage socializations for families.
This will be partially addressed by a social work intern who will work with the program in 1997-98.

Home Visiting

Home visits are tailored to individual family needs with literacy as a primary focus. Staff have had great success in establishing rapport in the home setting and providing a bridge to other resources for families.

Self Assessment has identified a need to:

Increase a parent's active role in shaping the home visit.
The parents often find it difficult to formulate and verbalize their concerns. This may be addressed by the use of a new parenting measure in 1997-98.

Adult Education

Adult Education is one program component that has been difficult to define and implement with our targeted population. The number of hours expended in teaching life skills and parenting skills has been excellent; however, the number of hours of adult basic education instruction is lower than desired.

Self assessment has identified a need to:

Extend time given to each adult participant.
(The plan for next year is to increase the number of staff, decrease the case loads, and increase the number of visits.)

Purchase more low level adult literacy materials. (In progress)

Revise the intake form to better establish family goals.

Develop a systematic testing protocol to better measure progress. (In progress)

Early Childhood

Most children experience sufficient ECE hours because of enrollment in outside programs. It is difficult to provide 60 hours per month in a home-based program.

Self-Assessment has identified a need to:

Increase the number of home visits with additional staff and lower caseloads.

Assign a part-time social work student to schedule additional group meetings.

Develop and distribute parenting materials that will provide literacy learning experiences for 6 and 7 year olds.
(We are working with a consultant to assist in writing these materials.)

Retention

Program has high retention of families.

Self-Assessment has identified the need to:

Develop a liaison with local businesses, agencies, and schools to actively pursue volunteer and/or fee-for-service opportunities.

Ask recent successful graduates to work with or mentor currently enrolled families based on interest.

Transitions

Staff spend a great deal of time encouraging transitions so that gains made in Even Start can be continued and augmented.

Self-Assessment has identified a need to:

Encourage the transition of our GED graduates to programs that provide additional employment training.
(Local Adult Learning Center is beginning a Work SCANS program in Fall 1997.)

Staff Development

Staff regularly attend staff development on varied topics provided by educational organizations, social service agencies, professional groups and other collaborators. Since May of 1993, staff have attended over 1300 hours of staff development.

Self assessment has identified a need to:

Attend additional staff development sessions in the area of adult education.
(Staff are participating in the planning of and will attend an adult ed conference in the fall of 1997 on family literacy.)

Find more staff time to share staff development experiences.
(More regularly scheduled staff meetings are planned for next year.)

Appendix G:

Completed *Evaluation Standards* Checklist

Citation Form

The Program Evaluation Standards (1994, Sage) guided the development of this (check one):

_____ request for evaluation plan/design/proposal
_____ evaluation plan/design/proposal
_____ evaluation contract
XXX evaluation report
_____ other: _____

To interpret the information provided on this form, the reader needs to refer to the full text of the standards as they appear in Joint Committee on Standards for Educational Evaluation, *The Program Evaluation Standards* (1994), Thousand Oaks, CA, Sage.

The *Standards* were consulted and used as indicated in the table below (check as appropriate):

Descriptor		The Standard was deemed applicable and to the extent feasible was taken into account.	The Standard was deemed applicable but could not be taken into account.	The Standard was not deemed applicable.	Exception was taken to the Standard.
U1	Stakeholder Identification	XXXX			
U2	Evaluator Credibility	XXXX			
U3	Information Scope and Selection	XXXX			
U4	Values Identification	XXXX			
U5	Report Clarity	XXXX			
U6	Report Timeliness and Dissemination	XXXX (Dissemination addressed in Recommendations)			
U7	Evaluation Impact	XXXX			
F1	Practical Procedures	XXXX			
F2	Political Viability			XXXX	
F3	Cost Effectiveness	XXXX			
P1	Service Orientation	XXXX			
P2	Formal Agreements	XXXX			
P3	Rights of Human Subjects	XXXX (Local staff did so for instruments they			
P4	Human Interactions	XXXX			administered)
P5	Complete and Fair Assessment	XXXX			
P6	Disclosure of Findings	XXXX			
P7	Conflict of Interest	XXXX			
P8	Fiscal Responsibility	XXXX			
A1	Program Documentation	XXXX			
A2	Context Analysis	XXXX			
A3	Described Purposes and Procedures	XXXX			
A4	Defensible Information Sources	XXXX			
A5	Valid Information	XXXX			
A6	Reliable Information	XXXX			
A7	Systematic Information	XXXX			
A8	Analysis of Quantitative Information	XXXX			
A9	Analysis of Qualitative Information	XXXX			
A10	Justified Conclusions	XXXX			
A11	Impartial Reporting	XXXX			
A12	Metaevaluation			XXXX	

Name Merrill L. Meehan (typed) Date: October 30, 1997

(signature)

Position or Title: Senior Research and Evaluation Specialist

Agency: Appalachia Educational Laboratory

Address: P.O. Box 1348, Charleston, WV 25325

Relation to Document: Author

(e.g., author of document, evaluation team leader, external auditor, internal auditor)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

NOTICE

REPRODUCTION BASIS

☑ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").