



Abt Associates Inc.

Cambridge, MA
Lexington, MA
Hadley, MA
Bethesda, MD
Washington, DC
Chicago, IL
Cairo, Egypt
Johannesburg, South Africa

**THIRD
NATIONAL
EVEN START
EVALUATION:
PROGRAM
IMPACTS AND
IMPLICATIONS
FOR
IMPROVEMENT**

Executive Summary

Date: May, 2003

Prepared for
U.S. Department of
Education
Planning & Evaluation
Service
Contract EA97049001

Prepared by
Robert St.Pierre
Anne Ricciuti
Fumiyo Tao
Cindy Creps
Janet Swartz
Wang Lee
Amanda Parsad
Abt Associates Inc.
Cambridge, Mass. and
Bethesda, Md.

Tracy Rimdzius
U.S. Department of
Education

EXECUTIVE SUMMARY

The Even Start Family Literacy Program, established in 1989, aims to simultaneously improve the literacy of children and their parents through (1) early childhood education, (2) parenting education, (3) adult education, and (4) parent-child joint literacy activities. The program's underlying premise is that eligible families need each of these four core instructional components, and that these services will be more effective when integrated in a unified program. During the period of this study, Even Start's guiding legislation stressed process factors such as collaboration with local service agencies and the recruitment and screening of eligible families, although it did require high-quality, intensive instructional components. The legislation was reauthorized in 2000 and 2001, and while all previous requirements have been retained, the legislation now stresses more strongly the importance of the quality of instructional content.

Key Findings in Brief

- ❑ While Even Start children and parents made gains on literacy assessments and other measures, children and parents in the 18 Even Start programs that participated in the EDS did not gain more than children and parents in the control group, about one-third of whom also received early childhood education or adult education services.
- ❑ Even Start serves a very disadvantaged population. Compared with Head Start, Even Start parents are much less likely to have a high school diploma, and Even Start families have substantially lower annual household income.
- ❑ Even Start children and parents made small gains on literacy measures and scored low compared to national norms when they left the program. Even Start children gained four standard score points on the Peabody Picture Vocabulary Test, the same amount gained by control group children and by children in the Head Start FACES study.
- ❑ Families do not take full advantage of the services offered by Even Start projects, participating in a small amount of instruction relative to their needs and program goals.
- ❑ While the early childhood classroom experiences provided by the EDS projects were of overall good quality, there was not sufficient emphasis on language acquisition and reasoning to produce measurable impacts and hence to achieve legislative outcomes. Further study is needed to better document the quality and content of Even Start's instructional services.
- ❑ The extent to which parents and children participated in literacy services is related to child outcomes.

This document reports findings from the third national Even Start evaluation. The Department of Education selected Abt Associates Inc. and Fu Associates, Ltd. to measure the effectiveness of the program and to provide information on program implementation. The evaluation included two complementary studies: (1) the Even Start Performance Information Reporting System (ESPIRS) which provided annual data on the universe of Even Start projects, and (2) the Experimental Design Study (EDS) which was an experimental study of Even Start's effectiveness in 18 projects.

The ESPIRS portion of the evaluation requested data from every Even Start project in each of four years (1997-1998 through 2000-2001) including program and family characteristics, participation rates, and family progress indicators. The EDS portion of the evaluation was conducted by collecting pretest, posttest, and follow-up data from families in 18 projects (one home-based project and 17 center-based or home/center-based projects) that were willing to randomly assign incoming families to participate in Even Start or to be in a control group.

This report draws on data collected in all four years of the ESPIRS as well as pretest and posttest data from 1999-2000 and 2000-2001 collected from the 18 EDS projects (see Exhibit E.1, below). Follow-up data from the EDS were not available in time to be included in this report. Hence, this document presents descriptive information on all Even Start programs and participants, and in addition discusses program impacts based on pretest and posttest data collected from the 18 EDS projects. Where possible, we have used data from studies of other programs with aims similar to Even Start (e.g. Head Start) in order to provide a context for the Even Start findings.

EXHIBIT E.1					
DATA COLLECTION SCHEDULE FOR THIRD NATIONAL EVEN START EVALUATION					
DATA BASE FOR THE EVALUATION	YEAR OF DATA COLLECTION				
	1997-1998	1998-1999	1999-2000	2000-2001	2001-2002
ESPIRS (all projects)	✓	✓	✓	✓	
EDS Cohort 1 (11 projects)			✓Fall 99 pretest ✓Spring 00 posttest	✓Spring 01 follow up (not included in this report)	
EDS Cohort 2 (7 projects)				✓Fall 00 pretest ✓Spring 01 posttest	✓Spring 02 follow up (not included in this report)

THE EVEN START FAMILY LITERACY PROGRAM

Even Start addresses the basic educational needs of low-income families including parents and their children from birth through age seven by providing a unified program of family literacy services, defined as services that are of sufficient intensity in terms of hours, and of sufficient duration, to make sustainable changes in a family, and that integrate:

- ❑ Interactive literacy activities between parents and their children (parent-child activities).
- ❑ Training for parents regarding how to be the primary teacher for their children and full partners in the education of their children (parenting education).
- ❑ Parent literacy training that leads to economic self sufficiency (adult education).
- ❑ An age-appropriate education to prepare children for success in school and life experiences (early childhood education).

Even Start's long-term purpose is to help break the cycle of poverty and illiteracy for low-income families. Local Even Start projects are meant to integrate the components of family literacy and build on services that already exist in their communities. The program has grown steadily over the past decade, both in terms of federal funding as well as the number of projects that are supported with those funds. From a small demonstration program in which \$14.8 million was used to fund 76 projects in 1989-1990, Even Start has grown ten-fold. In 2000-2001, \$150 million in funding was distributed to 855 projects serving 32,000 families in all 50 states (Exhibit 1.2)¹, and funding rose to \$250 million in 2001-2002. Even Start has been reauthorized and amended several times, most recently through the Literacy Involves Families Together (LIFT) Act of 2000 and the No Child Left Behind Act of 2001. The third national evaluation was designed before these reauthorizations, so this report's findings reflect Even Start as it was implemented pre-reauthorization.

KEY FINDINGS

While Even Start children and parents made gains on literacy assessments and other measures, children and parents in the 18 Even Start programs that participated in the EDS did not gain more than children and parents in the control group, about one-third of whom received early childhood education or adult education services.

On a wide variety of measures, Even Start children and their parents performed as well as, but not better than, control group children and their parents. The data show that children in the control group made the same kinds of gains as Even Start children on early literacy, language development, math skills, and social skills. Parents in the control group made the same kinds of gains as Even Start parents on assessments of adult literacy. And finally, families in the control

¹ References are to exhibits in the body of the report.

group made the same kinds of changes as Even Start families on economic self-sufficiency, parent-child reading, and literacy resources in the home (page 147).

Even Start serves a very disadvantaged population. Compared with Head Start, Even Start parents are much less likely to have a high school diploma, and Even Start families have substantially lower annual household income.

Even Start projects are required to identify, recruit, and serve the neediest families in their communities. This evaluation shows that projects take their mandate seriously, as Even Start families are poor, undereducated and underemployed by any standards. In 2000-2001, almost half of the parents who joined Even Start had less than a 9th grade education and 85 percent lacked a high school diploma or GED (Exhibit 4.10). In 1997, only 28 percent of Head Start parents lacked a high school diploma. During 2000-2001, 39 percent of new Even Start families reported annual household income of less than \$9,000 and 84 percent lived below the federal poverty line² (Exhibit 4.5). In 1997, 41 percent of Even Start families and 13 percent of Head Start families reported annual household income under \$6,000 (Exhibits 4.5 and 4.6).

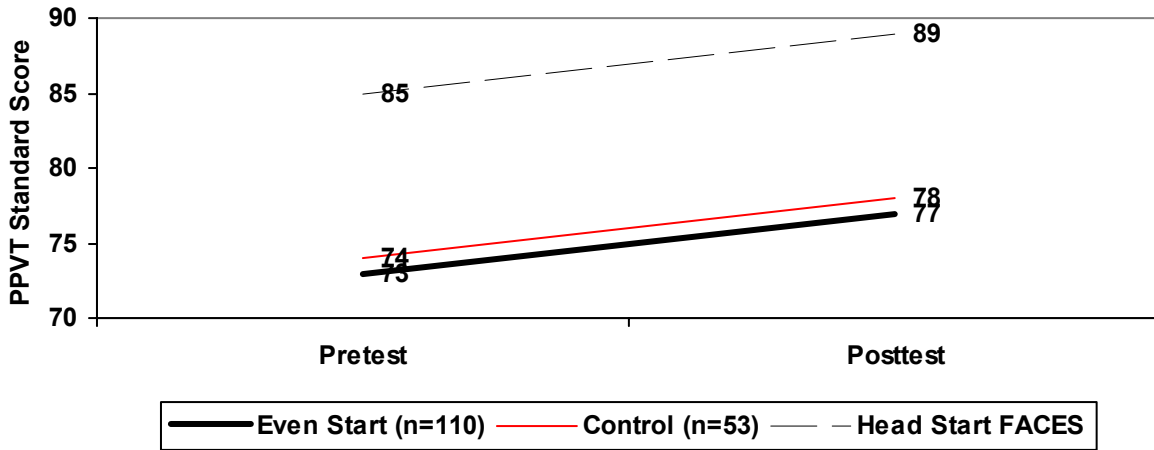
Even Start children and parents made small gains on literacy measures and scored low compared to national norms when they left the program. Even Start children gained four standard score points on the Peabody Picture Vocabulary Test, the same amount gained by control group children and by children in the Head Start FACES study.

Even Start and control group children each gained about four standard score points on the Peabody Picture Vocabulary Test (PPVT), a measure of receptive vocabulary, an amount that is comparable to the gain made by children in the Head Start FACES study (Exhibit E.2). Even Start children have literacy levels far below their counterparts in Head Start, and Even Start children and parents scored at the bottom when compared to national standards.

- Even Start children: When pretested with the PPVT, Even Start children scored at the 4th percentile, almost two full standard deviations below the national norm and one full standard deviation below Head Start children. The same children scored only at the 6th percentile on this measure at the posttest (Appendix 6.1, Exhibit 6.1.1). Even Start children score similarly low on several Woodcock-Johnson subtests (Letter-Word Identification, Dictation, Applied Problems, Incomplete Words, Sound Blending).
- Even Start parents: On the EDS pretest, Even Start parents scored at the 1st percentile (grade 2.9) on reading comprehension skills and at the 5th percentile (grade 4.1) on basic reading skills as measured by the Woodcock-Johnson. Even after making gains while in the program, Even Start parents moved only to the 2nd percentile on reading comprehension and to the 8th percentile on basic reading skills (Appendix 6.1, Exhibits 6.1.24 and 6.1.25).

² The federal definition of poverty considers both family income and household size. In 2000, a family of four (two adults, two children) was considered to be below the federal poverty line if it had annual income below \$17,463.

Exhibit E.2: Pretest and Posttest Standard Scores on the PPVT for Even Start and Control Children in the EDS, and for Children in the Head Start FACES Study

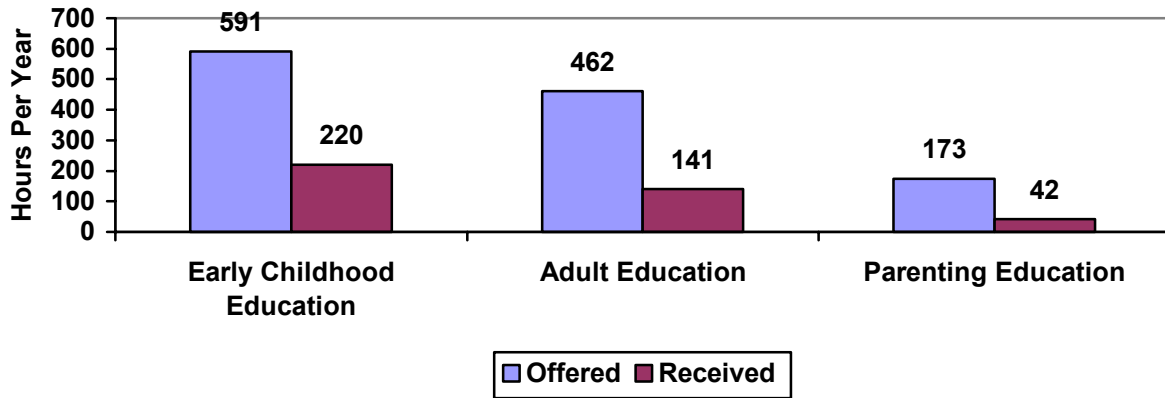


Families do not take full advantage of the services offered by Even Start projects, participating in a small amount of instruction relative to their needs and program goals.

The Even Start legislation requires projects to serve families that are “most in need” of educational services, and puts several requirements into place in recognition of the challenge of serving such a disadvantaged population. For example, Even Start is unique among federal programs in its ability to serve families with children from birth through age seven. While many other programs serving young children are meant to last only nine months to a year, Even Start has the potential to help children progress from infancy through the second grade. Further, projects are required to serve at least a three-year range within the birth through seven age span. Finally, the definition of family literacy services included in the legislation points out that services need to be of sufficient intensity and duration to produce meaningful change in families.

In response to these requirements, as well as to research showing that children who participate intensively in high-quality interventions are the ones who benefit the most (Ramey & Ramey, 1992), the Department of Education has provided technical assistance and encouraged projects to offer multi-year instructional services at high levels of intensity, and to improve retention in the program. This evaluation has documented increases over time in the amount of early childhood education and adult education offered to Even Start families. In spite of the increased amount of instructional services available, the average Even Start family received a low level of intervention services, both in terms of duration in months and total hours of participation, relative to their needs, relative to the goals of the program, and relative to the amount of instruction received by children in other programs that have generated large effects on child development. Exhibit E.3 compares the average annual hours offered to and received by Even Start children and parents who participated in early childhood education (birth through age five), adult education, and parenting education.

Exhibit E.3: Average Annual Hours of Early Childhood Education, Adult Education, and Parenting Education: Offered by Even Start Projects vs. Received by Even Start Families



- ❑ The average Even Start family remained in the program for 10 months and received instructional services in seven of those months (Exhibit 5.12).
- ❑ Approximately one-third of all families that joined Even Start during the four years of this evaluation participated for more than 12 months; conversely, two-thirds left the program with fewer than 12 months of participation. Of the Even Start families in the 18 projects that participated in the EDS, 35 percent did not participate enough to be included in the ESPIRS data collection. The remaining families were enrolled for an average of eight months, slightly less time than the national average (Exhibit 6.11).
- ❑ Each national Even Start evaluation has shown that families participate more intensively when they are in projects that offer higher amounts of instructional services. Over time, Even Start projects have increased the amount of early childhood education and adult education offered to children and parents (the amount of parenting education offered has not increased). However, in 2000-2001, parents and children actually participated in only a small fraction of the hours offered: 30 percent of adult education, 24 percent of parenting education, 25 percent of parent-child activities, and 30 to 62 percent of early childhood education (depending on the age of the child) (pages 127-130).
- ❑ In 2000-2001, parents received instructional services in an average of seven months. During that time they received an average of 42 hours of instruction in parenting education and 38 hours in parent-child activities, roughly equivalent to 1.5 hours per week of each. Parents received an average of 141 hours of adult education instruction, about five hours per week (Exhibit 5.5), and more than double the amount of participation in adult education programs nationally.
- ❑ In 2000-2001, children received instructional services in an average of seven months. Children birth to two received an average of 159 hours of early childhood education instruction (about six hours per week), children age three and four received an average of

254 hours (about eight hours per week), and children age five received an average of 246 hours (about seven hours per week) (Exhibit 5.9).

While the early childhood classroom experiences provided by the EDS projects were of overall good quality, there was not sufficient emphasis on language acquisition and reasoning to produce impacts that are greater than the control group and other early childhood programs and hence to achieve legislative outcomes. Further study is needed to better document the quality and content of Even Start’s instructional services.

Prior research has shown that high-quality early childhood programs can have large (although generally short-term) effects on the cognitive development of children from low-income families. So, if Even Start hopes to have large effects on the literacy and development of participating children, it is important to implement early childhood services of the highest possible quality with the best possible content, as identified by recent, scientific research. While this evaluation does not provide an in-depth assessment of the quality or content of Even Start’s instructional services, the early childhood services implemented by the EDS projects were comparable in overall quality to, but not appreciably better than, the early childhood services received by Head Start children and by children in other preschool programs.

The *Early Childhood Environment Rating Scale (ECERS)* was used to assess the overall quality of early childhood services in Even Start classrooms that participated in the EDS (Exhibits 3.7 and 3.8). The EDS early childhood classrooms were generally comparable in quality to Head Start classrooms, and were rated somewhat higher than some other types of early childhood classrooms. Even Start staff in most classrooms did a good job of supervising and encouraging children, using non-punitive discipline methods, and responding to children in a supportive and respectful manner. These characteristics help build positive relationships with children and guide them in adjusting to the social and behavioral rules of school. However, Even Start staff rarely expanded on information or ideas presented by children, there was often not a good balance between staff listening and talking to children, and staff in many classrooms did not talk with children about logical relationships. Thus, language was not frequently used to encourage children’s reasoning and communication skills.

The *Literacy Checklist*, a measure of reading and writing resources, was also used in the EDS (Exhibit 3.9). Most Even Start classrooms in the EDS had books displayed and available for children to use, and all had a library or reading corner. Most classrooms had an area set up for writing. However, Even Start classrooms scored somewhat lower than Head Start classrooms on the Literacy Checklist, meaning that they had fewer books available to children and were less likely to have writing areas and tools for writing or displays of children’s written work.

Teachers reported that almost all Even Start children in center-based classrooms had many different kinds of literacy-related activities available to them on a daily or almost daily basis including number concepts or counting (95 percent), letters of the alphabet or words (94 percent), and reading stories (90 percent). However, roughly the same percentage of control children also had these literacy activities available to them. Data from the Head Start FACES study show that Head Start children are exposed to the same literacy activities, with one

exception -- Even Start children are more likely than Head Start children to work on letters of the alphabet and words (94 vs. 69 percent).

The “good” overall score on the ECERS for early childhood classrooms in the EDS projects masks the fact that many EDS classrooms did not place sufficient emphasis on language acquisition and reasoning. While all aspects of early childhood classes are important, recent research has pointed out the particular relevance of language and reasoning skills as precursors and tools, both for reading and for general problem solving, especially for children from low-income families who often come from less enriched home learning environments than their middle- and upper-class age-mates. Thus, children are not getting what research says they need if they are to achieve the outcomes envisioned by the Even Start legislation.

The lack of emphasis on language and reasoning in the EDS early childhood classrooms is consistent with the findings of many other research studies which have shown that “...it is precisely on measures of the language environment that preschool programs serving poor children scored in the inadequate range” (Snow, Burns & Griffin, 1998, p. 147). Bredekamp & Rosegrant (1995) refer to this as the “early childhood error” – appropriate attention to traditional quality criteria such as space, materials, and child-caregiver ratio, but inadequate attention to the content of the instruction.

If we expect children to learn more in Even Start than they would otherwise learn (by participating in a control group), then the overall quality of Even Start early childhood education, and especially the emphasis on language acquisition and reasoning, will have to be better than, not the same as, the instruction offered by competing programs. Currently, this does not appear to be the case.

The extent to which parents and children participated in literacy services is related to child outcomes.

Consistent with the findings of prior research (e.g., Barnett, 1995; Ramey & Ramey, 1992; Ramey, Bryant, Wasik, Sparling, Fendt & LaVange, 1992) and with findings from the first national Even Start evaluation (St.Pierre, Swartz, Gamse, Murray, Deck & Nickel, 1995, pp. 175-180), data from the EDS show that children who participated more intensively in early childhood education scored higher on standardized literacy measures. Further, parents who participated more intensively in parenting education had children who scored higher on standardized literacy measures. On the other hand, there is no relationship between the amount of time that parents participated in adult education or parenting education and their scores on literacy outcomes.

Amount of participation was not manipulated experimentally. Instead, the extent to which families participate in Even Start is likely related to various family characteristics such as ethnicity and employment status, as well as program characteristics such as amount of service offered and the extent to which families are encouraged to participate. Therefore, the observed relationships between amount of participation and child literacy may be explained by factors such as differences in the motivation of families or their opportunity to participate in Even Start.

EDS STUDY CHARACTERISTICS

The evidence presented in Chapter 6 shows that (1) random assignment was carried out properly and resulted in statistically equivalent Even Start and control groups, (2) attrition of families from the study between pretest and posttest was low for studies of this type, and (3) outcome assessment was focused on the appropriate domains and used established measures.

Selection of EDS projects. The EDS called for pretest, posttest, and follow-up data to be collected from families in 18 projects. These projects were chosen because they minimally met Even Start’s legislative requirements³, had been in operation for at least two years, planned to operate through the length of the study, could serve at least 20 new families at the start of data collection, offered instructional services of moderate or high intensity relative to all Even Start projects⁴, and were willing to participate in a random assignment study. Projects were recruited from urban and rural areas, as well as projects that served varying proportions of ESL participants. Over the two recruitment years, 115 out of the universe of about 750 programs met the selection criteria, and 18 of these projects were willing to participate in the study.

Each of the 18 EDS projects was asked to recruit families as they normally do and to provide listings of eligible families to Abt Associates staff who randomly assigned families either to participate in Even Start (two-thirds of the families) or to be in a control group (one-third of the families). Assignment to the control group meant that the family could not participate in Even Start for one year. A total of 463 families were randomly assigned in the EDS -- 309 to Even Start and 154 to the control group, maintaining the planned 2:1 ratio.

Sample size and statistical power. The first Even Start random assignment study, conducted in the early 1990s by St.Pierre, Swartz, Gamse, Murray, Deck & Nickel (1995) was criticized because it had a small sample size (five projects with a total of fewer than 200 families) and resulting low power to detect effects. Compared with that study, the EDS has more than three times the number of projects and more than twice as many families. This provides sufficient statistical power to detect medium and large-sized effects, but relatively poor power to detect smaller effects. We argue that while such small effects may be interesting to researchers, they are not always relevant for policy making purposes. Hence, the statistical power offered by this evaluation is appropriate for determining the effectiveness of and improving Even Start.

Generalizability of the findings. Compared with the Even Start population, the 18 EDS projects over-represent Even Start programs that serve ESL Hispanic families in urban areas. While such over-representation means that care should be taken in applying the findings to Even

³ Visits to each EDS project confirmed that they were fully functioning. However, the fact that the EDS projects met Even Start’s legislative requirements and were fully functioning says little about the quality of the instructional services offered to children and their parents. While the EDS sites represent functioning Even Start projects, they were not selected to be models of excellence.

⁴ For this evaluation, projects were defined as “high-intensity” if they offered 60 or more hours per month of early childhood education, 60 or more hours per month of adult education, and 20 or more hours per month of parenting education. Details are presented in Exhibits 2.20, 2.22 and 2.25.

Start projects as a whole, almost 50 percent of the families served by Even Start are Hispanic and about 50 percent of the projects are in urban areas. Hence, the EDS findings apply to an important and growing part of the Even Start population.

Measurement. The EDS measured child and parent outcomes. Child outcomes were measured by direct assessment of the child's literacy skills (for children two years, six months through eight years of age), parent report on the child's skills, teacher report on the child's accomplishments and behaviors in school, and a review of school records. The child measures overlapped with the ESPIRS that was completed by all Even Start families and with measures for the Head Start FACES study. Outcomes for parents were measured through direct assessment of literacy skills and parent self-report. While we generally selected measures that were available in English and Spanish, we were advised that the goal of federally-funded adult education is to increase participant's skills in English, and therefore that we should attempt to assess all adult participants in English. We extended this recommendation to children and instructed data collection staff to attempt to assess all adults and children in English. If this was distressing to a parent or child, the Spanish version of the measure was administered.

Services received by the control group. In studies of educational and social services programs, control groups rarely, if ever, receive "no services." This was the case in the EDS, where control group parents reported that they and their children received various educational services between pretest and posttest. The premise behind Even Start is that it is important for a family to receive four different instructional services, and that the combination of these instructional services adds value to the literacy experience received by the child. Thus, comparing families who receive Even Start services with families who receive whatever services they obtain on their own (without Even Start) answers the policy-relevant question about whether Even Start's unified multi-service approach works better than an approach in which families find and use services on their own.

IMPLICATIONS OF THE FINDINGS

The fact that two experimental studies of Even Start show similar results, even though they were done at different times, one in the early 1990s at the very beginning of the program and a second after a decade of program implementation and many amendments to the program, points to the need to explore improvements if the Even Start model is to be an effective family literacy intervention. As implemented in the EDS projects, Even Start was not more effective than the mix of services that control group families obtained for themselves. Given Even Start's intuitive appeal as an approach for enhancing parent and child literacy, we interpret the lack of effectiveness as an indication that the Even Start approach needs to be strengthened. The remainder of this summary offers some ideas about this topic.

Family literacy is an important approach, but questions remain about its effectiveness.

Who can argue with the cornerstone of family literacy, that parents are their children's first and best teacher? A large research literature links levels of parental education to levels of child achievement (National Research Council, 1998, 2000, 2001). With regard to literacy and language development, a number of studies have shown a relationship between language-rich parent-child interactions and language development of young children (Hart & Risley, 1995; National Research Council, 1998, 2000, 2001; Powell & D'Angelo, 2000).

Building on these documented relationships, family literacy programs seek to improve the literacy development of young children by helping parents become more literate themselves, by helping parents understand more about how children learn, and by inculcating good teaching habits in parents, in addition to providing early childhood services directly to young children. However, no experimental evidence has been found to support the hypothesis that family literacy programs (or adult education programs more generally) can make large enhancements in parent literacy and parenting skills. Even assuming that it is possible to significantly alter parent literacy and parenting skills, research has not shown that these changes will translate into improved literacy performance among children in a timely manner.

In the current study, Even Start did not change the literacy skills or parenting skills of parents, nor did it change the literacy skills of children, over and above the changes that were seen in parents and children who did not participate in the program. This raises questions as to whether (1) Even Start families participated with sufficient intensity to derive the needed benefits, and (2) the services offered to Even Start participants were of sufficiently high quality and of the appropriate content.

Implication #1: Families did not participate long enough and did not get enough instruction to make the kinds of changes that are needed.

This evaluation has documented the difficult economic and educational circumstances faced by Even Start families. To design a family literacy program that meets the needs of families where half of the parents enter with less than a 9th grade education, where half of the families have annual income of less than \$12,000, and with parents and children who score at the very lowest levels on literacy assessments (on average, below the 5th percentile), we first must recognize that the changes required on the part of participating parents and children are much greater than previously realized.

Hence, it may well take much longer than the average of 10 months of participation for the changes in parent literacy, parent-child literacy interactions, and child literacy hypothesized by Even Start to occur, especially given the low literacy level of Even Start parents. Such low literacy would interfere with a parent's ability to be a successful teacher of his/her child, and years rather than months might be required before substantial improvements are seen. It is hard to imagine that parents who are having great difficulty with their own literacy skills will find it easy or natural to read and discuss books with their child or to talk to their child about the world

using varied vocabulary – two activities most strongly associated with family support for language and early literacy development (Bus, van Ijzendoorn & Pellegrini, 1995; Dickinson & Tabors, 2001; Hart & Risley, 1995). If it is unlikely that parents will become substantially better teachers of their children in a 10-month period, then there is little chance that improvements in child literacy will occur through their parents.

But what about early childhood education instruction? The most successful early childhood intervention programs have been able to make changes in child IQ scores of between 0.5 and 1.0 standard deviations (Ramey, et al, 1992; Barnett, 1995). Since the average Even Start child scores almost 2.0 standard deviations below national norms on the PPVT⁵, we might expect that the very best interventions developed to date could just about cut this deficit in half. Even if it were possible to improve child PPVT scores by a full standard deviation, Even Start children still would lag behind national norms by a substantial amount.

It is striking how few hours of early childhood education were received by children in Even Start families when compared with the hours received by children who participated in programs that have generated large effects on child outcomes. In 2000-2001, Even Start children birth to five years of age were offered an average of 591 hours of instructional service and received an average of 220 of those hours.⁶ This is only 20 to 25 percent of the amount of service offered to children participating in the Abecedarian project (Ramey & Campbell, 1988), in Project CARE (Wasik, Ramey, Bryant & Sparling, 1990), and in the Infant Health and Development Program (Ramey, Bryant, Wasik, Sparling, Fendt & LaVange, 1992). Each of these projects used the same curriculum in all sites and had large positive short-term effects of between 10 and 15 points (0.7 to 1.0 standard deviations) on standardized IQ tests.

The implication is clear – the early childhood programs that have produced the largest effects on child development are those that are able to engage children in full-time, year-round center-based instructional services using a single curriculum. Given the great needs of Even Start families, the best possible family literacy program would be able to engage families intensively for a very substantial period of time. Even the highest quality programs, those developed using research-based practices to teach instructional content, would not have an impact if parents and children do not attend sufficiently long or intensively. So one key question is whether Even Start's retention and participation rates can be improved.

Suppose that the participation levels seen in this study are the best that can be achieved by a family literacy program serving needy families. If so, then some of Congress' goals for Even Start are disconnected from the achievements that can reasonably be expected. For example, Congress wants states to use attainment of a high school diploma or a GED as an indicator of success. However, the low reading levels of Even Start parents suggest that few are likely to meet this criterion. Although parents in the EDS made significant improvements in

⁵ The average PPVT pretest standard score for Even Start children in the EDS was 72.9. In the first evaluation (early 1990s), the average PPVT pretest standard score was higher (79.8 points), but still more than a standard deviation below national norms.

⁶ Head Start offers children about 600 hours of instruction a year (four hours/day * five days/week * 30 weeks), just about the same as the average of 591 hours of early childhood instruction offered to Even Start children.

their literacy levels on most of the measures used, reading scores at the posttest were extremely low compared with national norms -- Even Start parents scored at about the 3rd grade level on the Woodcock Johnson test of reading comprehension (Exhibit 6.1.24). For children, Congress wants states to measure the extent to which children read on grade level. Even after significant improvement over a year of participation in Even Start, Even Start children in the EDS scored at the 6th percentile on the PPVT at posttest (Exhibit 6.1.1). Children with very low receptive vocabularies are not likely to be successful in learning to read when they start school. If we believe that participation is at a maximum level, then these goals ought to be modified.

An alternative view is that, with help, Even Start projects can do a better job of increasing the amount that families participate and the length of time they remain enrolled. This approach would require that the federal government and Even Start state coordinators provide explicit direction about the intensity of services that should be offered and the amount of participation that is expected on the part of enrolled families. In turn, local project staff must encourage families not only to join Even Start, but they also need to motivate families to remain in the program and to participate intensively. High expectations are important to the improvement of many areas of education, and participation in a family literacy program is no exception.

Increasing retention in Even Start also is dependent on the ability of federal and state governments to minimize conflicting requirements with welfare reform programs. In some states, Even Start fulfills welfare reform requirements, while in others, families have to leave Even Start to undertake other educational and job-related activities which comply with welfare reform requirements. Clearly, the former approach enhances retention of families in Even Start.

Implication #2: The quality and content of instruction on language acquisition is insufficient to meet Even Start's legislative goals and hence needs to be improved.

A fundamental hypothesis underlying the Even Start family literacy model, largely untested until this study, is that the presence and integration of all four instructional components will add value to literacy outcomes for children. The present evaluation prompts us to examine whether Even Start children, parents, and families are expected to have better literacy outcomes (1) because families participate in all of the instructional services, because the services are coordinated, and because some synergy is expected from receiving the combination of services, or (2) because Even Start instructional services are of higher quality than what exists elsewhere?

Instructional services need to be based on recent, scientific research. The evidence is that Even Start projects have spent the past several years operating under the first expectation listed above. The Even Start legislation and federal guidance point local projects in the direction of the first expectation. At the time of this study, the legislation contained 10 program elements as well as a myriad of other requirements. Of the 10 specific program elements, only one deals directly with the quality of the instructional services while the other legislative requirements deal with what might be termed "processes."

Even Start's process requirements specify that projects must provide four instructional components, as well as support services, on a year-round basis. In addition to center-based

services, Even Start projects must provide some educational services to families in their homes. Projects are to collaborate with other agencies to build on educational and support services that exist in their communities and provide an increasing local funding match. Projects have to integrate these services, including those not directly provided by Even Start, into a unified family literacy program. Even Start projects must identify, recruit, and serve families most in need of services, as well as screen and prepare those families for the substantial commitment involved in Even Start participation. Projects must serve children in at least a three-year age range. Projects must provide services for a wide range of adult learning levels, from adult basic education through GED preparation, as well as ESL classes. Projects also need to provide training to their staff, coordinate with related programs, provide for an independent local evaluation, cooperate with the national evaluation, and provide information for new state indicators of program quality.

As noted earlier, Even Start projects that participated in the EDS each implemented, at least minimally, all of these process requirements. However, there remains a legislative requirement mandating that Even Start projects “include high-quality, intensive instructional programs that promote adult literacy and empower parents to support the educational growth of their children, developmentally appropriate early childhood educational services, and preparation of children for success in regular school programs.” While the legislation provides no definition of high quality services and no guidance as to what intensive means, the evidence that this evaluation provides about quality shows that Even Start early childhood education represents mainstream instruction for children and is not of higher quality than the instruction received by control children, nor is it of higher quality than the instruction received by Head Start children. Adult education received through Even Start appears to be exactly the same as, not better than, adult education received on a stand-alone basis. And most parenting education delivered by Even Start relies on standard parenting curricula, typically Parents as Teachers.

Since the instructional approaches used in the Even Start EDS projects did not produce large gains in reading and literacy skills for Even Start participants, the program should move quickly to adhere to the No Child Left Behind legislation, where the programmatic emphasis is squarely on enhancing reading skills through application of rigorous, scientific research.⁷

Quality of instructional services should take priority over building on existing services. There is a tension in Even Start between spending the time and funds to deliver services directly as opposed to obtaining those services through existing agencies. Even Start projects are mandated to collaborate with local service agencies and build on existing services, in order to avoid duplication. This is a laudable goal, and Even Start projects have taken it to heart. According to reports from Even Start project directors, collaborating agencies often provide instructional staff, administrative or technical support, space and equipment, and community support. Perhaps the most important kind of collaboration involves the provision of instructional staff. Instructional staff were provided by public adult education programs to 51 percent of all Even Start projects, by public elementary schools to 40 percent, by Head Start to 35 percent, by community colleges to 33 percent, by state funded preschools to 33 percent, and by Title I preschools to 23 percent of Even Start projects (Exhibit 2.8).

⁷ Summaries of this research can be found by looking under Reading Resources on the Department of Education’s web site (www.ed.gov/offices/OESE).

However, Even Start can only be as effective as the services that families receive, whether they are provided by Even Start or by a collaborating agency. It is difficult for a local Even Start grantee to build a solid research-based program if it has not paid attention to, or has little control over, the quality of some or all of the educational services received by its families. Herein lies the tension. A project that develops its own program of instructional services has complete control over the quality of those services, but in doing so it may be duplicating services that are available elsewhere in the community and that might be used in a cost-effective manner. A project that builds on existing instructional services is complying with the legislative requirement to do so, but may be sacrificing service quality in the bargain. Given the results of this evaluation, quality of instruction ought to be the driving force in designing an Even Start program. Local projects ought to avail themselves of existing instructional services when those services are research-based and deemed likely to be effective. However, Even Start projects should not use existing services simply because they are available.

Quality might be enhanced by better targeting. Research shows that intensive, multi-year instructional services are better at enhancing children’s language development than a single year of service (e.g., Barnett, 1995). Building on this research, Even Start ensures that services are available for multiple years by requiring that projects design programs to serve children over at least a three-year age range. In practice, 96 percent of all Even Start projects offer services to children in a consecutive three-year span, 54 percent offer services to children from birth through age seven (Exhibit 2.37), and about half of all projects provide ESL, GED and ABE services to parents. Unfortunately, few families remain in Even Start long enough to take full advantage of what is being offered – the average family participates for 10 months.

Given that families generally do not participate for long periods of time, it is possible that by providing instructional services to parents with various needs and children of different ages, Even Start projects may be spreading themselves too thin. Perhaps more careful targeting of services to subgroups of families with similar backgrounds (e.g., families with parents that need ESL services, or parents that are close to getting their GED, or preschool age children) might either enable projects to focus instruction in a more effective manner, or allow projects to construct instructional services that would appeal to families for a longer period of time.

FUTURE DIRECTIONS

For more than a decade, Even Start has made programmatic improvements based, in part, on evaluation results. Several recent changes in the program are relevant to the findings from and implications of this evaluation.

State-level performance indicators. A 1998 amendment strengthened accountability in Even Start by requiring states to develop results-based indicators of program quality and to use these indicators to monitor, evaluate, and improve Even Start programs.⁸ All but a few local programs are administered by the states, so Even Start’s success is dependent to a large extent on

⁸ Title VIII of the United States Department of Labor, Health and Human Services, and Education Appropriations Act, enacted by the Omnibus Consolidated and Emergency Supplemental Appropriations Act, 1999, Section 101(f), P.L. No. 105-277 (1998).

the states' administration of the program, including making well-informed decisions about which programs to fund and which to continue. Development of performance indicators at the state level that call for local projects to collect data on child and adult outcomes, coordination of performance indicators across states, and using performance indicators to monitor and improve local projects could be a major step toward quality enhancements in Even Start.

All States were required to submit a plan for performance indicators to the Department of Education by June 2001. An analysis of these plans shows that they are of varying degrees of complexity and comprehensiveness. All states have developed ways to measure the six legislatively-mandated indicators. For some (e.g., high school completion, grade retention) the indicators are quite similar across states while for others (e.g., adult basic and literacy skills development, and child developmental gains), the measures and criterion standards are very different across states. Actual implementation of the indicators, including collection of data from local projects is beginning in most states in fall 2002.

Even Start's recent reauthorization. Even Start was reauthorized at the end of 2000 by the Literacy Involves Families Together Act and in 2001 by the No Child Left Behind Act, both of which attempted to improve the quality of Even Start projects. Even Start's purpose was amended to require that local projects build on existing community resources *of high quality* (the previous law did not explicitly require collaborator services to be of high quality). Further, Even Start's purpose now also includes promoting the academic achievement of children and adults, and using instructional programs based on scientifically based research.⁹

The new legislation contains five new required program elements, three of which are directly related to instructional quality. New program elements were established for the use of scientifically-based reading research in designing instructional services as well as the inclusion of reading-readiness activities for preschool children to ensure that children enter school ready to read. In addition, a new program element strengthens required staff qualifications.

Another new program element relates to attendance and retention in the program. Local projects are now explicitly required to encourage families to attend regularly and remain in the program a sufficient time to meet their program goals. The last new program element concerns continuity of family literacy. Projects must promote continuity to ensure that families retain and improve their educational outcomes. In addition to the new program elements, several existing program elements were amended. Projects now have to offer instructional (not just enrichment) services through the summer and local evaluations have to be used for program improvement.

The reauthorization of the program stimulates change by providing a more explicit focus on quality. These new programmatic requirements (i.e., using research evidence, building on existing resources of high quality, using local evaluations for improvement) will be best met if local projects continue to be provided with ongoing technical assistance from the Department of Education and state agencies.

⁹ Sec. 1201 of the Elementary and Secondary Education Act of 1965 (ESEA), as amended by P.L. 106-554.

For a decade, Even Start has worked on the administrative aspects of implementing a complex program (e.g., coordination and collaboration with local services, recruiting families most in need). The focus must now shift to improving the quality and intensity of instruction in each service component, even those components provided by a collaborator. The entire Department of Education is moving in this direction with the No Child Left Behind Act, which emphasizes that classroom literacy experiences must be developed based on scientific research. It will be critical to help local projects understand what has been learned from research and to provide guidance on quality and intensity standards.

Using evaluation to improve Even Start. The current evaluation has found that children from low-income families who attend traditional kinds of family literacy programs which are based on early childhood education programs of good general quality along with standard forms of parenting education and adult education, have no better literacy skills than control group children, nor do they gain more than expected when compared to national norms.

Future evaluation work will be most helpful to Even Start if it is designed to find, demonstrate or test effective family literacy practices – to identify and determine which practices and procedures work best and hence can be used as a template, or model, for improving Even Start projects across the nation. Assessing the quality and content of the instructional services offered by Even Start projects is another area where future research can improve on the work done in the past. And finally, future evaluations will need to carefully consider the best ways of assessing parents and children who have limited skills in reading, speaking and writing English.