



Counter to the *Third National Even Start Evaluation: Program Impacts and Implications for Improvement (2003)*

“Care should be given in applying the findings to Even Start projects as a whole” (p.9-10)

The William F. Goodling Even Start Family Literacy Program has undergone several national evaluations to measure its effectiveness. A key finding in the most recent evaluation, conducted by Abt Associates on behalf of the U.S. Department of Education, has led legislators to question the value of the Even Start program. The indicting key finding, “Even Start children and adults made gains on literacy assessments, but not more than adults and children in the control group, two-thirds of whom received no adult or early childhood education services” is addressed in this critique, which questions the validity of the Evaluation’s argument regarding the efficacy of Even Start. Further, it provides evidence of Even Start’s more recent success that has come since its reauthorization with the Learning Involves Families Together (LIFT) Act of 2000 and the No Child Left Behind Act of 2001.

The Evaluation preceded the development and implementation of Performance Indicators.

Programs in 1999-2001 were not held to the same accountability standard as currently funded Even Start programs. Even Start’s reauthorization provided an explicit focus on quality for instruction, collaborators, and participation. Programs now are required to provide instruction based on “scientifically-based reading research.” Further, staff qualifications have become much more restrictive, requiring the majority of existing teachers and all new teachers in adult and early childhood education classes to have an associate, bachelor or graduate degree in the appropriate field. Local projects are to collaborate with high quality partners (the previous law did not specify high quality). In addition, the LIFT Act of 2000 required states to develop and implement Performance Indicators, or standards, for Even Start participants outcomes, which have raised the level of gains for adults and children. Indeed, current funding decisions are determined by program performance against the state-set Performance Indicator benchmarks. Performance Indicators, generally, were not in place or implemented until 2002. The Evaluation itself cautions, “The current evaluation, covering the years 1997-1998 through 2001, reflects the program as it existed prior to the 2000 reauthorization” (p. 21).

The Evaluation does not reflect improvements to Even Start programs undertaken after changes in federal law in 2000 and 2001 and, therefore, is not currently applicable. Future efforts by the U.S. Department of Education to develop a national framework for program performance should lead to a “robust program performance system for Even Start programs.”¹

ESPIRS data, on which the evaluation was based, were not reliable.

The Evaluation was based partly on the ESPIRS data from 1997-1998 and 1998-1999, which were used to report on program and family characteristics, participation rates, and family progress. According to feedback from Even Start state and local evaluators², these data were suspect. Specifically,

¹ Senate Committee Labor/HHS/ED Appropriations Report, 2004.

² Evaluators participated in meetings hosted by the Goodling Institute for Research in Family Literacy at the 2003 and 2004 NESA and the 2004 NCFL conferences.

- Few programs reported outcomes for parents and children in ESPIRS, as these data were not required.
- During the time of the study, (1997-1999), the method of entering ESPIRS data was changed, and programs had difficulty with the new technology -- more data, thus, were lost.
- Data were reported directly to the federal government; thus states could not review it for accuracy.

Due to a flawed data collection process, the accuracy of the ESPIRS data used for the national evaluation was, therefore, not reliable. Most, if not all, findings based on those data are questionable.

The Experimental Design Study (EDS) design, on which the Evaluation was based, was problematic.

The EDS, employing 18 Even Start projects consisting of two program cohorts,³ was used to study the effectiveness of Even Start. The design structure is questionable due to the absence of a random sample of programs and the lack of sites representative of Even Start programs' national demographics.

- Projects volunteered for the study, instead of being randomly selected. Self-selection affects results. The Evaluation report, in fact, acknowledged that the evaluation plan could not be implemented perfectly.
- Selected programs over-represented Hispanic (75%) and urban (83%) populations in contrast to Even Start's national demographics of (46% and 55%), respectively. While these are growing populations, results cannot be generalized to the Even Start population, as stated in the Evaluation.
- Twenty-eight percent of the sites had been operating for only two years. Typically, new programs require several years to effectively implement the complex model. New Even Start grantees must locate and develop collaborative relationships with partners; provide or find high quality partners to provide the required educational components (adult education, parenting education, early childhood education, and parent-child interactive literacy); identify, hire, train and retain appropriate staff; and recruit families. Having over one-quarter of the EDS programs new and not fully developed to function effectively may have skewed the data.

The EDS did not include a representative sample of Even Start programs that should have been randomly selected. As a result, findings from the Study cannot and should not be applied to the Even Start program in general.

The EDS did not evaluate the effectiveness of the Even Start family literacy model.

The Evaluation states, "While the EDS sites represent functioning Even Start projects, they were not selected to be models of excellence" (p. 9). Since the purpose of the study was to evaluate the effectiveness of the Even Start model as opposed to services families obtain for themselves, it was critical that the model was a well-specified intervention and implemented in a specific way. Integration of services, the core of the Even Start model and the focus of the study, was not a criterion for site selection. Tables in the Evaluation show major differences among the EDS sites in their services. The importance of adhering to a particular model in a controlled experimental study (such as the EDS) cannot be overemphasized. Selecting projects that effectively integrate the family literacy model with fidelity is a necessary prerequisite in evaluating the efficacy of Even Start.

By selecting volunteer Even Start programs, not those that are implementing research and evidence-based practices and effectively integrating components, the Evaluation did not, in fact, study the model as it was intended to operate.

³ Cohort 1 (n=11), 1999-2001, had no follow up in Spring '01 and Cohort 2 (n=7), 2000-2001, had no follow-up in Spring '02.

According to recent statewide evaluations, Even Start is having a positive effect on participating families.

Results taken from recent statewide evaluations provide evidence of the effectiveness of Even Start. While only a sample are provided, the data represent results from at least ten states for which recent statewide evaluations were available. Over 600 projects are represented, far more than the 18 of the EDS. As evidenced below, in many states, Even Start adults and children are making statistically significant gains, suggesting that the gains are due to the program intervention and not due to chance.

Pre-School Children

- A study by the Goodling Institute for Research in Family Literacy at Penn State found that children between the ages of 3-5 who had participated in family literacy for at least 90 days exhibited statistically significantly⁴ more growth in all developmental areas than children the same age who were just beginning their participation.
- In Nebraska, children taking the Teacher Rating of Literacy and Language (TROLL) had statistically significantly greater gains in language and literacy for oral and reading skills.
- In Colorado's 14 programs, 94% of the Even Start infants and toddlers and 88% of pre-school children were at age appropriate levels, suggesting that Even Start children are not far behind the national norm as indicated in the Evaluation.
- In California, 67% of children entering kindergarten were rated as "fully mastering" or "almost mastering" reading readiness behaviors.
- In Kentucky, 95% of the children were on target for reading readiness, far exceeding the Performance Indicator benchmark of 75%.

School-Age Children

- Colorado conducted a follow-up study of Even Start children to determine long term effects on Even Start Children. Using results from the Colorado Student Assessment Program, teachers reported that Even Start students scored higher overall than non-Even Start comparison students. Over one-half (53%) of Even Start children were reading above grade level and 47% at grade level, contrasted to only 28% of the comparison children reading above grade level and 44% at grade level. A full 28% of the comparison children read below grade level while 0% of the Even Start children did.
- In California, 64% of children in kindergarten to second grade met grade level content standards in reading and math and 70% of the English language learners made progress in English skills.
- About 75% of Even Start elementary students in Massachusetts were on or above grade level in reading; 86% of their teachers rate their attitude as good-excellent, which will support them throughout their school years.
- In Connecticut 100% of school age children met the federally required Performance Indicator for attendance. In North Carolina and New York, school age children exceeded the Performance Indicator for attendance.
- Promotion to the next grade level is the third mandated Performance Indicator for children. A full 100% of children in Connecticut were promoted. Even Start children in other states are also doing well: 98% in California, 97% in Colorado, 96% in Pennsylvania, 94% in North Carolina, and 93% in Kentucky.

Adults

- Pennsylvania's adults (nearly 3000) in 2002-2003 met or exceeded Performance Indicators for seven of the nine adult assessments, marginally missing two. Over half (53%) with the goal of a diploma or GED met it, compared to only 11% in 2000-2001.
- New York's 2002 statewide evaluation of 70 programs (3155 adults) reports that adults exceeded all Performance Indicators in all of its assessments. Further, 79.9% entered post secondary education or training and 83.3% obtained employment, far surpassing the Performance Indicator of 50% for both.

⁴ Statistically significant means that the results occurred due to the intervention and not by chance.

- Adults in California exceeded the adult education Performance Indicators for English GED (61%), Spanish GED (64%), and high school diploma (61%).
- Adults in Nebraska and North Carolina far exceeded all Performance Indicators for adults at all levels--English speakers as well as English language learners.
- In Kentucky, 69% achieved a GED and 96% a high school diploma. A full 100% with the goal of entering post-secondary education or training did so.
- Colorado's adults made significant achievements with 41% earning a GED and 63% entering higher education or training. A special population of teen parents resulted in 80% staying in school and 78% graduating from high school.
- Oregon (eight programs) reports that 61% of enrolled parents improved their literacy skills, gaining at least one level or completing some or all of the GED tests
- Over 80% of Massachusetts Even Start parents made significant academic gains in communication, reading, and understanding children's learning and writing and two-thirds made strong gains in English language acquisition, math and problem-solving skills.

Other Outcomes

While Even Start is an educational program for adults and children, its primary focus is on the "Parent as the child's first and most important teacher." Statewide data indicate that parents take this concept seriously. Further, Even Start is intended to help families become better community members. The data from statewide evaluations provide evidence of the following:

- Parents read more to their children, have more books at home, and take their children to the library more often than before participating.
- Parents are now more informed about children's development and age appropriate expectations.
- Parents are now more active in their children's classroom, volunteer more and talk more with teachers.
- Parents now take better care of their and their children's medical and dental health.
- Parents have registered to vote or voted for the first time.
- Many parents obtained a driver's license.
- Parents are now more active in their community.

Recent statewide and local evaluations provide evidence of the effectiveness of Even Start as an intervention and therefore a national evaluation employing a different design is both timely and essential to further attest to its effectiveness. "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" (Third National Even Start Evaluation, 2003, p. 17).