

GAO

Testimony

Before the Subcommittee on Children and Families,
Committee on Labor and Human Resources, U.S. Senate,
and the Subcommittee on Early Childhood, Youth and
Families, Committee on Education and the Workforce,
House of Representatives

For Release on Delivery
Expected at 9:30 a.m.
Thursday, March 26, 1998

HEAD START

Research Insufficient to Assess Program Impact

Statement of Carlotta C. Joyner, Director
Education and Employment Issues
Health, Education, and Human Services Division



Head Start: Research Insufficient to Assess Program Impact

Messrs. Chairmen and Members of the Subcommittees:

I am pleased to be here today to discuss what is known about Head Start's impact on children and their families. I also want to discuss the adequacy of the Department of Health and Human Services' (HHS) current research plans to provide additional information on Head Start's impact.

For the past 30 years, Head Start has provided a comprehensive set of educational, health, mental health, and social services to low-income preschool children—an array of services generally not offered by other programs when Head Start began. Its ultimate goal is to improve the social competence of preschool children in low-income families. Head Start defines social competence as children's everyday ability to deal with both their current environment and later responsibilities in school and life. During these 30 years, the program has served over 15 million children at a total cost of more than \$30 billion. Head Start's funding has grown substantially in recent years, and further increases have been proposed. For example, between fiscal years 1990 and 1997, funding more than doubled—from \$1.5 billion to almost \$4 billion. The administration's goal now is to expand the program's annual enrollment to one million children by 2002.

Although Head Start has long enjoyed both congressional and public support, opinions about the program's impact have been divided. We define impact as differences in outcomes, such as improved school readiness or health status, caused by Head Start participation. Implicit in this definition is the concept that differences found would not have occurred without program participation. Conflicting information on program impact and the focus on results-oriented program performance information required by the Government Performance and Results Act (Results Act) of 1993 have renewed interest in the outcomes and impact of the current Head Start program. In response to this interest, in a 1997 study,¹ we reviewed the research literature on Head Start to determine what was known about the impact of the current program.

In my statement today, I will discuss the results of that research review, HHS' current initiatives to assess program impact, and possible ways to improve HHS' efforts. This discussion is based on our past work and our assessment of information from HHS about its new and planned initiatives, although we have not independently reviewed these initiatives.

¹Head Start: Research Provides Little Information on Impact of Current Program (GAO/HEHS-97-59, Apr. 15, 1997).

In summary, the Head Start program has provided comprehensive services to millions of low-income children and their families—services that in the program’s early years participants probably would not have otherwise received. Little is known, however, about whether the program has achieved its goals. Although an extensive body of literature exists on Head Start, only a small part of that involves program impact research. Because of these research studies’ individual and collective limitations, this body of research is insufficient for use in drawing conclusions about the impact of the national program.

HHS has the following initiatives it describes as impact assessments: (1) development of performance measures focusing on program outcomes, rather than just processes; (2) a national longitudinal study of a representative sample of Head Start children and their families (Family and Child Experiences Survey—FACES); and (3) a collaborative effort with the National Center for Educational Statistics (NCES). These efforts are headed in the right direction for Head Start to evaluate the impact of its program. It is unclear, however, whether these efforts will meaningfully compare the outcomes achieved by Head Start children and their families with those achieved by non-Head Start children and families, leaving unanswered questions about Head Start’s impact.

Background

Since 1965, Head Start’s primary goal has been to improve the social competence of children in low-income families, that is, their everyday ability to deal with both their current environment and later responsibilities in school and life. This considers the relationships between cognitive and intellectual development, physical and mental health, nutritional needs, and other factors. Head Start delivers, or provides access to, a wide range of services—educational, medical, dental, nutrition, mental health, and social services. HHS administers the Head Start program through its Head Start Bureau within the Administration for Children and Families (ACF).

Public and private nonprofit agencies that receive their funding directly from HHS provide Head Start services at the local level. These agencies include public and private school systems, community action agencies, government agencies, and Indian tribes. In fiscal year 1996, about 1,400 local agencies, called grantees, received Head Start grants. Grantees are required to obtain additional funding from nonfederal sources to cover 20 percent of their program costs. Head Start grantees work with various community sources to provide services. For example, some grantees

coordinate with public health agencies to obtain health services, while others contract with local physicians. Although all grantees operate under one set of performance standards, they have a great deal of discretion in implementing those standards, resulting in programs that vary.

In addition to providing services to children and families, Head Start sees one of its roles as a national laboratory for child development. Consequently, Head Start uses much of its discretionary research funding for demonstrations and studies of program innovations. The amount of funds allocated to research, demonstration, and evaluation has represented about 2 percent of the Head Start budget over the years. About \$12 million (about 0.3 percent of the Head Start budget) was so allocated for fiscal year 1997.

The main focus of the program's research, according to Head Start Bureau officials, has been to improve the program by exploring ways to maximize and sustain Head Start benefits. In addition, Head Start funds studies designed to answer questions on the effectiveness of new or innovative service delivery strategies. Such studies typically involve special program efforts and demonstration projects conducted on a trial basis at a few Head Start sites that focus on practices or services not typically found in regular Head Start programs.²

The passage of the Results Act in 1993 has heightened the importance of the type and direction of this research. The Results Act is designed to hold federal agencies accountable for achieving program results. The act specifically requires that agencies clearly define their missions, establish long-term strategic goals as well as annual goals linked to them, measure their performance according to their performance goals, and report on their progress. Agencies are also expected to perform discrete program evaluations and to use information from these evaluations to improve their programs.

The Results Act encourages a focus on delineating desired outcomes and developing performance measures to assess achievement of those outcomes. In addition, the Results Act focuses on objective and systematic assessments of the manner and extent to which programs achieve their intended objectives. In assessing outcomes, we are referring to achieving program purposes, such as promoting child wellness. As noted, we define

²The term "regular" Head Start refers in this testimony to programs that operate within the scope of established Head Start program options and under normal Head Start requirements. Regular programs are to be distinguished from demonstrations and other special programs that serve populations or offer services not normally found in Head Start.

impact as differences in outcomes caused by Head Start participation. Essentially, impact evaluations are the only way to answer the question, “Is this program making a difference?” Impact evaluation is a form of program evaluation that assesses the net impact of a program by comparing its outcomes with an estimate of what would have happened without the program. This form of program evaluation is used when external factors are known to influence the program’s outcomes; it isolates program contributions from other factors that may affect the achievement of program objectives. The most reliable way to determine program impact is to compare a group of Head Start participants with an equivalent group of nonparticipants. The preferred method for establishing that the groups are equivalent at the outset is to randomly assign participants to either a Head Start group or a comparison group, although other methods are valuable for estimating a program’s net impact.

Evaluation of Past Research

In 1997, we reported the results of our work on identifying what existing studies suggest about Head Start’s impact. To conclude that impact has been demonstrated, one would expect to see either (1) a sufficient number of reasonably well-designed individual studies whose findings could appropriately be combined to provide information on national impact or (2) at least one large-scale evaluation using a nationally representative sample. After locating and screening 600 studies and consulting with many early childhood researchers and officials at the Head Start Bureau, we identified only 22 studies that met the criteria for inclusion in our analysis.³

Of these 22 studies, many had individual methodological and design weaknesses, such as noncomparability of comparison groups, which raised questions about the usefulness of the findings. In addition, no single study had used a nationally representative sample so that findings could be generalized to the population of Head Start children.⁴ Because of our findings, we recommended that the Secretary of HHS include in HHS’ research plan an assessment of the regular Head Start program’s impact. This type of assessment is especially important because a large amount of funds are devoted to the Head Start program and other programs are competing for shrinking federal resources. Furthermore, the number of

³Our basic criteria were that Head Start participation took place in 1976 or later, that the studies compared Head Start participation with no preschool or some other kind of preschool, and that tests of statistical significance were reported. We limited our review to the current Head Start program, that is, the program in 1976 or later because of the substantial program changes occurring in the early to mid-1970s.

⁴In the late 1970s, HHS contracted for a national evaluation of the educational services component of Head Start. The study was implemented but never completed.

other early childhood programs for low-income families has been growing. Thus, the Congress needs to know with some certainty whether the federal investment in Head Start is making a difference.

In commenting on our earlier report, HHS said that the existing research on Head Start's impact was substantial and that the Department's strategy to expand this research was appropriate for determining both the program's impact and its quality. HHS also indicated plans to evaluate the feasibility of conducting impact studies such as we recommended.

HHS supported its claim that the existing research was substantial by noting the findings from a 1985 research synthesis of studies conducted in the 1960s and 1970s and two more recent studies. We disagreed, however, that findings drawn from studies more than 20 years old adequately support claims about the current program's impact. As noted, the current Head Start program operates in an environment that has changed in the last 20 years, when other, non-Head Start comprehensive early childhood services were not as available. Similarly, the findings from the two more recent studies did not support conclusions about program impact that can be generalized to the national program. Even though these two studies were larger than others we had found, both had significant methodological limitations.

Current Initiatives

HHS' current initiatives reflect its opinion that a randomized control group is not necessary to measure Head Start's impact. The current initiatives HHS describes as assessing impact include (1) the development of new performance measures, (2) a longitudinal study called FACES, and (3) a collaborative effort with NCES.

More specifically, HHS has described its performance measures as methods for annually—and over longer periods—assessing the quality and effectiveness of programs. As required by the Results Act, these measures will focus on both the program's results and the methods used to achieve these results. Throughout its history, Head Start's quality assessment efforts have focused on process indicators such as the number of teachers with degrees. Head Start measured these indicators by monitoring grantee compliance with mandatory performance standards. The new performance measures will begin to shift the focus from processes to the outcomes that Head Start children and their families are experiencing. This is an important effort, not only to improve program performance, but also to begin to lay a foundation for possible impact evaluations that could

assess the net impact of the Head Start program. It will allow Head Start to define and assess program outcomes, such as improved language skills, that it could then use to compare Head Start participants' outcome results with those of a control group to determine impact.

Another HHS initiative, *FACES*, is a study of a representative sample of Head Start children and their families intended to show whether Head Start is reaching its goal of improving children's social competence. According to HHS, for the spring 1997 pilot, data were collected from a sample of 2,400 families with children enrolled in 160 randomly selected centers in 40 Head Start programs nationwide. The full study will collect data from 3,200 families at program entry, exit, and at the end of kindergarten. HHS will conduct a more comprehensive validation substudy of 120 families. Researchers will use well-established and widely used scales, assessments, and observational protocols and specially tailored questionnaires to collect data on children's vocabulary, emergent literacy and mathematical skill, perceptual-motor development, and social and communicative competence before and after Head Start participation.

Head Start officials describe *FACES* as a way to draw conclusions about Head Start's impact in part because it will use nationally normed instruments.⁵ In addition, some of the *FACES* data elements will be the same as those in a Department of Education national household education survey.⁶ This will allow for comparing certain *FACES* results with a nationally representative sample of low-income children. It is not clear from our work so far how HHS will use the nationally normed data. According to HHS officials, the study will not compare Head Start children and their families with a randomly assigned control group of other children and families or with any other group.

In addition, the Head Start Bureau is collaborating with NCES on its Early Childhood Longitudinal Study. This study, implemented in fall 1997, after a 5-year planning effort, is collecting data on a nationally representative sample of kindergarten children in public and private schools, according to an NCES official. The Head Start Bureau has participated on the planning committee and provided some funding for this study. The study will

⁵Norms are obtained by administering a test to a sample of people and deriving the distribution of scores for that group. Some of the tests used by Head Start have been normed using samples selected to represent the national population for a particular age group.

⁶In 1993, the survey interviewed parents of a national probability sample of 4,423 3- to 5-year-old children, including 2,000 4-year-olds. Among these 4-year-olds, 244 from low-income families were reported to be attending Head Start; another 181 from low-income families had never attended any center-based preschool program.

collect data from parents and children, including descriptions of children's preschool experience and standardized tests in areas such as achievement and psychomotor development. This database will be available as a public-use tape for Head Start as well as other researchers. Head Start could use this database to compare groups of children in non-Head Start preschool programs with those in Head Start programs to assess program impact.

Improvement on Initiatives

Head Start's initiatives are headed in the right direction because of their increased focus on outcomes and research that could be expanded to compare outcomes for children in Head Start with those for similar children and families not served by the program. It is not clear how or whether Head Start will make these comparisons, however, using nationally normed tests or comparison group data from NCES. In addition, either of those research designs provides a much weaker basis for drawing conclusions about impact than a study with randomized assignment. For example, if Head Start uses NCES data for comparisons, the results could provide some indication of program impact. Some question will always remain, however, about the degree to which preexisting differences in the groups may have affected study results. True experimental designs, also called randomized trials, eliminate such questions. Randomized trials are comparison group studies that randomly assign study participants to either a treatment or control group. In the case of Head Start, these studies would require recruiting more eligible children than the program can serve. From these recruits, some would be randomly assigned to Head Start; the rest, the unserved children, would constitute the control group. HHS officials cited ethical considerations of assigning children to an unserved control group as one of the difficulties in conducting randomized trials.

Randomized trials, however, could be appropriately applied to Head Start research as long as Head Start lacks the resources to serve all eligible preschool children.⁷ While acknowledging the difficulties of random assignment, some early childhood researchers with whom we spoke

⁷In our 1997 report, we discussed several alternative research designs that use random assignment. We stated that a research design that randomly accelerates or delays rather than withholds services could be used. This would involve selecting a study group and randomly assigning some children to Head Start the first year, while the remainder would serve as a control group. The control group would receive Head Start services the following year. Another strategy that could be used to study specific parts of the program would be to use an alternative treatment design. In this case, some randomly assigned participants would receive the full Head Start program, while others would receive partial services. For example, if the study interest is in school readiness and cognitive issues, the control group might receive only nutritional and health services.

suggested that Head Start conduct randomized trials to study regular Head Start programs because this type of study provides the most conclusive information on program impact. In fact, the evaluation of the Early Head Start program, now under way, has randomly assigned potential participants to Early Head Start or a control group that has not received Early Head Start services. Control groups of randomly assigned participants are important to determining impact because they prevent mistakenly attributing outcomes to program effects when these outcomes are really caused by other factors. For instance, a recent evaluation of the Comprehensive Child Development Program, a demonstration project involving comprehensive early childhood services like those of Head Start, found positive changes in the families participating. The study had a control group, however, and researchers discovered that the control group families also had similar positive changes. They concluded therefore that the positive changes could not be attributed to the program.

Conclusions

Although impact research can be costly and time consuming, the federal government has made a considerable financial investment in the Head Start program; therefore, Head Start warrants a close examination to determine what the public is getting for its investment. Head Start has devoted substantial resources to research and evaluation activities, including some long-term studies and studies involving comparison groups. Although these have been worthwhile efforts, they have not sufficiently focused on evaluating results. HHS is taking steps that may help lay the groundwork for efforts to evaluate the net impact of Head Start program services. Identifying performance measures is an important first step in building a research and impact evaluation base for Head Start. In addition, this effort could yield a set of common measures upon which a body of research, including impact research, could be built. Similarly, the information gained in FACES should be extremely useful, especially to the extent that it is nationally representative.

HHS efforts, however, do not include plans for a research study or set of studies that will definitively compare the outcomes achieved by Head Start children and their families with those achieved by similar non-Head Start children and families. Although definitive results could take years to obtain, questions about Head Start's impact will remain unanswered unless these plans are expanded.

**Head Start: Research Insufficient to Assess
Program Impact**

Messrs. Chairmen, this concludes my statement. I would be happy to answer any questions you or members of the Subcommittees may have.

Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are \$2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. VISA and MasterCard credit cards are accepted, also. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:

**U.S. General Accounting Office
P.O. Box 37050
Washington, DC 20013**

or visit:

**Room 1100
700 4th St. NW (corner of 4th and G Sts. NW)
U.S. General Accounting Office
Washington, DC**

**Orders may also be placed by calling (202) 512-6000
or by using fax number (202) 512-6061, or TDD (202) 512-2537.**

Each day, GAO issues a list of newly available reports and testimony. To receive facsimile copies of the daily list or any list from the past 30 days, please call (202) 512-6000 using a touchtone phone. A recorded menu will provide information on how to obtain these lists.

For information on how to access GAO reports on the INTERNET, send an e-mail message with "info" in the body to:

info@www.gao.gov

or visit GAO's World Wide Web Home Page at:

<http://www.gao.gov>

**United States
General Accounting Office
Washington, D.C. 20548-0001**

**Bulk Rate
Postage & Fees Paid
GAO
Permit No. G100**

**Official Business
Penalty for Private Use \$300**

Address Correction Requested
