

Head Start's Lasting Benefits

W. Steven Barnett, PhD; Jason T. Hustedt, PhD

The benefits of Head Start are under increased scrutiny as Congress debates its reauthorization. How effective is Head Start, and how can it be improved? We provide a current overview and critical evaluation of Head Start research and discuss implications of this research with an eye toward informing debate. There has been a good deal of controversy over whether Head Start produces lasting benefits, dating back to its early years. Our review finds mixed, but generally positive, evidence regarding Head Start's long-term benefits. Although studies typically find that increases in IQ fade out over time, many other studies also find decreases in grade retention and special education placements. Sustained increases in school achievement are sometimes found, but in other cases flawed research methods produce results that mimic fade-out. In recent years, the federal government has funded large-scale evaluations of Head Start and Early Head Start. Results from the Early Head Start evaluation are particularly informative, as study participants were randomly assigned to either the Early Head Start group or a control group. Early Head Start demonstrated modest improvements in children's development and parent beliefs and behavior. The ongoing National Head Start Impact Study, which is also using random assignment, should yield additional insight into Head Start's effectiveness. We conclude with suggestions for future research. **key words:** *early education, Head Start, long-term benefits, policy*

HHEAD START is our nation's foremost federally funded provider of educational services to young children in poverty. Since 1965, more than 21 million children have participated in this comprehensive child development program (U.S. Department of Health and Human Services, 2003a). As a comprehensive program, in addition to its educational services, Head Start also provides social, health, and nutritional services to children and their low-income parents. When Early Head Start was established in 1994, the program was expanded to serve even younger children (from birth to age 3) and their families. By 2002, the Head Start program reported funding more than 910,000 children with a budget of \$6.5 billion (U.S. Department of Health and

Human Services, 2003a). However, Head Start remains a promise unfulfilled. Nearly 10 years after Congress authorized full-funding, Head Start's budget is still insufficient to serve all eligible children or deliver uniformly high-quality services to all enrolled.

As one of the most prominent educational and social programs in the United States, Head Start has attracted both proponents and detractors. How effective is Head Start as an early education program for disadvantaged children? What are the long-term benefits associated with participation in Head Start? These are questions that are reconsidered in each authorization cycle, when the program's benefits come under increased public scrutiny. Head Start was most recently reauthorized by Congress in 1998 and was scheduled for reauthorization again in 2003, although this process has not yet been completed. This article critically reviews the research on Head Start and other early education programs for at-risk children. We also discuss the implications of this research for issues that are likely to arise during reauthorization. Finally, we present recommendations for future studies of Head Start.

From the National Institute for Early Education Research, Rutgers, The State University of New Jersey, New Brunswick.

We extend our thanks to the anonymous reviewers of this article for their helpful suggestions.

Corresponding author: W. Steven Barnett, PhD, National Institute for Early Education Research, Rutgers, The State University of New Jersey, 120 Albany St, Suite 500, New Brunswick, NJ 08901 (e-mail: sbarnett@nieer.org).

THE WESTINGHOUSE STUDY

Controversy over the benefits of Head Start dates back to its earliest years, when a study by Westinghouse Learning Corporation and Ohio University (1969) reported that the program produced few sustained effects. This was the first prominent effort to investigate Head Start's impacts over time. Former Head Start children identified in first, second, and third grades were compared to schoolmates within the same grades who had not participated in Head Start, with a focus on cognitive and social-emotional development. Children from the Head Start and comparison groups were matched within grades on other important characteristics including ethnic group, gender, socioeconomic status (SES), and kindergarten attendance.

The Westinghouse study was immediately and widely criticized on methodological grounds (Condry, 1983). However, no one appears to have noticed at the time the most serious methodological flaw. The post hoc selection of the 2 groups literally equated the children on grade level. This biases between-group comparisons to the extent that differences in grade retention rates and special education placements truncated the samples, thereby eliminating the higher percentage of lower performing children from the comparison groups. The most obvious evidence that the comparison group does not represent comparable cohorts is that an increasing age gap is found moving across the grades with children in the third-grade comparison group significantly older than the third-grade Head Start children. Despite this and other evidence of methodological problems, the Westinghouse study continues to be cited in policy debates as evidence that Head Start does not produce sustained educational benefits for children in poverty.

FINDINGS FROM SHORT- AND LONG-TERM STUDIES OF HEAD START

Since the publication of the Westinghouse study, Head Start has continued to draw re-

searchers' attention. A number of longitudinal studies have followed former participants over time to gather more information about the benefits associated with Head Start. This research can be divided into 2 general categories: short-term and long-term studies. For the purposes of this review, we consider studies with immediate outcome measures and longitudinal studies with outcome measures taken earlier than third grade to be *short-term* studies, and consider studies with outcomes measured in third grade or later to be *long-term* studies.

A brief summary of the key findings from the short-term studies follows, since our primary interest is in Head Start's long-term benefits. Evidence of short-term benefits of preschool programs including Head Start has been thoroughly reviewed elsewhere (Barnett, 2004; McKey et al., 1985; Nelson, Westhues, & MacLeod, 2003; Ramey, Bryant, & Suarez, 1985; White & Casto, 1985). Studies have generally shown that programs for children at risk, including Head Start, result in increases of 0.5 standard deviations in IQ and achievement. Estimated impacts on measures of social behavior, self-esteem, and academic motivation typically are slightly smaller.

A recent short-term study by Abbott-Shim, Lambert, and McCarty (2003) is particularly notable for using random assignment of eligible 4-year-olds who had applied to a large Head Start program with a waiting list. This procedure allowed the researchers to rule out selection bias as an influence on results. Abbott-Shim et al. (2003) found that Head Start participants benefited substantially compared to nonparticipants in the areas of receptive vocabulary and phonemic awareness and had more positive health-related outcomes, for example, they were more likely to be current on their immunizations. And the parents of Head Start children reported more positive health and safety habits than the parents whose children did not attend Head Start. Because of the strength of the research design used in this study, these outcomes provide strong support for the short-term effectiveness of Head Start.

Some past reviewers (Haskins, 2004; McKey et al., 1985; White & Casto, 1985) have found that positive impacts of Head Start and early childhood programs for disadvantaged children decrease over time and eventually fade altogether. However, recent meta-analyses of longitudinal studies (Gorey, 2001; Nelson et al., 2003) suggest that effects persist over time although there may be some diminution of effects over the long term. These findings are consistent with the work done by Barnett, Young, and Schweinhart (1998), who used causal modeling to show that long-term effects of early childhood education are built upon short-term effects.

Reviews focused on long-term studies of early education programs serving economically disadvantaged children (eg, Barnett, 1998, 2004) find that the evidence regarding Head Start's long-term outcomes is mixed. In a recent examination of Head Start's long-term cognitive effects, Barnett (2004) identified only 39 studies in which educational programs included treatment and control groups, served children from low-income families, began during or before the preschool years, and were followed up with cognitive or academic measures at least through third grade, of which 15 were studies of "model" programs and 24 were studies of large-scale public programs. Twelve of the public program studies focused on Head Start, and an additional 4 included both Head Start and public school programs. Several of the model program studies, but none of the large-scale public program studies, employed random assignment.

Studies of model programs typically show initial gains in children's IQ scores that fade out over time (Barnett, 2004). Studies of large-scale programs have less often measured IQ, although the Peabody Picture Vocabulary Test has sometimes been used as a proxy for verbal IQ, making it more difficult to evaluate whether Head Start produces persistent IQ gains. However, it is likely that initial increases in IQ scores by Head Start children also fade out over time. Findings regarding other types of benefits are more promising.

Studies of both model and large-scale programs find achievement effects. In some studies the effects persist, in others effects on achievement cease to be statistically significant over time. Fade-out is frequently associated with high attrition over time or with other design flaws that affect the collection of achievement test data. Yet, decreases in children's later rates of grade retention and special education placements are found in most studies of model and large-scale programs. This apparent inconsistency often can be explained by differences in data collection procedures that lead to greater, more-biased attrition for achievement test data (Barnett, 2004). Few studies have measured impacts on high school graduation, but those with the largest samples reported statistically significant positive impacts (Barnett, 1998).

Overall, it appears that model programs and large-scale programs such as Head Start have similar *types* of effects, but the studies of model programs found effects of *greater magnitude* (Barnett, 2004). Given the variation in populations, programs, and contexts across studies, it is difficult to identify a single cause for this difference in effectiveness. Yet, it seems highly plausible that programs such as Head Start lack the type of funding necessary to produce the levels of intensity and quality achieved in better funded model programs with the direct result that they are less effective. Several studies provide direct evidence in support of this argument (Barnett, 1998). Some of this is discussed below in the context of findings from key studies of non-Head Start preschool interventions.

OUTCOMES FROM OTHER PRESCHOOL INTERVENTION PROJECTS

The Carolina Abecedarian Project (Campbell & Ramey, 1994, 1995) is one of the most notable studies of a model program to provide high-quality early education services to at-risk children. Participants were identified in the 1970s as infants, on the basis of their parents' low-income status as well as other risk factors predictive of cognitive difficulties

in childhood. The sample ($N = 111$), which was primarily African American, was divided into experimental and control groups by random assignment. Experimental group children attended the full-day, year-round Abecedarian program until age 5. Another randomization took place before children started school, with half the members of both the control and the experimental groups receiving an additional 3-year intervention. Thus, participants in this study received from 0 to 8 years of intervention services, with variation in its timing. Follow-up results have now been reported through age 21 (Campbell, Pungello, Miller-Johnson, Burchinal, & Ramey, 2001; Campbell, Ramey, Pungello, Sparling, & Miller-Johnson, 2002).

Findings from the Abecedarian Project show that the program produced large initial effects that persisted long after the intervention ended (Campbell et al., 2001, 2002; Campbell & Ramey, 1994, 1995). At the age 21 follow-up, Campbell et al. (2002) found that program effects were strongest for young adults who had taken part in the (5-year) preschool phase of the intervention. When compared to the preschool control group, these adults showed stronger performance on measures of academic skills and IQ. At age 21, they also were more likely to be enrolled in 4-year colleges, were better educated overall, and were more likely to hold skilled employment. Further, cost-benefit analysis of the Abecedarian Project (Masse & Barnett, 2002) shows that its overall benefits outweigh its costs, on the order of \$4 saved for every dollar spent on the preschool intervention (present value discounted at a real rate of 3%).

Research on the Chicago Child-Parent Centers (CPC; Reynolds, Temple, Robertson, & Mann, 2002) provides evidence of the long-term effects of a public-school-operated preschool program. The CPC program began in 1967 and classrooms are located in or near public schools in Chicago's highest poverty neighborhoods. From age 3 until age 5, participants attend 2.5 hour classes 5 days a week during the school year and a 6-week summer program is also generally provided.

After attending kindergarten, participants receive less intensive services through the public schools until age 9. Longitudinal follow-ups of the CPC cohort born in 1980 have been completed through age 21, on the basis of 2 study groups created beginning in 1985: former participants in the preschool and kindergarten phases of the CPC program ($N = 989$) and a comparison group of nonparticipants ($N = 550$). Members of the comparison group were matched to former preschool participants using SES and other demographic factors. Reynolds and colleagues (2002) report positive long-term outcomes from CPC across a wide range of domains. These include persistent gains in reading achievement (age 14), lower rates of grade retention and special education, lower rates of reported child maltreatment (ages 4-17), lower rates of juvenile arrests, and higher rates of educational attainment. A cost-benefit analysis estimates that the CPC preschool program yields an economic return far exceeding its cost (Reynolds et al., 2002).

RECENT RESEARCH ON HEAD START'S LONG-TERM OUTCOMES

Although long-term longitudinal evaluations of benefits associated with the Head Start program have been rare, several recent studies have sought new evidence. In a follow-up to the Head Start Planned Variation study conducted from 1969 to 1972, Oden, Schweinhart, Weikart, Marcus, and Xie (2000) compare 22-year-olds who attended Head Start at age 4 to others who had not attended, in 2 communities, 1 in Florida ($N = 424$) and 1 in Colorado ($N = 198$). Former Head Start participants were located as young adults, and a post hoc comparison group was constructed using young adults who had lived on the same streets or in the same high-poverty neighborhoods (Census tracts) as the Head Start participants. Members of the comparison group had not attended Head Start or any other early education program. However, perhaps because many children from the communities' lowest SES families had attended

