DOCUMENT RESUME

ED 443 573 PS 028 736

AUTHOR Abbott-Shim, Martha; Lambert, Richard; McCart, Frances
TITLE A Study of Head Start Effectiveness Using a Randomized Design.
SPONS AGENCY Administration for Children, Youth, and Families (DHHS), Washington, DC.
PUB DATE 2000-06-00
NOTE 49p.; Paper presented at the Head Start National Research Conference (5th, Washington, DC, June 28-July 1, 2000).
CONTRACT 90YD0016/02
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Behavior Problems; Child Development; Child Health; Emergent Literacy; Health Promotion; Parent Attitudes; Parent Background; *Preschool Children; Preschool Education; *Preschool Evaluation; *Program Effectiveness; *Program Evaluation; Research Methodology; Social Development; Vocabulary Development
IDENTIFIERS Phonemic Awareness; *Project Head Start; Randomized Experiments

ABSTRACT

Although an extensive body of literature exists on the impact of Head Start, very few studies have used an experimental design with random assignment, a key methodological component needed to increase the weight of evaluation findings. This study used a randomized design with a wide range of outcomes related to school readiness to assess the impact of the Head Start program. Participating in the study were all eligible 4-year-old applicants and their parents within selected centers in a Head Start program in a southern urban setting. Participants were assigned randomly to Head Start or control group conditions after eligibility determinations had been made. Head Start participants attended the program for at least 7 months. Child assessments were made in the areas of health, social skills, cognitive skills, and preliteracy skills. Analysis techniques included growth curve modeling and traditional analysis of variance. The initial status of the two groups was equivalent; the growth rates for the Head Start children were statistically faster than for the control children for receptive vocabulary and phonemic awareness measures. There were no significant differences between the groups or change over time in social skills or positive approach to learning. For Head Start children, there were no differences in ratings of problem behaviors between the fall and the following spring; ratings of problem behaviors significantly declined for the control group during the same time period. A greater proportion of Head Start parents than control parents reported improvement in some of their child's and their own health habits. (Contains 30 references.) (KB)

A Study of Head Start Effectiveness Using a Randomized Design

Martha Abbott-Shim
Georgia State University

Richard Lambert
University of North Carolina at Charlotte

Frances McCarty
Georgia State University

Correspondence regarding this paper can be sent to:
Martha Abbott-Shim, Ph.D.
Department of Educational Psychology and Special Education
Georgia State University
University Plaza
Atlanta, GA 30303
phone: 404.651.0399
fax: 404.651.1415
email: mabbottshim@gsu.edu

This research was funded by the Administration on Children, Youth, and Families, U.S. Department of Health and Human Services (Award No. 90YD0016/02). The authors wish to thank other members of the research team and the Head Start staff whose major contributions made this study possible.

Abstract

A randomized design study with a wide range of outcomes related to school readiness, including health, social skills, cognitive skills, and pre-literacy skills was conducted with eligible four-year-old applicants and their parents within a southeastern Head Start program. Children and their families in the Head Start treatment and control groups were given a battery of assessments. The study used growth curve modeling and traditional analysis of variance when only two measurements of outcomes were available. Initial status was equivalent and the growth rates for the Head Start children were statistically significantly faster than the control children on the receptive vocabulary and, phonemic awareness measures. There was a statistically significant time by group interaction and main effect of time for the problem behavior index of the social functioning measure. The parent report of health outcomes also showed statistically significant differences between the two groups.

Introduction

Head Start has been widely regarded as one of the most successful social programs of the last thirty years. Many research studies, using a variety of methods and outcome measures, have provided useful information about program characteristics and the effects of participation. Numerous reviews and quantitative research syntheses have documented the cumulative evidence for the value of the program (McKey, Condelli, Ganson, Barrett, McConkey, & Plantz, 1985). Despite the weight of this evidence and the collective expertise and experience of parents, children, educators, and researchers, the U. S. General Accounting Office Report (1997) put forth the conclusion that little evidence exists to illustrate the impact of the program.

Several themes recur within the U. S. General Accounting Office (GAO) 1997 report. First, it is suggested that the Head Start community should be held to a high standard as to what constitutes conclusive evidence of program effectiveness. The suggested standard is similar to the randomized clinical trial standard. However, in acknowledgement of the realities of conducting educational research in field settings, the GAO report contains thoughtful and reasoned explanations of the difficulties in applying such a standard to a program like Head Start. The second theme addresses the nature of what constitutes effectiveness. The GAO report takes the view that in order to be considered effective, the program must have far reaching and long term impact. The implication is that the existing studies on Head Start effectiveness have taken too narrow a view of outcomes, do not have a sufficiently long term perspective, and have not focused enough on sub-group comparisons (GAO, 1997), all making definitive statements about effectiveness difficult to present.

In support of this position, the Head Start Amendments of 1998 (P.L. 105-258) directed the Secretary of Health and Human Services to establish an expert panel charged with providing

recommendations for a study that provides a national analysis of the impact of Head Start. The Advisory Committee on Head Start Research and Evaluation provided a framework for studying the impact of Head Start. Impact was defined as "a difference in an outcome for a participant in the program that would not have occurred without participation in the program" (U.S. Department of Health and Human Services, 1999, p. 2). Subsequently, a request for proposals to conduct a national impact study was announced (U.S. Government Printing Office, 2000).

Although an extensive body of literature exists on Head Start, very few studies have utilized an experimental design with random assignment, a key methodological component needed in order to increase the weight of findings. Only one Head Start study, The Effects of Head Start Health Services (Fosburg & Brown, 1984), has used random assignment to form the Head Start and non-Head Start groups. This study was conducted in four large Head Start programs located in four regions of the country representing both urban and rural settings. The sample included pre/post-test data for 208 children and post-test only data for 609 children. It was found that Head Start children were more likely to receive preventative and remedial health services than other low-income children in their community. While an important finding on health services offered and utilized throughout Head Start, this study does not report findings on cognitive and social outcomes. The National Early Head Start Evaluation Study has also used an experimental design with the random assignment of eligible families into treatment and control groups, however findings are not yet available (Kisker, Love, & Raikes, 1999).

Another Head Start study that employed random selection of centers is the Effects of Mainstreaming Handicapped Children studies (Applied Management Services, Inc., 1978; Roy Littlejohn Associates, Inc., 1985). It included 55 randomly selected Head Start centers and 44 non-Head Start programs. The sample was comprised of 391 Head Start children, 321 non-Head

Start children receiving services from other programs, and 121 unserved children. The developmental gains for children who were physically handicapped, mentally retarded, and health or developmentally impaired in Head Start and non-Head Start programs were generally not greater than those of unserved children. However, the developmental gains in physical, self-help, academic, and communications skills were greater for children identified as speech impaired in the Head Start and non-Head Start programs as compared to the children who were not attending a program.

A small body of literature addresses the effects of Head Start by examining comparative gains for children in Head Start to children in control and comparison groups. The Educational Testing Service Head Start Longitudinal Study (Lee, Brooks-Gunn, & Schnur, 1988) included a sample of 414 Head Start children, 112 children enrolled in other preschool settings, and 225 children not enrolled in preschool from two urban locations. Group assignment was subject to family choice and was not under experimental control. There were large initial group differences between the Head Start, and control and comparison groups, with the Head Start children scoring lower on nearly every demographic and cognitive measure. When looking at the effect sizes of gain scores, Head Start children gained significantly more than children in the "no preschool" control group and the "other preschool" comparison group. The largest Head Start advantages were for the PPVT and Motor Inhibition tests. When controlling for race and background, Head Start had a significant and positive effect for African-American children on the Preschool Inventory, Motor Inhibition, and Toy Sort measures.

Zigler, Abelson, Trickett, & Seitz (1982) studied changes in intelligence quotient scores which were assessed at pre-test, retest, and post-test for 59 children attending Head Start and 25 comparison children from low-income families in a northeastern metropolitan community. It

was found that both groups increased from pre-test to retest, a result which was attributed to familiarity with the testing situation. The Head Start group continued to show improvement on the post-test which was administered in kindergarten. This improvement was attributed to gains in motivation that children acquire while attending a preschool intervention program.

Most of the Head Start effectiveness literature includes studies using post-test only designs with Head Start and matched comparison groups. In two studies of Head Start children compared to children with other preschool experience and no preschool, the preschool groups (Head Start and other preschool) performed better on cognitive measures (Bryant, Peisner-Feinberg, & Clifford, 1993; Currie & Thomas, 1995). In other studies Head Start children have scored higher on reading achievement (Nystrom, 1988; Williams, 1988), and vocabulary, language skills, concept development, and social adaptive behavior (Texas Instruments Foundation, 1996) than non-Head Start children. In one study Head Start children's performance was compared to children who had been on the Head Start waiting list. No significant differences were found on reading readiness scores in kindergarten, first or second grades (Bee, 1981). Another study compared Head Start children to children who had applied to Head Start but did not attend. At the end of third grade the Head Start group had a higher percentage of students who scored above the 80th percentile on one of the sub-tests of the Cognitive Abilities test. On the other Cognitive Abilities sub-tests, the California Achievement Test, special education placements, and grade retention there were no differences (Hebbeler, 1985).

However, other studies have found no differences between the Head Start and non-Head Start children on developmental measures (Esteban, 1987), on standardized achievement measures (Hunt, 1987), and on measures of locus of control, self-concept, and cognitive abilities (Roberts, 1984). Reedy (1991) found no difference between Head Start and non-Head Start

children on reading, math, written language, and receptive language measures; however the Head Start children did score higher on the measure of general knowledge and lower on the measure of maladaptive behavior.

The Head Start effectiveness research includes only one experimental design with random assignment of children to treatment and control conditions. Many of the other studies suffer from relatively weaker designs, specifically the noncomparability of comparison groups. In addition, the studies report varied results in response to the question of Head Start's impact. Therefore, the Georgia State University Quality Research Center (GSU QRC), in partnership with a Head Start program, conducted an experimental design study aimed at examining the impact of Head Start in a southeastern, metropolitan community. The purpose of this article is to describe and discuss both the implementation and the findings of this randomized design research effort with a wide range of outcomes related to school readiness, including health, social skills, cognitive skills, and pre-literacy skills.

Method

Participants

The participants in this study included all eligible four-year-old applicants and their parents within selected centers of the participating Head Start program located in a southern urban setting. The three participating centers were selected because the number of families recruited in the communities served by the centers offered the opportunity to form treatment (Head Start) and control (waiting list) groups. A random assignment procedure, developed in cooperation with the program director and several of the program staff, was used to place participants in the treatment and control group conditions after eligibility determinations had

been made. Consent to participate in the on-going research effort with the Head Start program was obtained from all families at the time of their eligibility determination.

The program's recruitment efforts were not altered for this study and encompassed an extensive penetration of the defined geographic area. The recruitment process included the completion of an eligibility form that delineated information about the child's and family's need for services. Recruitment efforts were initiated in January 1998 and continued through the summer. It was felt that it would be critically important to wait as long as possible for the recruitment efforts to saturate the communities. However, there needed to be adequate time allocated after the assignments were made to contact the families and inform them of their status before the school year began. Therefore, the random assignment procedure used to select Head Start and control group participants was completed in late summer.

In conducting the random assignments of children, the following exclusion categories were identified in collaboration with the Head Start program: children with disabilities, returning or transferring Head Start children, siblings of a child already assigned to either the Head Start or control group, and children with a very high "need" score on the Head Start program's eligibility scale. Children identified as being in one of these categories were not subject to the random assignment procedure. Instead, they were automatically accepted into the program. The remaining slots in each of the participating centers were then filled using a random assignment procedure that was carried out in two phases in accordance with the program's normal recruitment time line. The first phase took place at the end of July and the second phase in the middle of August.

The random assignment procedure was conducted at the center level rather than the program level. It was felt that this would serve to strengthen the design by allowing for greater

equivalence of geographic and demographic variables for children and families in the treatment and control conditions. The randomization procedure that was used to create the Head Start and control groups involved assigning each four-year-old child who was eligible for Head Start, had applied for enrollment, and was not automatically placed in the program, a randomly generated identification number. These identification numbers were then used to make the Head Start and control group assignments. In an effort to address some of the ethical implications associated with the implementation of this design, it was decided that control group children would be placed on a waiting list and would migrate into the Head Start classroom with the availability of slots according to their previously assigned random number.

The Head Start program sent letters to the Head Start families selected in the first phase, and subsequently sent letters to both the additional Head Start families selected and the control families based on the second phase selections. The control group letter indicated that the family was on the Head Start waiting list and could chose to participate in the Head Start Effectiveness study, a sub-study of the on-going research effort. Subsequently, 86 children were assigned to seven Head Start classrooms (Head Start group) and 80 children were placed on the wait list (control group). In addition, 43 slots in the seven classrooms were filled by children who were granted automatic acceptance into the Head Start program because they fell into one of the previously described exclusion categories. The majority of these children had a disability (11 children) or were returning or transferring Head Start children (27 children). There were three children with a very high "need" score on the Head Start program's eligibility scale that were also automatically accepted. In addition, as other eligible children and families requested Head Start services, they were informed about the Head Start Effectiveness sub-study, assigned a

randomly generated number, and placed on the waiting list (control group). Therefore, the Head Start and control groups were not initially equal in size and fluctuated in size until November 1.

Procedures

Head Start Treatment.

The Head Start participants attended a nationally recognized Head Start program comprised of 46 (98%) high quality classrooms which have been accredited by the National Association for the Education of Young Children (NAEYC). The Head Start program implements a theme-based curriculum with a wealth of classroom resources available to the teachers. The Head Start partner and researchers decided that a child must be enrolled for a minimum of 7 months (November – May) in order for the child to have received the Head Start "treatment." All of the Head Start teachers and family service workers were blind to the design and the random assignment procedures used in this study. The staff employed by the Head Start partner had participated in research over the last three years, and their involvement during this study year was no different than other years. They were aware that research was being conducted in the classrooms as part of the partnership, but they did not know about the Head Start Effectiveness sub-study in particular.

Child Assessments.

The child measures were administered by trained assessors to the treatment and control groups at three times: September-October, 1998; January-February, 1999; and late March-early May, 1999. The child outcome measures included the following instruments. The Peabody Picture Vocabulary Test-Third Edition (PPVT-III) (Dunn & Dunn, 1997) is a test designed to measure the child's receptive vocabulary. The PPVT-III scores have high reliability, with

internal-consistency (alpha) coefficients ranging from .92 to .98 and test-retest reliability ranging from .91 to .94 with a median of .95.

The M-KIDS Preliteracy Inventory (Nurss, 1995) is a preliteracy measure designed to assess print concepts, story retelling, and pre-writing skills in four to six year olds. In a national standardization sample (stratified by socio-economic status, race, and geographical distribution), the Preliteracy Inventory had Kuder-Richardson reliabilities ranging from .89 to .92. It also has a concurrent validity correlation of .62 with the Metropolitan Readiness Test composite score (Nurss, 1995).

The Early Phonemic Awareness Profile (Dickinson & Chaney, 1997) includes two composites: phoneme deletion, comprised of eight "judgment" and six "corrects" test items, and rhyming, comprised of eight "recognition" and five "production" test items. The Cronbach's alpha estimates from the total test items for two administrations in the fall (n=261) and the spring (n=241) are .95 and .95 for the total (Dickinson, 1999).

Parent Assessments

The parent measures were administered to both the Head Start and control groups in November-December, 1998, with a brief follow-up interview in the spring of 1999. Information about the child and family was collected using The Family and Children's Experiences Survey (FACES) Parent Interview (Administration on Children, Youth, and Families, 1997) which was initially developed by the Head Start Quality Research Consortium. The interview contains questions regarding household composition, demographic background variables, out-of-home care, services the child has received, satisfaction with Head Start services, parent involvement in the program, home learning activities, disabilities, the child's development, the transition to kindergarten, household routines, health and safety related issues, a broad range of home and

neighborhood characteristics, and caregiver depression, locus of control, and social support. The parents' rating of their children's social functioning as measured by the social skills and positive approach to learning scale and the problem behavior index was also included. The interview with the Head Start parent or primary caregiver is conducted over approximately one hour by a trained interviewer.

Predictions and Research Questions

It was hypothesized that the growth trajectory and post-test outcomes for the children and families in the Head Start condition would be different than that for the children and families in the control group condition. In this context, growth trajectory is defined as a statistical function that is fit to multiple measurements over time for each group. Differences in height and slope of the trajectories of each group were tested. The following specific research questions were addressed: 1) What is the growth rate or slope of the growth trajectory for each outcome measure for all children and for each group? 2) Are there differences between the groups in the growth rate of outcome variables? 3) Are there differences between the groups in pre/post-test gain scores for outcome measures?

Data Analysis

This study followed a specific analytical strategy, growth curve modeling, which is a special case of hierarchical linear modeling (HLM). A two level model was employed that treated change over time as nested within person in the level one model and person characteristics as predictor variables in the level two model (Bryk & Raudenbush, 1987). Three measurements were taken across various outcome measures enabling the researchers to estimate two parameters in the level one model, initial status and growth rate. A predicted growth trajectory was created for each individual, each group, and for the whole sample by estimating

the extent to which a time metric could predict performance on each outcome. The time metric in this case was the number of days since the beginning of the school year that had transpired at the time that each measurement took place. This has the effect of estimating a per day growth rate using a time metric that is centered on the beginning of the school year. This technique has the advantages of handling missing data, that is varying numbers of measurements across subjects, and the fact that the measurements may not have occurred at exactly the same time for each subject. As seen in Table 1, the mean number of days since the beginning of the school year that had transpired by the time of each measurement was not exactly the same for each group.

It is often unrealistic in social science research to make the assumptions of repeated measurements analysis of variance that considers each measurement occasion to occur at the same time for each subject. Even in the cases where this is achieved, the subjects may not all be the same age at the time of assessment. Age can be entered into models such as these as another form of time metric. In this case, the subjects were randomly assigned to groups which in theory would have the effect of equally distributing the range of ages across the groups, and as the analysis would indicate in Table 1, the mean ages of the groups were not different at the time of each assessment. In addition, we were interested primarily in the effect of the treatment and not of natural growth and development. Therefore, the time metric was chosen to represent the amount or duration of treatment received by the Head Start group. This technique assumes a linear growth pattern, as taking only three measurements does not allow for estimation of terms in the within-person growth model that might represent curvilinear growth patterns. The first measurement was taken approximately two months into the school year, and therefore, the initial

status and growth patterns within the first two months of the school year are predicted from the model only and can not be compared to actual data points.

First, we examined the estimation of the variance components for each outcome measure: total variance, within-person variance over time, and between-person variance. The differences in initial status between the groups were then examined followed by testing the between-group differences in growth rate. When there was not a statistically significant difference between the groups in initial status, only the growth rate parameter estimate was modified by the group membership variable (Head Start or control group), following the pattern illustrated in Figure 1. For the models that were created for measures that showed a between-group difference in initial status but no difference in growth rate, only the initial status parameter estimate was modified by the group variable, following the pattern illustrated in Figure 2. It should be noted that by including the group variable only in those level two models where it had a statistically significant relationship to the outcome variable, some pooling of variance accounted for could result. However, the final model for each outcome measure was chosen to best describe the patterns present in the data. A traditional analysis of variance was used to examine the social functioning of the children in the Head Start and control groups because only two measurements were obtained for this outcome. In addition, chi square analyses were used to examine differences in parent perceptions of health outcomes.

Figures 1 and 2 about here

Results

As seen in Table 1, the analyses included 82 children at Time 1, 85 children at Time 2, and 80 children at Time 3 for the Head Start treatment group. Initially at the beginning of the

school year, 87 children were assigned to the Head Start treatment group. The decreased number of children at the Time 1 assessment is due to parents not following through with the enrollment of their children. The increased number of children from Time 1 to Time 2 is due to three children migrating from the control group into Head Start before November 1. The attrition between Times 2 and 3 is due to the transfer of children to other centers or the family moving beyond the JCCEO service area. The control group initially had 86 children assigned and 60 children were assessed at Time 1. Attrition of the 26 children is due to: four children were accepted into Head Start before November 1; four children moved out of the JCCEO service area; five families did not have a working telephone number and family service workers were not able to contact them; six parents scheduled interview/assessment times and had repeated (two or more) "no shows"; four parents said that they did not have the time to participate in interviews/assessments; and three parents indicated they were not interested in participating. As seen in Table 1, the analyses included 60 children at Time 1, 42 children at Time 2, and 41 children at Time 3 for the control group. The research team was very persistent in trying to maintain the size of the control group. The 19 children and families who participated in previous assessments but not in the final assessment are accounted for as follows: three moved into Head Start before November 1 and became a part of the treatment group; nine moved into Head Start after November 1 and therefore, were dropped from the study; one child moved from one foster home to another outside of the JCCEO service area; one family participated only in the first assessment with 13 follow-up calls attempting to schedule another interview; three families participated in the first and second assessments with approximately 10 follow-up calls to each family seeking their continued participation; one family participated in the first and third

assessments with 11 follow-up calls; and one family participated in the first, second, and third assessments with seven follow-up calls.

Table 1 about here

The equivalence of the Head Start treatment and control groups with respect to family characteristics is portrayed in Table 2. The appropriate statistical tests were conducted to examine the differences between the proportions and means, and none indicated a statistically significant difference. The percentages of respondents to the parent interview who were in the mother, father and other relative categories are fairly similar across the Head Start and control groups and across time. There are slightly more fathers in the Head Start group reporting as the primary caregiver than in the control group. The percentages of respondents in the various categories of current marital status are similar in both groups and across time. The mean number of years for the respondents' educational level is almost identical across groups and time. The mean number of adults and children in the household in the two groups and across time are very similar.

Table 2 about here

For the measure of receptive vocabulary, the PPVT, the analysis of standard scores in Table 3 showed group equivalence in initial status, no growth over time for the control group, and a statistically significant faster rate of growth for the Head Start group ($p=.027$). This analysis is a test of the change in relative position as compared to norm samples. Figure 3 depicts the predicted growth patterns of the Head Start and control groups. As shown in Table 3, when the raw scores were analyzed, both groups were growing over time, as expected, while the Head

Start group was growing at a faster rate. This analysis was performed to make sure that the expected pattern of growth in the raw scores that could be attributed to natural growth and maturation was observed.

Table 3 about here

Figure 3 about here

In this same table a similar pattern of initial equivalence, all children growing, and a faster rate of growth for the Head Start children, was observed for the Phonemic Awareness total score. The growth rate for the Head Start children was statistically significantly faster (p=.035) than the control children. For the measure of print concepts, the model predicted Head Start children to be much higher at initial status and both groups grew at a parallel rate across the year as shown in Table 4. For the Print Concepts Total scale the growth rate for both groups was statistically significant (p=.000), however the two groups were significantly different at the initial status (p=.000).

Table 4 about here

Table 5 shows the correlation between initial status and growth rate for each outcome measure. These correlations ranged from near zero for the PPVT raw score to a fairly high negative correlation for the Print Concepts measure. Most of the correlations were modestly negative, as would be expected, indicating that the highest growth rates tended to be achieved by the children who started out lower at the beginning of the academic year. The majority of the variance (60%) in the scores for the PPVT and Phonemic Awareness measures, was found to be

between children. This between-child variance includes both within and between group differences but does not include within-child, or growth related, variance. The majority of the variance for the Print Concepts measure was found to be within children.

Table 5 about here

As reported in Table 6, a three-way analysis of variance with two between factors (treatment and gender) and one within factor (pre-post) was used to examine the parents' ratings of their child's social functioning. No significant interactions or main effects for the social skills and positive approach to learning (SSPAL) scale were found. There was a statistically significant time by group interaction ($F=(1,79)=6.143$) and a statistically significant main effect for time ($F=(1,79)=9.447$) on the problem behavior index (PBI). A simple main effects analysis indicated that for the Head Start group there was no difference between the fall and spring ratings of problem behaviors. However, for the control group there was a statistically significant difference between the fall and spring ratings with the mean rating in the spring being lower than the mean rating in the fall. This finding would indicate a reduction in perceived problem behaviors.

Table 6 about here

Health related outcomes by group are reported in Table 7. The parent report to all health questions about well care, health screenings, immunizations, and dental examinations showed statistically significant differences between the two groups. There was a statistically significant greater percentage of the Head Start parents indicating that they had addressed the specific health issues by the spring of the year when these data were collected.

Table 7 about here

It is also interesting to note that there are also differences in the parent report of health habits in the home. As seen in Table 8, there are no statistically significant differences between the two groups in common health habits such as tooth brushing, washing hands before meals, and exercising and staying fit. However, there is a statistically significant greater proportion of Head Start parents than control parents reporting that there is improvement in their children's habits of washing their hands after using the toilet, and eating nutritious and healthful foods. There is also a statistically significant greater proportion of Head Start parents than control parents reporting improvement in their own health habits of tooth brushing, eating nutritious and healthful foods, exercising and staying fit, and using seat belts more regularly.

Table 8 about here

Limitations

It is important to note that this study was not intended to represent an ideal strategy for evaluating the Head Start program as a whole. This study was conducted in the context of a close partnership with the Head Start program. The research design strategy was based, in part, upon decisions that were made within the partnership, intentionally to satisfy thoughtful input from program personnel. There are several unique features that this research introduced. First, the research team has worked in close partnership with this particular Head Start program for many years across several major research projects including the National Head Start/Public School Transition Demonstration Project and the Head Start Quality Research Consortium. There are long standing connections between our research team and this program including a

member of our staff who is housed in the Head Start program. Because of the previously established relationship, the type of challenges that would be faced by a researcher entering a set of randomly selected programs around the country for the first time may not be adequately represented in this study.

Second, the community that is served by this Head Start program contains several unique features. There is no publicly funded, state or local, pre-kindergarten program available to low income families, nor are there any viable alternatives to Head Start for low income families in this community. The community is very homogeneous and somewhat self-contained in a defined geographic area. The community is located within an almost entirely African-American, low income, inner city environment in one county in the southeast. Third, the particular Head Start program under study would be considered high quality, having NAEYC accredited classrooms, a comprehensive recruitment and penetration strategy, and a nationally recognized management team.

The decision was made to allow children randomly assigned to the control group to act as a waiting list, so they could migrate over to the Head Start group as slots became available in the classrooms under study. This decision was made based upon several considerations. First, it was our strong desire to minimize disruptions to the treatment under evaluation, in order to evaluate Head Start as it naturally occurs in a specific program. Migration from a waiting list is a typical part of the Head Start year in the program under study. Second, to not allow migration, that is to consider children once assigned to a group as always assigned to that group, would have the effect of changing the normal class size in the program. Under that strategy, once children leave the program or move out of the area their slots would not be filled, since there would not have been a waiting list. Given the well established connection between group size

and the quality of early interventions, such a strategy would have been fatally flawed in giving the Head Start group an advantage of a smaller group size than the program would normally maintain. Third, deliberately leaving Head Start slots unfilled would have been unacceptable to the program personnel who have a strong commitment to serving as many of the needs of the community as possible within the funding they are given. The migration did result in some loss of statistical power as some children migrated from the control group to the Head Start group until November 1. However, attrition is inevitable in field research of this kind, and it is difficult to speculate upon the effects that attrition would have had on the control group under any design conditions. It is possible that more subjects were retained in the control group than would have occurred under other design options. The possibility of moving into Head Start was a real possibility for those in the control group given their waiting list status and the random process which gave all children an equal chance to migrate into open classroom slots. It is important to note that class size could have been maintained at normal operating levels without allowing for migration from the control group. This would have required a randomly assigned holdout sample of children who could have been used to fill slots in the classrooms as attrition began to influence class size. However, this strategy would have required more children than the total number who registered with the program and therefore, was not considered a feasible option.

Finally, the quality of care for the control group children was not collected. As seen in Table 9, at the four data collection points approximately half of the children were in some other care arrangement. For those children who were in care, approximately one quarter were in care at the home of a family day care provider or relative and approximately three quarters of the children were in center-based care at private or non-profit child care or nursery school programs. The overwhelming majority of these children were in full-time care of 40 hours per week.

Therefore, the results of this study are limited by the researchers' inability to describe these alternative care settings.

Table 9 about here

Discussion

In summarizing the results, the initial status was equivalent and the growth rates for the Head Start children were statistically significantly faster than the control children on the receptive vocabulary, and phonemic awareness measures. For the Print Concepts measure, the initial status of the Head Start children was statistically significantly higher than the control children and the growth rate for both groups was statistically significant. The social functioning measure showed no significant interactions or main effects for the social skills and positive approach to learning scale, whereas there was a statistically significant time by group interaction and main effect of time for the problem behavior index. In regard to health outcomes, there was a statistically significant greater percentage of the Head Start parents indicating that they have addressed their children's well care, health screenings, immunization, and dental examinations. There is also a statistically significant greater proportion of Head Start parents than control parents reporting that there is improvement in some of their children's health habits and some of their own health habits.

It is important to note that the curious findings for the Print Concepts measure which showed differences in initial status may be due more to growth that occurs in the first two months of the school year than in true differences in initial status. The nature of this measure is such that it is conceivable that if a child is read to everyday for two months, this child would

easily begin to understand some of the basic concepts about books and the reading process. It is possible that the growth pattern in this measure is curvilinear with a period of acceleration early in the year, given the initial repeated and regular exposure to the reading process, and then levels off fairly quickly. The strategy of employing three measurements, delaying the first measurement until the second month of the school year, and assuming a linear growth trajectory, may simply have not allowed us to examine the actual pattern of the growth in this area of functioning. Given the consistent finding of equivalence in initial status on other variables, this seems to be a plausible explanation.

The positive impact of the Head Start treatment for this program as compared to the control condition is clearly evident on the measures of receptive vocabulary and phonemic awareness which showed a statistically significant faster rate of growth for the Head Start children. The results regarding the social functioning of the two groups are not surprising, given that the measure used in this study was the parent's rating of the child's behavior. There is little reason to suspect that the parents for the two groups would rate their children differently. However, it is interesting to note that the Head Start parents gave consistent ratings of problem behaviors, while the control group parents initially (in the fall) rated their children as having more problem behaviors. However, at the spring assessment the problem behavior ratings were very similar.

An important component of Head Start services is the provision of health services. Clearly, the results of this study suggest that Head Start continues to have a positive impact on preventive health outcomes. Participation in Head Start increases the likelihood that a child will receive appropriate health screenings, immunizations, and dental examinations. This finding is

similar to that found in the Effects of Head Start Health Services study by Fosburg and Brown (1984).

This research study not only offers some encouraging results regarding the effectiveness of the Head Start program, but does so in the context of an attempt to address some of the methodological shortcomings of previous studies. While the findings of this study are confined to a comparison of Head Start children with non-Head Start children within the specific context of a southern urban setting where relatively few publicly funded alternatives to Head Start exist, several important features of the study strengthen the results. This study used multiple outcome domains in an attempt to assess the comprehensive nature of the mission of the Head Start program. Multiple measurements over time were used to attempt to capture patterns of growth and a type of randomized design was used in assigning children to treatment and control conditions. Descriptive information about the quality of the alternative care settings for the control children is also needed. Further study is also needed to expand the use of randomization to include a holdout sample that could be used to maintain class size and avoid the process of migration to the treatment group. Future research is needed to replicate these findings in other contexts with a more nationally representative sample.

References

Administration on Children, Youth, and Families (1997). The Family and Child Experiences Survey. Washington, D.C.: Author.

Applied Management Services, Inc. (1978). Evaluation of the process of mainstreaming handicapped children into project Head Start, Phase II, Executive summary (ED168291). Washington, D.C.: Education Resource Information Center.

Bee, C. K. (1981). A longitudinal study to determine if Head Start has lasting effects on school achievement. Ann Arbor, MI: UMI Dissertation Services.

Bryant, D., Peisner-Feinberg, E., & Clifford, R. (1993). Evaluation of public preschool programs in North Carolina. Chapel Hill, NC: University of North Carolina.

Bryk, A. & Raudenbush, S. (1987). Application of hierarchical linear models to assessing change. Psychological Bulletin, 101(1), 147-158.

Currie, J. & Thomas, D. (1995). Does Head Start make a difference? The American Economic Review, 85 (3), 341-364.

Dickinson, D. (1999, February). E-mail: ddickins@edc.org. Message: EPAP psychometrics - internal consistency.

Dickinson, D. & Chaney, C. (1997). Early Phonemic Awareness Profile. Newton, MA: Education Development Center, Inc.

Dunn, L. & Dunn, L. (1997). Peabody Picture Vocabulary Test -Third Edition. Circle Pines, MN: American Guidance Service, Inc.

Esteban, M. (1987). A comparison of Head Start and non-Head Start reading readiness scores of low-income kindergarten children of Guam. Ann Arbor, MI: UMI Dissertation Services.

Fosburg, L. & Brown, B. (1984). The effects of Head Start health services: Executive summary of the Head Start health evaluation. Cambridge, MA: Abt Associates.

Head Start Amendments of 1998, 42 U.S.C. 9801 et seq. (1998).

Hebbeler, K. (1985). An old and a new question on the effects of early education for children from low income families. Educational Evaluation and Policy Analysis, 7 (3), 207-216.

Hunt, E. W. (1987). A comparison of the academic achievement of urban second grade pupils with different forms of public preschool experience. Ann Arbor, MI: UMI Dissertation Services.

Kaufman, A. & Kaufman, N. (1994). K-FAST: Kaufman functional academic skills test. Circle Pines, MN: American Guidance Service, Inc.

Kisker, E., Love, J., & Raikes, H. (1999). Leading the way: Characteristics and early experiences of selected Early Head Start programs. Washington, DC: U.S. Department of Health and Human Services

Lee, V. Brooks-Gunn, J., & Schnur, E. (1988). Does Head Start work? A 1-year follow-up comparison of disadvantaged children attending Head Start, no preschool, and other preschool programs. Developmental Psychology, 24 (2), 210-222.

McKey, R. H., Condelli, L., Ganson, H., Barrett, B., McConkey, C., Plantz, M. (1985). The impact of Head Start on children, families, and communities: Final report of the Head Start evaluation, synthesis, and utilization project. DHHS Publication No. (ODHS) 90-31193.

Nurss, J. (1995). Metropolitan Early Childhood Assessment Program: Pre-literacy Inventory. San Antonio: The Psychological Corporation.

Nystrom, P. (1988). A longitudinal study to determine the effects of Head Start participation on reading achievement in grades kindergarten through six in Troy public schools. Ann Arbor, MI: UMI Dissertation Services.

Reedy, Y. B. (1991). A comparison of long range effects of participation in Project Head Start and impact of three differing delivery models.

Roberts, J. H. (1984). A study of duration in Head Start and its impact on second graders' cognitive skills. Ann Arbor, MI: UMI Dissertation Services.

Roy Littlejohn Associates, Inc. (1985). Follow-up evaluation of the effects of mainstreaming hendicapped children in Head Start. Washington, D.C.

Sciarappa, Rogers, & Chamberlin (1994). Consumer Constructed Empowerment Scale (Making Decisions). Center for Psychiatric Rehabilitation, Boston University: Boston, MA.

Texas Instruments Foundations. (1996, June) Learning by leaps and bounds. Paper presented at Head Start's Third National Research Conference, Washington, D.C.

U.S. Department of Health and Human Services Advisory Committee on Head Start Research and Evaluation. (1999). Evaluating Head Start: A recommended framework for studying the impact of the Head Start program. Washington, D.C.: Author.

U.S. General Accounting Office. (1997). Head Start: Research provides little information on impact of current program. (GAO Publication No. GAO/HEHS-97-59).

U.S. Government Printing Office. (2000). National Head Start Impact Study. Federal Register. (RFP # 282-00-0022). Washington, D.C.: Author

Williams, M. T. (1988). Early childhood education intervention: An analysis of Nicholas County, Kentucky, Head Start program impacts from 1974 to 1986. Ann Arbor, MI: UMI Dissertation Services.

Zigler, E., Abelson, W., Trickett, P., & Seitz, V. (1982). Is an intervention program necessary in order to improve economically disadvantaged children's IQ scores? Child Development, 53, 340-348.

Table 1. Comparison of Time Related Variables By Group.

Group		Age			Days from First Day of School		
		Time 1	Time 2	Time 3	Time 1	Time 2	Time 3
Head Start	Mean	55.012	58.875	60.747	53.927	170.341	223.025
	Median	55	59	61	59	168	222
	Std. Dev.	3.560	3.354	3.402	13.452	14.394	14.803
	Minimum	47	48	50	31	151	193
	Maximum	69	65	66	109	195	251
Control	Mean	54.967	58.952	60.659	62.067	182.857	240.488
	Median	55	60	60	62.5	187	242
	Std. Dev.	3.701	3.920	4.060	16.598	12.635	13.982
	Minimum	41	44	46	37	165	222
	Maximum	61	65	67	97	209	271
	t	0.074	0.114	0.124	3.122	4.794	6.257
	p	0.941	0.909	0.901	0.002	0.000	0.000

Note. - Head Start n = 82(T1),85(T2),80(T3). Control n = 60(T1),42(T2),41(T3).

Table 2. Background Variables by Group.

Variable		Completed Fall Interview Head Start (n=66)	Completed Fall Interview Control (n=60)	Completed Fall and Spring Interview Head Start (n=47)	Completed Fall and Spring Interview Control (n=41)
Respondent's relationship to child					
(percent in each category)	mother	83.3	91.7	83.0	92.7
	father	9.1	1.7	8.5	2.4
	other relative	7.5	6.6	8.5	4.9
Respondent's current marital status					
(percent in each category)	single, never married	60.6	61.7	63.9	56.1
	married	19.7	20.0	10.6	24.4
	separated/divorced	18.2	16.7	23.4	19.5
	widowed	1.5	1.6	2.1	0.0
Respondent's education level					
(number of years)	mean	13.4	13.2	13.6	13.4
	standard deviation	1.5	1.7	1.5	1.8
Household composition					
number of adults (18 and over)	mean	1.7	1.5	1.7	1.5
	standard deviation	0.9	0.7	1.0	0.7
number of children (under 18)	mean	2.4	2.7	2.3	2.7
	standard deviation	1.1	1.3	1.2	1.4
Child's gender					
(percent in each category)	male	51.5	58.3	44.7	65.9
	female	48.5	41.7	55.3	34.1

Table 3. Predicted Mean Growth Trajectory by Group for Outcome Measures with no Difference in Initial Status.

Outcome	Group	Predicted Initial Status	Day 60	Day 170	Day 230	Predicted Monthly Growth Rate
PPVT St.Sc.	Head Start	78.426	80.225	83.525	85.324	0.900*
	Control	78.426	78.976	79.985	80.536	0.275
PPVT Raw Sc.	Head Start	32.206	36.524	44.441	48.759	2.159*
	Control	32.206	35.000	40.124	42.919	1.397***
Phonemic Awareness	Head Start	9.133	10.430	12.808	14.105	0.649*
	Control	9.133	9.856	11.182	11.906	0.362**

* p<.05, ** p<.01, *** p<.001

Table 4. Predicted Mean Growth Trajectory by Group for Outcome Measures with no Difference in Growth Rate but Differences in Initial Status.

Outcome	Group	Predicted Initial Status	Day 60	Day 170	Day 230	Predicted Monthly Growth Rate
Print Concepts	Head Start	15.932	17.502	20.381	21.952	0.785 *
	Control	10.068	11.638	14.517	16.087	

p = .000

Table 5. Additional Characteristics of Each Model.

Outcome	Proportion Variance Within Persons	Proportion Variance Between Persons	Within Persons Variance Reduction	Initial Status Reliability	Monthly Growth Rate Reliability	Correlation Between Growth And Init. Status
PPVT Standard Score	0.395	0.605	0.174	0.536	0.224	-0.280
PPVT Raw Score	0.399	0.601	0.387	0.509	0.242	0.022
Phonemic Awareness	0.395	0.605	0.433	0.704	0.493	-0.634

Table 6. Social Functioning by Group and Gender

Group	SSPAL Fall	SSPAL Spring	PBI Fall	PBI Spring
Head Start				
male (n=22)				
mean	12.14	12.50	3.59	3.50
sd	1.67	1.65	2.89	3.49
female (n=26)				
mean	12.46	12.12	4.00	3.73
sd	1.56	1.90	2.64	3.09
Grand Mean			3.70	3.60
Control				
male (n=24)				
mean	12.25	13.08	5.79	4.25
sd	1.87	0.93	2.50	2.13
female (n=11)				
mean	12.36	13.18	4.64	2.82
sd	2.16	1.4	2.29	1.83
Grand Mean			5.20	3.50

Note. - SSPAL: Social skills and positive approach to learning, PBI - Problem behavior index.

Table 7. Health Related Outcomes by Group (percent in each group responding "yes" to the question).

Question	Head Start	Control	Chi-sq.
In the last year, did you take child for well care/check-ups?	100.0	7.7	72.69***
Has child received the following health screenings?			
Lead	95.6	73.0	8.29**
Hematocrit	95.6	51.4	21.52***
Tuberculosis	88.9	68.4	5.30*
Growth Assessment	100.0	74.4	13.10***
Blood Pressure	97.8	79.5	7.31**
Hearing	100.0	82.1	9.00**
Vision	100.0	82.1	9.00**
Urine	100.0	64.9	19.16***
Do you have a record of child's immunizations?	100.0	10.3	70.18***
Is child up-to-date on immunizations/shots?	100.0	2.6	78.07***
Has child had a dental examination in the past year?	88.6	48.7	15.65***

*p<.05, **p<.01, ***p<.001

Table 8. Improvement in health habits over the past year by group (percent in each group responding "yes").

	Head Start	Control	Chi-sq.
Child health habits:			
Tooth Brushing	97.8	92.3	1.43
Washing hands before meals	100.0	97.4	1.19
Washing hands after using toilet	100.0	82.1	9.00**
Eating nutritious and healthful foods	87.0	61.5	7.33**
Exercising and staying fit	87.0	71.8	3.03
Parent health habits:			
Tooth brushing	97.8	76.9	8.88**
Eating nutritious and healthful foods	84.8	64.1	4.85*
Exercising and staying fit	80.4	43.6	12.37***
Using seat belts more regularly	95.7	74.4	7.89**
Improving safety in your home	97.8	87.2	3.54

*p<.05, **p<.01, ***p<.001

Table 9. Control group child care for those completing three assessments, including the first and last.

Question	Interview 1 (n=41)	Interview 2 (n=35)	Interview3 (n=39)	Interview 4 (n=41)
Is child in preschool or child care now? (percent in each category)				
yes	46.3	60.0	53.8	56.1
no	53.6	40.0	46.2	43.9
Where is care provided? (percent in each category for those indicating yes to above question)				
family day care/home	21.1	28.6	28.6	30.4
center-based care	78.9	62.0	66.7	65.1
another Head Start program		9.5	4.8	4.3
How many hours/week is care used? (for those responding yes to first question)				
mean	40.4	35.6	37.5	38.1
sd	7.2	10.9	13.4	5.4
minimum	25.0	8.0	7.0	20.0
maximum	50.0	50.0	80.0	40.0

Figure 1. Models Tested For Outcome Measures with no Difference in Initial Status, Shown In Equation Format.

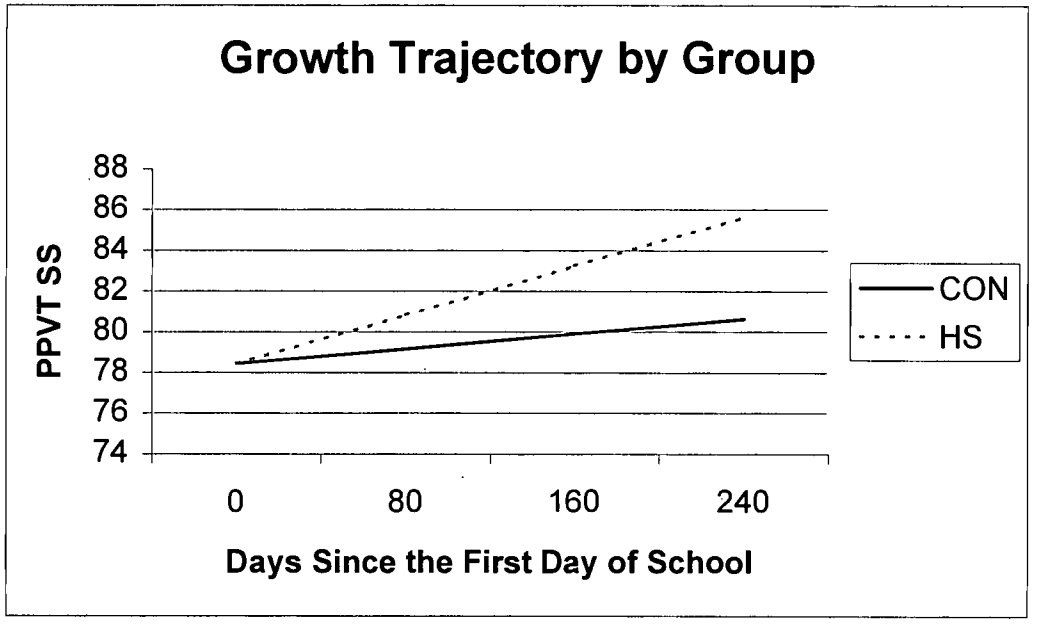
Within Person Model.	$Y_{ti} = \pi_{0i} + \pi_{1i}(DAY) + e_{ti}$
Initial Status Model.	$\pi_{0i} = \beta_{00} + r_{0i}$
Between Person Model.	$\pi_{1i} = \beta_{10} + \beta_{11}(GROUP) + r_{1i}$

Figure 2. Models Tested For Outcome Measures with Difference in Initial Status and No Difference in Growth Rate, Shown In Equation Format.

Within Person Model.	$Y_{ti} = \pi_{0i} + \pi_{1i}(DAY) + e_{ti}$
Initial Status Model.	$\pi_{0i} = \beta_{00} + \beta_{01}(GROUP) + r_{0i}$
Between Person Model.	$\pi_{1i} = \beta_{10} + r_{1i}$

Note. DAY is expressed as days since the first day of the Head Start school year.
GROUP is coded 0 for the Control group and 1 for the Head Start group.

Figure 3


Growth Trajectory by Group
PPVT SS
88
86
84
82
80
78
76
74
0
80
160
240
Days Since the First Day of School
CON
HS

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

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

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