

Child Outcomes Findings from the Head Start Impact Study and Family and Child Experiences Survey (FACES)

Chair: Jennifer Brooks

Presenters: Ronna Cook, Nicholas Zill

Discussants: Jim O'Brien, Gayle Cunningham

Cook: The National Head Start Impact Study is a longitudinal study that involves approximately 5,000 3- and 4-year-old preschool children across 84 nationally representative grantee/delegate agencies in communities where there are more eligible children and families than can be served by the program. The children participating were randomly assigned to either a treatment group (which had access to Head Start services) or a comparison group (which did not have access to Head Start services, but could receive other community resources).

The study follows a 3-year-old and 4-year-old cohort of newly entering children who had not experienced Early Head Start. One of the strengths of the study is a high response rate at close to 80 % of the initial sample throughout every round of data collection. Information is collected from a variety of sources including parent interviews, child assessments, and classroom observations.

Forty percent of the children are at home, which poses issues about collecting information that is comparable to those children who are in centers or day care homes. Information exists for children in day care homes, as well as in home-based situations with relatives or nonrelatives.

At baseline, there were no differences between the Head Start group and the non-Head Start group in terms of age, gender, race/ethnicity, parent language, or income eligibility. This sample has a large Hispanic population more evident in the 4-year-old than the 3-year-old sample. There were few differences between the Head Start group and the non-Head Start group on maternal characteristics, including age, marital status, race/ethnicity, and education level. A little more than one third of the mothers had no high school education, and approximately one third of mothers are employed full-time.

There were a number of statistically significant results for the Head Start children in cognitive assessments, including prereading, prewriting, and vocabulary for the 3-year-olds. No statistical significance was found for oral comprehension and phonological awareness or early math, showing areas that may need intensive attention. The impact of Head Start on reducing the achievement gap shows an almost 50% gap between the Head Start group and non-Head Start group on children's prereading, compared to the national norm.

In terms of spelling for 4-year-olds, the gap between the Head Start group and the national norm, compared to non-Head Start children, was reduced by about 28%. For vocabulary, the gap was only reduced about 8%. Results for the social-emotional domain are only based on parent reports at this time. Eventually teacher reports in kindergarten

and first grade will be added to broaden the view of social-emotional issues and approaches to learning, critical to understanding school readiness. Among the 3-year-olds, parents reported reduced behavior problems in terms of total behavior and in hyperactivity; however, no statistically significant impacts were detected on social skills, approaches to learning, or social competencies. No social-emotional impacts were found among 4-year-olds.

In the health domain, Head Start parents reported higher use of dental care for their children than non-Head Start parents. Head Start parents of the 3-year-old cohort also reported that the health status of their children was better.

In the parenting practices domain, Head Start parents read significantly more to their children than the non-Head Start parents. The 3-year-old Head Start group also reported positive impacts on involvement in enrichment activities and reduced use of physical discipline. No impacts for physical discipline or safety practices were found in the 4-year-old Head Start group.

In both age cohorts, Head Start children were twice as likely to have attended a center-based program as control group children. The control group children were five times more likely to have been exclusively in parent care, and Head Start children, as a beginning measure of stability, were more likely to be in the same setting both fall and spring than the non-Head Start children.

Early findings from the Head Start Impact Study show modest but statistically significant effects of the program on children's early literacy skills, health, and parental practices and perceptions. For the 3-year-olds, statistically significant findings were measured across all the domains, but not necessarily for every test within the domain. For 4-year-olds, impacts were measured for 3 out of the 4 domains. Impacts of the Head Start program on children's social-emotional development and on achievement in early elementary school are still to be determined. Head Start seems to meet a need for low-income families who want center-based care, as many of the control group children spent the year at home.

Zill: FACES is a mechanism for generating successive longitudinal findings about Head Start program performance. It is based on stratified national probability samples of the Head Start programs, families, and children. All children in Head Start have an equal probability of being in the survey; therefore, larger programs are somewhat overrepresented compared to other studies such as the Head Start Program Information Report (PIR), where all programs large and small have an equal probability.

FACES is not an experimental study of the impact of Head Start, and it has no randomized control group. However, some of the measures are compared to the national norms. Neither is FACES a study of interventions of a particular design to improve Head Start. Sample programs are not selected on the basis of curriculum or level of performance, but they are representative of what is happening in Head Start programs.

The findings from FACES are correlational and suggestive rather than definitive. There is no control group, but there is much to learn from cohorts over time. The data collection schedule to date of 3 cohorts was sampled in the years 1997, 2000, and 2003. Some of the data are now available for secondary analysis and the 1997 and 2000 year data are available in restricted public use files. Each cohort is an independent national sample of programs, families, and children, to enable comparisons over time. Children were followed from Head Start entrance until the end of kindergarten, although in 1997 they were followed to the end of first grade.

The child assessment measures include both nationally norm-referenced measures, publisher norms that are available to compare with the children in Head Start, and criterion-referenced measures that cover basic skills that children typically learn in preschool, such as phonological awareness, identifying colors, and counting. No general population norms are available for those measures.

The comparison strategy used in FACES is to use the cognitive assessment measures that have national norms, and to calculate standard scores for Head Start children based on these norms. The study then examines the extent to which Head Start children move toward national average scores during Head Start and kindergarten years. Some limitations exist for this comparison strategy. First of all, not all assessment measures used in FACES have national norms, and the published norms for the preschool children are often based on a small subsample of the general norming population. Some of these norms are relatively unstable when making comparisons within small age groups, and they also may be dated. Some evidence from an influx of immigrant children from relatively low-parent-education backgrounds shows that the norms of the past may be more rigorous than represents the general population of young children today. Therefore, some Head Start children may be compared with a standard that is no longer as timely. Another limitation is the comparison of FACES scores with preschool children in the United States, not necessarily with those from low-income families.

It is difficult to conduct FACES baseline assessments at the beginning of the program year. The roles of the Head Start programs have not stabilized. Head Start teachers are often not yet familiar with the children, which is important for the behavioral information collected. There has been some instruction happening before the baseline survey and, if that instruction has an effect, then the current strategy may miss gains in achievement that occur early in the Head Start year. However, those would not be missed in the Impact Study because the spring comparisons do indicate where children are in terms of growth.

In the area of prereading, FACES has found significant growth over cohorts in number of letters known by end of year. Growth was also seen in the number of letters known at baseline, which seems to indicate instructional gains in letter recognition early in the program year. Nevertheless, there still has been a significant change in the closing of the gap between Head Start children and the national norm in this prereading area.

One of the largest impacts in the Head Start Impact Study was in the area of parents' perceptions of children's early literacy skills, and that has also been a consistent finding

in FACES. A sizeable gain from fall to spring has been seen in what parents report their children can do in areas of letter recognition, color naming, early writing, and attempting to do writing.

In social skills, Head Start children show significant fall to spring gains in cooperative behavior and in social skilled behavior, as reported by teachers in all three cohorts of FACES. National norms are not available for teacher ratings of cooperative behavior and peer relationships, but ratings by the Head Start teachers are predictive of later ratings by elementary school teachers on the social adjustment of the children.

In terms of problem behavior, Head Start children showed small but significant fall to spring declines in withdrawn behavior in FACES 2003, a finding not seen in earlier cohorts. Head Start children also showed small but significant declines in hyperactive behavior in both FACES 1997 and 2000, but that finding was not replicated in 2003.

Head Start graduates showed continued significant gains toward national norms by the end of kindergarten in both FACES 1997 and 2000. The Impact Study should provide data from the kindergarten and first grade follow-ups and shed light as to whether these gains are attributable to Head Start. The FACES 2003 follow up includes broader, composite measures of reading skills, math skills, and general knowledge.

In summary, FACES shows gains in Head Start children reaching national norms, in the Head Start year and beyond, and in parent perceptions of children's early literacy skills. The gains have increased in some areas, notably letter recognition and pre-reading skills. The results seem to be related to more emphasis on these skills by programs. FACES has not found similar increases in the gains across the three cohorts of FACES for vocabulary and early math skills. These areas are important because the Early Childhood Longitudinal Study Kindergarten cohort (ECLS-K) indicates that vocabulary and math skills become increasingly important as children move from learn to read to reading to learn, as they try to relate what they are reading to their general knowledge about the world. FACES has found gains in social skills and some diminution of problem behavior, but comparative norms are not available for those measures. The Impact Study may provide information for some level of comparison to national norms in those areas.

O'Brien: The way FACES was developed through consultation with Head Start staff and partners has ensured the importance of the questions in both FACES and the Impact Study for the Administration, Office of Head Start, and most Head Start grantees. The interest in whether the findings close the gap or represent changes in children's or family growth after Head Start are reflected in the beliefs about program purpose.

When reading the study findings, one must consider the value of the way that the samples were collected in order to be generalizable to Head Start. A generalized study with an average of performance impacts or descriptions of the programs cannot illuminate much about the program at that level of intervention, but it does offer useful descriptive information. There were some surprises in terms of the range of needs for children and families, and most of the research goes beyond a child-focused range of family needs.

Another lesson learned is that when a clear message is communicated to programs about a goal and interventions, curricula, and classroom activities are addressed over several years, significant changes result in children's performance. This is especially effective when goals are broad enough to allow an area such as letter identification to include familiarity with books, a building block for reading success, a way to explore language and writing. The areas of vocabulary and numbers are not as discrete a skill, and it is more difficult to clearly describe the interventions, but many Head Start programs are exploring those types of interventions.

Finally, the policy level or interventions that the Administration for Children and Families (ACF) and the Department of Health and Human Services has evidenced over the past few years may have contributed to some of the growth described. Most likely, an important part of that is the Head Start Child Outcomes Framework that assesses children's progress. Other programs may not actively manage information, and that might be where our technical assistance efforts are best invested to move up this average and allow the highest performing programs to be imitated by other programs. It is also important to showcase evidence-based practices that are feasible within a Head Start program.

Cunningham: Programs ask what can be learned from these important studies that helps programs do a better job with teach children what they need to know. These two studies describe what Head Start programs should think about, do, and try. However, few Head Start program personnel attend research conferences. How can Head Start move this research to practice in ways that make a difference across Head Start programs across this nation?

Head Start programs respond when the message shared is clear. For example, a clear message was conveyed about children not learning alphabets, and now there has been a dramatic increase in the number of letters Head Start children recognize when they leave Head Start. The message about vocabulary and math is filtering through to programs, and measures are underway to teach these skills to children.

A small percentage of Head Start programs have found ways to take research on how children in their programs are doing, and about what is best for children, and use it for staff development and strategic planning. Head Start should put staff development in its policy requirements and its funding. This mechanism for interpreting research and evaluation information can help them learn how to interpret data generated from within their own programs.

If Head Start truly is the nation's child development laboratory, the gap between research and practice can be bridged by developing a real system that supports the interpretation of both research and local program data for the purpose of staff development.

Brooks: A question has been asked about whether children in the FACES 2000 data did meet national norms for early reading and math and whether this was true in 2003 at the end of the kindergarten year.

Zill: Preliminary analysis shows that is correct, but FACES 2003 data is still being collected on the second year of kindergarten. Head Start seems to be doing well in terms of early preparation, particularly the learning to read early skills.

It is challenging to assess phonological awareness in preschool-aged children, particularly those from low-income families, but the instruments used are workable and do produce reasonable results. In the area of math, the approach used in ECLS-K and also in the National Reporting System of putting together a relatively small group of items from several different domains can satisfy multiple parties.

Cook: It has been a challenge to get a grasp on phonological awareness, and the measures used have similar results. Many FACES measures were used to build the Head Start Impact Study for comparability of the study populations. It is important to evaluate which measures are best and also important to look at the breadth of findings, not just on individual tests.

Zill: The future is in computer-assisted adaptive assessment. This method uses a large bank of items covering several different domains tailored to the child. An item offers a baseline test of where the child is, and then moves to a second stage so that the child who is ahead of the Head Start norm receives more difficult items. The children who are behind receive easier items; therefore, a larger set of domains can be sampled while keeping the time relatively short.

Brooks: There is also a new research to practice work zone that is a collaboration between the Office of Planning, Research, and Evaluation, Office of Head Start, and providers of technical assistance. This collaboration looks at the full circle of what is known, how this knowledge is disseminated, and even how these needs are identified and researched. The outputs from this session will also be put online in the proceedings so that practitioners can see it from their own perspective.

Brooks: It is also essential to address use of local assessment data to encourage thinking about research in a more dynamic way instead of just doing large studies, putting out a report, producing a research brief, and expecting that all will be implemented as intended.