

## **Discussion of the Head Start Impact Study: Year 1 Findings**

**Chair:** Mary Bruce Webb

**Discussant:** Maggie J. Stevens

**Presenters:** Ronna Cook, Martha Zaslow

**Cook:** Several things make the Impact Study unique and are critical to its implementation. First, we were examining an ongoing program where people are passionate about its value. Therefore, building a partnership with the Head Start community was essential for full participation. Second, it is not only a randomized study, but nationally representative. Third, not only were grantees randomly selected, but also centers within the grantees, and children within each center. The children will be followed through the first grade. The data presented are from the first year (i.e., 8 or 9 months of intervention). About 15 % of the Head Start centers were eliminated from participation because they did not have enough children.

Outcomes were studied across four basic domains: cognitive, social/emotional, health, and parenting practices. Quality of services in relationship to outcomes was also studied. When the study was released to the public, debate arose around whether the effects found were in the correct areas, whether there were enough effects, and whether the sizes of the effects were large enough. To summarize, Head Start children had significantly higher early literacy skills than non-Head Start children. An interesting issue arose with the health domain. No differences were found between intervention and control groups in access to health services. However there were significant differences in access to dental care and Head Start parents were more likely to report positive health status for their children. In the parenting practices domain, Head Start parents reported that they were more likely to read to their children and less likely to use physical discipline, than non-Head Start parents. Also interesting is that children in Head Start were much more likely to be in quality preschool environments, i.e. higher scores on the ECERS), in particular, more frequent use of curriculum and better child-teacher relationship. Future analyses will show if what is occurring in Head Start has an impact on school readiness in kindergarten and Grade 1. So far, results indicate that Head Start is meeting the need for low-income families who want center-based care. This speaks to the continued effort of the Head Start community to utilize the Program Performance Standards and evidence-based practice.

**Zaslow:** The results of the Impact Study present several issues for the Head Start learning community. “Learning community” is defined as those researchers and practitioners that are involved in the mutual task of evaluating the effectiveness of Head Start. Together, they have a track record of looking at Head Start research, and other research on early childhood development more broadly, and using it meaningfully to guide practice. The body of Head Start research is growing, from the national descriptive picture of family and child experiences within Head Start programs (FACES), to the data from the National Reporting System and the research from the Quality Research Centers. The FACES data illustrate a system-wide response to the concerns about children’s early literacy development: learning from the initially flat or limited gain scores and tracking change over time system-wide in the extent of fall to spring gains. The Head Start Quality Research Centers serve as a vanguard for the kind of learning that the Head Start community engages in. This coordinated set of projects seeks to identify, pilot and evaluate

new strategies for improving quality in Head Start. A key feature of these projects is that researchers and practitioners work closely together on developing strategies for quality improvement. It is within this context that the findings of the Head Start Impact Study are viewed: what are the issues that it poses for the learning community to take stock of, to ponder, and to respond to?

There are several features of the Impact Study that need to be taken into account in any discussion of the results. These include: 1) the study has a nationally representative sample of programs are broadly implemented in a wide range of circumstances; 2) the results are based on about 9 months of intervention; and 3) the results reflect “intent to treat” rather than a contrast of exposure to Head Start vs. controls. A substantial minority of the control group found their way into other Head Start programs, while some of those families assigned to the Head Start program did not participate. As the report itself noted, these crossovers made it more difficult to find impacts, and the report is, in this sense, a conservative examination of the effects of the program. Further, a majority of the children in the control group, about 60, participated in some form of non-parental care. Thus, the control group was not a no-service control group.

Taking these features of the evaluation into account, what then are some of the issues posed by the first year results for the learning community? How do they pose challenges for practice, or for further examination in research? In harmony with the FACES results, the first year results show that the growing emphasis on children’s early literacy is being felt in children’s outcomes. However the range of outcomes affected thus far is not yet the full set of important early predictors of later literacy. The National Early Literacy Panel is carrying out an extensive review and synthesis of the research on early predictors of later literacy outcomes. One of the questions they are focusing on is: What are the skills and abilities in the preschool years that are linked to outcomes in reading, writing, and spelling from kindergarten forward? The findings point to a set of strong predictors of such later conventional literacy outcomes as decoding, reading comprehension, and spelling. Strong predictors include alphabetic knowledge, concepts about print, phonological awareness, and oral language ability. Impact Study findings thus far are heavily concentrated in the area of alphabetic knowledge with some more limited indications of program impacts on oral language ability. It is encouraging and important that the 3-year-old sample showed positive impacts on a direct assessment of receptive vocabulary and on a task involving naming colors. However, the children in the 4-year-old sample did not show these impacts. Direct assessments of some other key predictors of later literacy outcomes did not reveal program impacts. These include measures of phonological awareness and of oral comprehension. Thus, a key question for the learning community is whether the first year findings, taken together with the results from successive waves of FACES, point to a need to extend the work on early literacy to a greater focus on vocabulary, oral comprehension, and phonological awareness?

Marilou Hyson and Heather Biggar, in their chapter on the new NAEYC standards in the book *Critical Issues in Early Childhood Professional Development*, note that substantial progress has been made beyond the view that child-led and teacher-led instruction stand in opposition. A more complex perspective is emerging involving an integration of these approaches; one in which early childhood educators are seen as intentionally introducing content, but in ways that are appropriate for how young children learn. According to the new NAEYC standards, early

childhood educators need preparation not only in the processes of children's development, but also in academic content areas such as early literacy and early mathematics. Thus, a priority has been to work towards this integration in children's early literacy development.

Another important question is whether attention should focus on how to introduce early mathematics content more systematically into Head Start classrooms. The first year results did not detect statistically significant impacts in either the 3-or 4-year-old group on an assessment of children's proficiency at solving simple word problems that involved counting, simple arithmetic, and basic measurement. No significant impacts were found on a counting task examining children's understanding of one-to-one correspondence. Research by Herbert Ginsburg and his colleagues indicates that informal or everyday mathematics can be built upon through a set of organized activities intended to introduce mathematical ideas in ways that children enjoy. Their work closely follows a joint statement by the National Association for the Education of Young Children and the National Council of Teachers of Mathematics regarding specific topics that should be covered in working with young children. Thus, the question for the learning community is: should the first year results point in the direction of strengthening the focus on early mathematics as well as early literacy?

In terms of social/emotional outcomes, there have been mixed patterns of results over the years. Many studies have found that children who participated more in center-based care in the early years scored higher on measures of language and cognitive development at 2 and 3 years, and measures of language and memory at 4 ½ years, and were more socially engaged, but exhibited more conflict with peers. Results from Early Head Start, however, pointed to positive impacts on children's behavioral outcomes at 36 months along with positive cognitive impacts. Perhaps specific program characteristics underlie the pattern. Within this context, it is important to note that the first year results also show favorable impacts of the program on total behavior problems for children who entered the program as 3-year-olds, specifically on hyperactive behavior. Four-year-olds did not show the overall pattern of reduced behavior problems, but children from families with English language backgrounds did show significantly lower aggressive behavior scores. Together, the findings raise the key issue of what features of these programs underlie the reductions in problem behavior? The Head Start Impact Study included direct observations of classroom quality, carried out in Head Start and in center-based programs. These are not experimental analyses because they are restricted to those children in the sample who participated in formal early care and education settings. However, the results indicate differences, on average, in the patterns of social interaction in Head Start classrooms vs. other classroom settings. Specifically, teachers in Head Start classrooms were rated as more sensitive, more promoting of independence, and less harsh than teachers in other center programs. The question for the learning community is: What more specifically are Early Head Start and Head Start programs doing to help reduce problem behaviors in young children, and can this information help more broadly in supporting positive developmental outcomes across both the cognitive and social domains for children participating in formal early care and education programs?

Stepping back from the individual domains of development, a noteworthy pattern to study is the breadth of impacts. Impacts occurred across domains. A summary of the Impact Study prepared for the Society for Research in Child Development called particular attention to the breadth of

the first year impacts, placing findings in the context of other research indicating that children's school readiness rests on positive development across multiple domains of development. The confluence of positive impacts, particularly for 3-year-olds, is in itself an important pattern. This focus on a range of outcomes poses a challenge to the field: to retain the focus on multiple aspects of children's development rather than to report only on cognitive outcomes. As the disturbing mixture of cognitive and social outcomes summarized above indicates, early care and education experiences can have positive effects on one aspect of development while simultaneously having negative effects on another. Focus needs to be on multiple aspects of development in studying the effects of early care and education programs. The results also pose a challenge to the learning community in terms of understanding in greater detail for whom and why the positive impacts tended to cumulate. In the results thus far, why do we see impacts across more areas of development for 3-year-olds? Are there certain subgroups of children in which the pattern of breadth is especially likely to occur?

The issue of the size of effects in the Head Start Impact Study has tended to be discussed in terms of the cup being half full or half empty. The issue needs to be reframed. Those focusing on the cup being half empty note that the effect sizes in the study are modest. They range from .10 (in the 3-year-old group, number of times the child was spanked), to .34 (again in the 3-year-old group, for two parent-report measures: parent report of the child's literacy skills and parent report that the child had dental care). The report adopts the convention that effect sizes of less than .2 are small, effects sizes between .2 and .5 are moderate, and those over .5 are large. Following this convention, most of the statistically significant impacts that the study detected were small in magnitude, though about 1/3 were moderate in size.

Those focusing on the cup being half full note that the impacts documented were uniformly positive; that these impacts were documented only about 9 months after enrollment in the evaluation, in a program that many children participated in only part-day; and that, as we have noted, the program was a widely implemented public program rather than a tightly controlled and small demonstration project. They note further that the control group was not a no-service control, but that a majority had some form of non-parental childcare, including in some instances, another Head Start program.

The report itself begins to move away from looking at effect sizes from the half full—half empty perspective, to the “when fuller” perspective in the reporting of subgroup analyses. Some effect sizes were larger within key subgroups (for example, effect sizes for the Woodcock-Johnson III Letter-Word Identification and PELS for the African-American subgroup). One of the most important results that yet to be reported is looking at the effects in light of program quality. A challenge for the learning community will be to carry out these analyses, identifying the quality criteria to be used to distinguish among Head Start programs. Should this be the Early Childhood Environment Rating Scale-Revised total score? Would the expectation be that certain aspects of quality are more closely related to outcomes in particular domains of development? In an extremely thought-provoking article, Kathleen McCartney and Robert Rosenthal note the importance not only of the magnitude of an effect, but also its practical importance. The work may not be over when the effect size is labeled as small, moderate, or large: Implications that effect sizes of any magnitude have in real world context need to be examined. One hypothesis for the possible practical importance of small effects during early childhood was proposed over a

decade ago. In 1995, Doris Entwisle asked the question of how the relatively small effects of early childhood intervention programs on assessments of children's intelligence documented to that point, effects that faded out over time, nevertheless appeared to have lasting effects on such outcomes as retention in grade and referral for special education services. There is the possibility that effects that are small or moderate in magnitude can have practical importance through altering the elementary school teacher's perceptions of the appropriate ability grouping for the child. For the learning community, it will be important to follow up on the first year Head Start impacts by exploring whether and how findings of the magnitude documented, in fact, have practical importance through such decisions.

It is an explicit goal of Head Start to support the development of children from a range of different backgrounds, respecting their home cultures. The descriptive statistics about the analysis sample in the Impact Study make it clear how often Head Start programs are facing the challenge of supporting the development of children who are learning English. English was the primary home language for about 70 % of the children in the 3-year-old sample, but just over 60 % in the 4-year-old sample. And while about 23 % of the fall child assessments were completed in Spanish (with some children then going on to complete the assessments in spring in English), the parallel figure for the 4-year-old sample was nearly a third of the sample. Fewer positive impacts on child outcomes were found for children whose primary language was Spanish than those whose primary language was English in both age groups, but the pattern was particularly striking for 4-year-olds. In this age group, while positive impacts were found in all domains of development for children whose primary language was English, for children whose primary language was Spanish, impacts were found only in the area of health. A key question for the learning community is what approaches best support the development of children whose primary language is Spanish or another language other than English? We have much to learn in this area, from choice of measure, to how to help children remain connected to their home culture and language as they master English, to what factors and experiences predict the best initial but also longer-term academic progress in this group of children.

Putting Early Head Start findings together with the pattern of positive impacts in the first year of the Head Start Impact Study opens the lens to what precedes participation in Head Start. Would effects be stronger or more pervasive if children and families participated in Early Head Start from birth to 3 and then Head Start from 3 to 5? The Early Head Start impact study points to delayed or sleeper effects for the highest risk families, surfacing only at the end of the follow-up period. It may take sustained work with particularly high-risk families to show effects. Would the combination of Early Head Start and Head Start be beneficial for high-risk families? For all Head Start-eligible families?

Although the learning community was defined as comprising Head Start researchers and program providers, it is clear that the issues raised by the first year results of the Head Start Impact Study have implications that are important for all those interested in early childhood practice and research.

**Stevens:** how does the Head Start community of practitioners perceive the Head Start Impact Study? How do the results fit in with what programs are currently doing in their day-to-day work? What is helpful for developing next steps? How is research in general helpful to Head

Start practitioners? Research often gets released through the media and the nuances are lost. For example, recent research on health concerns of caffeine use reveals that it will increase cholesterol, but protect against Parkinson's. People make their decision based on what is right for them. Sometimes as parents, we use the folk methods that are tried and true, that have been passed down from our family. So sometimes in the Head Start community, decisions are made on the bases of what people are comfortable with and know. Additionally, how the results of research or evaluation are framed will make a difference. For example, it is disheartening to a teacher, who has spent an entire year working with children to get them ready for kindergarten, to learn that there were "small to moderate gains" in child cognitive outcomes. However, the Society for Research in Child Development Report took the same results and reported that Head Start had narrowed the gap in specific aspects of pre-reading abilities between children in poverty and U.S. children as a whole by 45 %. That sounds a lot better, even if the results are the same. Therefore, how research is translated to the community—the Head Start community and the community at large—is incredibly important.

The Head Start community believes that research is a useful tool. It is a tool that helps inform practice. However, we have also learned to be extraordinarily cautious. Oftentimes it is because of how the media makes quick sound bites out of information that is usually more complex.

The inclusion of "crossovers" is going to be something that the Impact Study researchers are going to need to be able to clearly explain to the Head Start practitioner. As soon as the control group has families who have found their way into a Head Start program regardless of their being eliminated from the "Head Start intervention" group, then random-assignment is called into question along with the results.

Head Start practitioners can whole-heartedly support the results on children's phonological awareness. These results point to the need for professional development and changes in classroom practice.

At this point in time, programs have a wealth of information at their disposal: the Program Information Report, child outcome data, self-assessment data, and the NRS information. All of this information had at one time been sent only to others. More and more Head Start programs realize that this is their data. They make a difference to their program, and these data are now being used in a more thoughtful and analytical ways. In addition, a growing numbers of programs have added people to their staff that have a variety of titles, such as outcome managers, accountability managers, monitoring experts, and quality assurance managers. Their work developed from looking at child outcome data and program data, finding strengths, finding weaknesses, and helping to improve the quality of the program. It is more frequently to inform classroom practice, to inform program decision-making, and to drive professional development.

Staff asks questions like: how do parents view their health services and how we can improve that? Would home visits by our family and community staff that focused on shared reading techniques improve literacy scores? Learning more about how research may be used has empowered staff to explore new ways of thinking about their jobs and the families they serve.

Research results become more important to them, and much less about just going through the motions of data collection for someone else.

Finding common ground is important. Researchers and practitioners suffer from time lags. Researchers collect data, then they have to analyze the data, and then they have to publish the results. Oftentimes practitioners get impatient. But practitioners have time lags too. While they recognize the importance of research results, they need time to recognize its usefulness and find its relevancy to their program.

The missing piece seems to be the connection and the communication between the research community and the Head Start community. We have some of it in place now, but we need better translators. This conference has helped enormously, but the number of Head Start practitioners at these conferences is still relatively small. The National Head Start Association added a research track at their annual conference. That has been well attended by practitioners. The Society for Research in Child Development Reports help translate large reports into bite-sized, understandable, digestible pieces. The ACF's website that has the annotated bibliography gives programs a place where they know they can find all research results. The Head Start Quality Research Centers are closely tied to programs. The Head Start -- University Partnership Grants and the Early Head Start research grants have helped to develop understanding between researchers and practitioners. Some new players that have helped to translate research into practice are the statewide Head Start associations that have formed research committees, such as in Florida, California, and Kentucky. These groups do most of their communication through conference calls. As learning communities, they support each other and learn from each other. For example, Florida holds an annual research conference where the topic is always research to practice, helping to interpret research for the Head Start community in Florida. Media is also developed. For example, for the past 4 years, the Florida association has published a booklet called "The Florida Portrait of Children's Outcomes," available on [floridaheadstart.org](http://floridaheadstart.org). The challenge is for researchers to contact programs and associations and other learning communities within their states to start these conversations and to help develop a common language to support a common goal. Both the research community and the Head Start practitioner community need to be knowledgeable about research and what works best for children, and also what supports the highest quality service for young children and their families.