

Issues for the Learning Community

Martha Zaslow, PhD

The Head Start program has a history of using research to guide its efforts. In this way, Head Start practitioners, policy makers, and researchers comprise a “learning community.” The release of the first-year results of the Head Start Impact Study provides an important opportunity for reflection by the learning community, and for identification of strategies to further strengthen the program. This article begins by illustrating how Head Start has responded to previous research. It then underscores some key features of the Head Start Impact Study that set it apart from other important evaluation studies. Within this framework, the article goes on to identify a set of key issues that the first-year results of the Head Start Impact Study pose for the learning community to consider in charting the future course of the program as well as to others focusing on strengthening young children’s school readiness.

THE Head Start program has a history of using research to guide its efforts. In this way, Head Start practitioners and researchers comprise a “learning community.” The release of the first-year results of the Head Start Impact Study (U.S. Department of Health and Human Services, 2005) provides an important opportunity for reflection by the learning community and for identification of strategies to further strengthen the program.

This article begins by illustrating how Head Start has responded to previous research. It

then underscores some key features of the Head Start Impact Study that set it apart from other important evaluation studies. Within this framework, the article goes on to identify a set of key issues that the first-year results of the Head Start Impact Study pose for the learning community to consider in charting the future course of the program.

A LEARNING COMMUNITY

The body of Head Start research is multifaceted, including multiple “strands” in addition to the Head Start Impact Study (U.S. Department of Health and Human Services, 2005) and the Early Head Start Impact Study (Love et al., 2005; U.S. Department of Health and Human Services, 2002, 2006). These strands include the national descriptive picture of family and child experiences within the Head Start program (the Family and Child Experiences Survey [FACES]; U.S. Department of Health and Human Services, 2003), the data from the National Reporting System, the research of the Quality Research Centers, and a range of further studies focusing on specific issues.

While the focus of this article is on the Head Start Impact Study, findings from the FACES research (U.S. Department of Health and Human Services, 2003) provide a striking illustration

Author Affiliation: *Child Trends, Washington, District of Columbia.*

This article is based on a presentation given at a plenary session at Head Start’s Eighth National Research Meeting on June 27, 2006.

The author thanks her colleagues, Kristin Moore, Tamara Halle, Kathryn Tout, and Laura Lippman at Child Trends for their helpful feedback on early drafts of the presentation and the paper. The author is very grateful to Jessica Vick, graduate student at the University of Maryland, and research intern at Child Trends, for her help in revising the article in light of reviewer comments. The author also thanks the editor of this special edition and the anonymous reviewer for the many helpful suggestions they provided for clarifying and strengthening the article.

Corresponding Author: *Martha Zaslow, PhD, Child Trends, 4301 Connecticut Ave NW, Suite 350, Washington, DC 20008 (mzaslow@childtrends.org).*

of Head Start as a learning community. The first cohort of FACES data, collected in 1997–1998, showed no or only small gains from fall to spring on some key aspects of early literacy development. For example, children showed no significant gains from fall to spring of the year in which they participated in Head Start on knowledge of book and print conventions. The findings were taken very seriously, and Head Start put into place a systemwide focus on children’s early literacy development. The 2000 wave of FACES data provided an opportunity to ask whether there were any systemwide improvements. This subsequent wave of data showed significant fall to spring gains in children’s knowledge of book and print conventions where there had been none earlier. At the same time there was also a significantly larger gain from fall to spring in the later cohort than in the earlier cohort in children’s letter recognition.

The FACES data illustrate a systemwide response to the concerns about children’s early literacy development: learning from the initially flat or limited gain scores and tracking change over time systemwide in the extent of fall to spring gains. At the same time, the 2000 FACES data provided a cautionary note: In 2000, there was no greater improvement from fall to spring than in the earlier cohort on children’s scores on an assessment of vocabulary. That is, the systemwide change had penetrated to some measures of children’s early literacy, but not to others.

Turning to another strand of Head Start research, the Head Start Quality Research Centers have served as a vanguard for the kind of learning that the Head Start community engages in (Abbott-Shim, 2006). This coordinated set of projects has sought to identify, pilot, and evaluate new strategies for improving quality in Head Start. It is a key feature of these projects that researchers and practitioners have worked closely on developing strategies for quality improvement. Approaches to improving quality have been piloted in selected sites and then replicated in further sites, and there is coordinated data col-

lection across the multiple intervention and evaluation studies.

In sum, the Head Start community has a track record of absorbing messages from its research, including cautionary ones, piloting and testing the efficacy of approaches to strengthening practices within the program, and carrying out ongoing documentation of whether or not there are systemwide responses to new approaches. It is within this context that this article asks: what are the issues that the results from the Head Start Impact Study pose for the learning community to ponder, and to which to respond, at this point in time?

KEY FEATURES OF THE EVALUATION TO KEEP IN MIND IN CONSIDERING THE RESULTS

There are some features of the Head Start Impact Study that need to be taken into account in any discussion of the results. Before turning to a discussion of the issues posed by the results of the Head Start Impact Study, it is important to note first these distinctive features of the evaluation.

By way of background, the Head Start Impact Study was congressionally mandated. It assesses the impact of Head Start across 84 nationally representative agencies, and includes a total of 4667 newly entering children (2559 three-year-olds and 2108 four-year-olds) whose families applied for Head Start. Children were randomly assigned to a group that had access to Head Start services, or a group that could take advantage of whatever non-Head Start community resources were available to them. Statistical analyses revealed that there were no significant differences between children randomly selected to be in the treatment and control groups. Data collection, including interviews with parents, direct child assessments, ratings of children by care providers, surveys of teachers, interviews with center directors, and direct observations of quality, began in fall 2002 (US DHHS, 2005). Children will be followed through the third grade. A report on impact

results 2 years after children and families enrolled in the study is anticipated shortly.

The Head Start Impact Study is one of an increasing number of experimental evaluations focusing on the effects of programs and policies on children. Yet some features of the study set it apart. Among other recent and ongoing experimental evaluations are a set of studies considering the impact on children's development of differing approaches to reforming welfare policy (Hamilton et al., 2001; McGroder, Zaslow, Moore, & LeMenestrel, 2000; McGroder, Zaslow, Moore, & Brooks, 2006; Zaslow, McGroder, & Moore, 2000; Tout et al., 2004), and experimental evaluations focusing on the impacts on children of differing program approaches to strengthening the quality of early care and education settings (such as the Quality Interventions for Early Care and Education, or QUINCE evaluation, which is evaluating on-site consultation approaches as a means of strengthening quality in home-based as well as center-based child care; Bryant & Wesley, 2006).

While the Head Start Impact Study shares with these other evaluations the central feature of experimental design, permitting strong causal attributions to assignment to the program or intervention group as the source of effects on children, a feature that sets the Head Start Impact Study apart is that it involves not only random assignment of participants but also selection of a representative sample of program sites, permitting generalization to all Head Start programs that met the sampling requirements (including the requirement of not having enough spaces for all those who applied), and giving the study strong external as well as internal validity. It is a distinctive feature of the study that results can be seen as generalizing to Head Start programs of the kinds sampled. While, as noted, there are a number of other experimental evaluations of early childhood programs now in progress, it may be argued that this study is the largest in terms of the number of programs and children to which its results will generalize.

In relation to this feature, it is very important to acknowledge that the Head Start Im-

pact Study is a study of the impacts of a program as it was broadly implemented in a wide range of circumstances. This is not an evaluation of a small, tightly controlled demonstration program with uniform high quality.

It is also critical to acknowledge that the first-year impact study reflects outcomes for children only about 9 months after enrollment. Only children who were newly entering Head Start were randomly assigned to groups within the study. There has been a tendency to juxtapose the findings of the Head Start Impact Study first-year results with those of such studies as the Abecedarian Project evaluation (Campbell & Ramey, 1995; Campbell, Ramey, Pungello, Sparlling, & Miller-Johnson, 2002; Ramey et al., 2000). Yet this was a multiyear intensive intervention that started in infancy. Subsequent findings will provide more information about the impact of 1- to 2-year enrollment in Head Start on children's outcomes.

As others have noted (eg, Ludwig & Phillips, 2007), it is also important to take into account that the impact results reflect "intent to treat" rather than a contrast of exposure to Head Start versus not. Families were randomly assigned to be able to enroll their child in a particular program or not to be able to enroll the child. But it was not feasible or ethical to prevent families from seeking out other Head Start programs if random assignment resulted in the child not being eligible to enroll in a particular program, nor were families forced to participate in Head Start once random assignment resulted in their having access to a program. 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 notes, 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. Furthermore, as Ronna Cook has pointed out (June 2006) and the Impact Study report makes clear, a majority of the children in the control group, about 60%, participated in some form of nonparental care. Thus, the

control group was not a no-service control group.

In sum, the Head Start Impact Study provides not only an experimental evaluation but also data that can reflect on the program nationally. This is an examination of a public program implemented in a wide range of circumstances and with varying quality, rather than a small and tightly controlled demonstration. The first-year results reflect about 9 months of participation rather than multiple years of participation. And finally, the results are a conservative test of the effects of participation in Head Start in the pure “intent to treat” analyses. These are all important factors to keep in mind in interpreting the findings.

ISSUES FOR THE LEARNING COMMUNITY

Taking these features of the evaluation into account, what then are some of the issues posed by the first-year results? How do they pose challenges for practice, or for further examination in research?

Extending the literacy focus: From 10 letters to a thousand words?

In harmony with the FACES results summarized above, the first-year results of the Head Start Impact Study point to the conclusion that the growing emphasis of the Head Start program 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 (2006) 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, phonologic awareness, and oral language ability (Westberg & Molfese, 2005).

Findings from the Head Start Impact Study first-year results (U.S. Department of Health and Human Services, 2005) show favorable impacts for some but not all of these important early predictors of later literacy. The findings thus far are heavily concentrated in the area of alphabetic knowledge with some more limited indications of program impacts on oral language ability. Children who entered Head Start as 3-year-olds as well as children who entered Head Start as 4-year-olds were more advanced at the end of the first year of the evaluation than children in the control group on the Woodcock Johnson-III letter-word identification task (effect size for 3-year-olds = 0.24, for 4-year-olds = 0.22) and in a letter naming task (effect size for 3-year-olds = 0.19, for 4-year-olds = 0.24).*

It is encouraging and important that the 3-year-old sample also showed positive impacts on a direct assessment of receptive vocabulary ($d = 0.12$), and on a task involving naming colors ($d = 0.10$). However, the children in the 4-year-old sample did not show these impacts in receptive vocabulary or naming of colors. Direct assessments of some of the further key predictors of later literacy outcomes did not reveal program impacts. These include measures of phonologic awareness (Comprehensive Test of Print and Phonological Processing) and of oral comprehension (Woodcock Johnson-III Oral Comprehension Test). Jean Layzer, in a presentation at the Brookings Institution (August 1, 2005), asked whether Head Start has shown impacts on the skills that will help with an early and fundamental literacy task: that of decoding written language through recognizing letters and understanding their correspondence with sounds; but few or no impacts on those skills that will surface as of greater importance later on related to comprehension of text.

*Note that the Woodcock-Johnson letter-word identification task goes beyond the issue of letter recognition.

A key question for the learning community is whether the first-year findings of the Head Start Impact Study, 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 phonologic awareness? N. Zill (personal communication, May 9, 2005) posed the issue in this way: Should Head Start progress from a focus on mastery of 10 letters to mastery of 1000 words?

As noted earlier, the Quality Research Centers have focused on testing approaches to strengthening quality in Head Start. Two of the Quality Research Centers have had a specific focus on early literacy development. Efforts to introduce literacy-focused curricula in the work of Fischel and colleagues (Fischel & Bracken, 2006; Fischel, Spira, Shaller, Fuchs-Eisenberg, Katz & Storch, 2004), and literacy-focused in-service training in the work of Clark-Chiarelli, Dickinson, and colleagues (Clark-Chiarelli, Dickinson, Bolte, & Buteau, 2004; Clark-Chiarelli, Grospen, & Bolte, 2006) are being rigorously evaluated in these centers. There is also work focusing specifically on strengthening vocabulary growth within Head Start programs, for example, the studies by Wasik and colleagues (June 2006). These studies will provide important guidance should the learning community determine that a broader range of literacy outcomes should be targeted.

And what about numbers?

Marilou Hyson and Heather Biggar, in a chapter on the revised NAEYC standards in the book *Critical Issues in Early Childhood Professional Development* (Hyson & Biggar, 2006), 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. For example, they may structure the classroom environment so that children can explore specific content in their

play. They note that the revised NAEYC standards build in the view that 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.

The first priority thus far has been to work toward this integration in children's early literacy development. The first-year results of the Head Start Impact Study raise the important question of whether attention should also be focused 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-on-one correspondence. Mean scores on both measures were higher for Head Start children, but differences were not statistically significant (U.S. Department of Health and Human Services, 2005).

Several research groups have been deepening our understanding of how young children develop early mathematical skills, and whether and how it is possible to strengthen these skills. In one such group, Herbert Ginsburg and his colleagues at Columbia Teachers College note that young children "have a spontaneous interest in mathematical ideas" (2006, p. 174). For example, young children enjoy saying counting words, determining which block tower is higher, and developing and extending patterns. Their research indicates that this 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 (April 2002) regarding specific topics that should be covered in working with young children.

Positive effects on children's knowledge and skills have now been found in 2 evaluation

studies involving providing professional development for Head Start teachers in early mathematics (Arnold, Fisher, Doctoroff, & Dobbs, 2002; Sophian, 2004). Interestingly, in 1 of these studies, the researchers found increases in child self-report of interest in math. In addition, teachers in the intervention group had higher scores at posttest on measures of liking for and sense of competence in teaching math. These measures of positive engagement by both teachers and children appear to be very important complements to the evidence on increased understanding of early mathematics in the children.

With these encouraging results, the question for the learning community is: Should the first-year results of the Head Start Impact Study point in the direction of strengthening the focus on early mathematics as well as early literacy?

Shifting the focus to social outcomes

A number of years ago, the researchers of the NICHD Study of Early Child Care and Youth Development presented evidence of a mixed pattern of results regarding the outcomes over time of children's participation in child care during their early years: Controlling for other aspects of care (including quality), children who had participated more in center-based care in the early years scored higher on measures of language and cognitive development at 2 and 3 years (NICHD, 2000) and measures of language and memory at 4¹/₂ years (NICHD 2002a). Children who spent more time in center care were found not only to be more socially engaged but also to have more conflict with peers. Again controlling for other aspects of child care participation, including quality, children with more extensive exposure to child care over their first years of life were reported by both mothers and teachers to show more problematic social behavior, including more acting-out aggressive behavior (NICHD, 2002b). A more recent publication by this research group indicates that teacher-reported behavior problems remain slightly elevated through sixth grade among children with greater early expo-

sure to center-based child care (Belsky et al., 2007).*

Recent work by Tamara Halle and colleagues extends the examination of this pattern to a nationally representative sample. This research group has found a combination of more advanced reading achievement, but higher scores on measures of problem behavior, for children who had participated in center-based care in the year prior to kindergarten in the Early Childhood Longitudinal Study-Kindergarten Cohort data (Halle & Hair, 2006).†

Yet in 2003, John Love and a group of colleagues engaged in the Early Head Start Impact Study (along with other studies) raised the critical question of whether this pattern reflected on all forms of formal early care and education (Love et al., 2003). This group noted that Early Head Start overall, including its center-based and mixed-center- and home-based approaches, showed positive impacts on children's behavioral outcomes at 36 months along with positive cognitive impacts. This group questioned whether formal center-based settings in and of themselves are predictive of the mixed pattern, or whether formal settings with specific characteristics underlie the pattern.

The most recent data from the Early Head Start Impact Study continue to raise this compelling question. In a follow-up study, developmental outcomes for children who had been assigned to participate in the Early Head Start program or a control group have been tracked through the preschool years. When all children are considered irrespective of whether they had been assigned to the Early Head Start or control group, those who participated in formal programs after age 3 showed

*In contrast with the findings at earlier follow-up points, however, in sixth grade, vocabulary scores were predicted by early child care quality rather than by type.

†The Early Childhood Longitudinal Study-Kindergarten Cohort has data on extent of participation in child care in the year prior to kindergarten as well as on participation in different types of child care, but lacks data on child care quality.

the mixture: they had not only better early reading skills but also more aggressive behavior. But in a very important new finding, the follow-up study indicates that children who *had* been in Early Head Start and then proceeded to participate in formal settings after age 3 showed the enhanced early reading skills, but did not show the increase in aggressive behavior. That is, Early Head Start participation protected against the pattern noted now in multiple studies of elevated behavior problems following more extensive participation in formal settings (U.S. Department of Health and Human Services, 2006).

Within this context, it is very important to note that the first-year results of the Head Start Impact Study also show favorable impacts of the program on total behavior problems for children who entered the program as 3-year-olds ($d = -0.13$; ie, a reduction in total behavior problems), and specifically a reduction in their hyperactive behavior ($d = -0.18$). 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.

So here as for the Early Head Start Impact Study, children on average showed reductions in reported problem behaviors while also showing more advanced scores on some measures of cognitive development. Together, the findings of the Early Head Start and Head Start Impact Studies 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. But they do provide important indications that there are differences, on average, in the patterns of social interaction in Head Start classrooms and other classroom settings that children in the study sample participated in. For example, results from the Arnett (1989) measure of caregiver sensitivity indicate that for 3-year-olds, teach-

ers in Head Start classrooms received higher scores than teachers in center programs on ratings of sensitivity (eg, speaking warmly to the children and listening attentively when they speak) and promoting of independence (eg, allowing children to follow their own agenda, valuing children's input). For 4-year-olds, Head Start teachers were rated as less harsh and again as promoting more independence on average. Clearly there is a range of quality in both Head Start and centers. But these findings raise the possibility that in the Head Start Impact Study sample, there may be differences on average in the supportiveness of caregiver-child interaction for those attending Head Start and those attending other formal early care and education settings.

While the Head Start Impact Study included "state of the art" observational measures of quality, we are currently engaged in a period of examination of our measures of quality (Roundtable on Measuring Quality in Early Childhood and School-Age Settings, 2006). Multiple measures have been developed recently, looking at more specific aspects of stimulation in early childhood settings (see Halle & Vick, 2007, for a compendium of early childhood measures of quality). Just as the field has been working toward more in-depth measures of early language and literacy stimulation in early care and education settings (as in the Early Language and Literacy Classroom Observation; Smith, Dickinson, Sangeorge, & Anastasopoulos, 2002), there may also be a need to develop observational measures focusing in greater depth on the guidance provided to children on social interactions that may help children in Head Start and other programs handle difficult emotions and social situations.*

The question here for the learning community is: What more specifically are Early Head Start and Head Start programs doing to help

*One such measure, Supports for Social-Emotional Growth Assessment (SSEGA; Smith, 2004) for Early Childhood Programs Serving Preschool-Age Children, is under development at New York University by Sheila Smith and colleagues.

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?

Breadth of impacts

Thus far this article has focused on impacts in selected outcome areas. But stepping back from the individual domains of development, a noteworthy pattern is the breadth of impacts in the first-year results of the Head Start Impact Study. This is especially apparent in the results for children who entered the program as 3-year-olds. For this age group, there were favorable impacts on prereading skills (Letter-Word Identification, $d = 0.24$; Letter Naming, $d = 0.19$), prewriting skills ($d = 0.13$), vocabulary (PPVT-III Adapted, $d = 0.12$; Color Naming, $d = 0.10$), (reduced) parent-reported problem behaviors ($d = -0.13$), and parent report of child health status ($d = 0.12$). There were also program impacts on parent participation with the child in enrichment activities (the number of times children read to, $d = 0.18$, Family Cultural Enrichment Scale, $d = 0.11$) and children having greater access to dental healthcare ($d = 0.34$).

Here again, juxtaposing the impact findings from Early Head Start and Head Start, the breadth of impacts in Early Head Start is also noteworthy. At 36 months, after participation in Early Head Start, children had higher Bayley Mental Development Scores, higher receptive vocabulary scores, and lower parental ratings of aggressive behavior. The home environments of families participating in Early Head Start were observed to be more supportive and stimulating, parents were observed to be more supportive in interacting with their children, and parents reported reading to their children more often and spanking them less often. The children were observed to be more engaged when playing with a parent, and to show more sustained attention to objects during play (Love et al., 2005; U.S. Department of Health and Human Services, 2002).

The summary of the Head Start Impact Study prepared by Hiro Yoshikawa on behalf of the Society for Research in Child Development (Yoshikawa, 2005) called particular attention to the breadth of the first-year Head Start Impact Study impacts. This summary places the findings in the context of other research indicating that children's school readiness rests on positive development across multiple domains of development. Findings summarized in *From Neurons to Neighborhoods*, reporting on the conclusions of the National Academy of Sciences Committee on Integrating the Science of Early Childhood Development (Shonkoff & Phillips, 2000), indicate that doing well across multiple aspects of development in the early years predicts better later adjustment to and progress in school. The confluence of positive results in the Head Start Impact Study, particularly for 3-year-olds, mirroring the findings from the Early Head Start Impact Study, is in itself an important pattern.

This focus on a range of developmental outcomes for children in the Head Start Impact Study as well as the Early Head Start Impact Study 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 mixture of cognitive and social outcomes summarized above from the NICHD Study of Early Child Care and Youth Development indicates, early care and education experiences can have positive effects on one aspect of development while simultaneously having negative effects on another. We need to continue to focus 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? For example, does this reflect on the responsiveness of younger children to the effects of early childhood programs, on differences in families and children who enrolled in Head

Start earlier as opposed to later, or perhaps on wider discrepancies in the early care and education experiences between the program and control groups for 3-year-olds than for 4-year-olds, given the greater access 4-year-olds in the control group may have had to prekindergarten programs? Are there certain subgroups of children in which the pattern of breadth is especially likely to occur? As we follow this sample over time, do we see particular implications for subgroups of children who showed positive outcomes early across a range of developmental outcomes as opposed to in a more limited set?

The issue of size of effects

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. Laura Lippman (personal communication, June 24, 2006) from Child Trends has suggested that we rephrase the question to be "When is the cup fuller?"

Those focusing on the cup being half empty note that the statistically significant 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: 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

the conventions adopted in the report, most of the statistically significant impacts that the study detected were small in magnitude, although about a third were moderate in size.

Those focusing on the cup being half full note that the impacts documented were uniformly positive; that these impacts were found 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 child care, including in some instances, Head Start.

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 (eg, the effect size for the Woodcock-Johnson III Letter-Word Identification task with African American children).

Some of the most important work focusing on the question of when the glass is fuller remains to be done. One of the most important steps that has not yet been reported on (and that is in progress) is a set of analyses looking at impacts for children in light of program quality. A related set of analyses in the Early Head Start Impact Study focused on implementation of the Early Head Start performance standards (Love et al., 2005; U.S. Department of Health and Human Services, 2006). Three groups were identified: those that fully implemented the performance standards early on, those that implemented them fully but later in the evaluation, and those that implemented them incompletely. All 3 groups had favorable impacts on child outcomes, but the 2 groups with full implementation, either early or later, had more impacts on children. A challenge for the learning community will be to carry out parallel analyses in the Head Start Impact Study, identifying the implementation or quality criteria to be used to

*The author thanks the reviewers of this article for suggesting these alternative interpretations.

[†]Although these interpretations differ from Cohen's (1988) conventions (<.2 = trivial, .2-.5 = small, .5-.8 = moderate, >.8 = large), the study's authors (U.S. DHHS, 2005) reason that the findings represent the impact of participation over a short period of time, some children randomly assigned to be eligible for a Head Start program did not go on to participate while some of those assigned not to be eligible for a Head Start program ended up participating in another one, and impact estimates should be viewed within the context of the effect sizes found in other studies of other early childhood and educational research.

distinguish among Head Start programs. Should this be the total score on the Early Childhood Environment Rating Scale-Revised (Harms, Clifford, & Cryer, 1998)? Or, 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, McCartney and Rosenthal (2000) 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: We may need to continue to examine the implications that effect sizes of any magnitude have in the real-world contexts in which children spend time.

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 (Entwisle, 1995). On the basis of her research into the transition to formal schooling of low-income children in Baltimore, she hypothesized that even a small impact of an early childhood intervention program on a child's IQ could result in a teacher in the first grade tracking a child into a higher reading group. Such groups are formed in the large majority of elementary schools, and are generally based on the teacher's perception of the child's relative standing within the class rather than on any more absolute or objective measure. According to Entwisle, such early ability-grouping decisions can have lasting implications. Children in higher groups are taught more words, and stories are read to them with more intonation and emotion. Children in higher reading groups tend to get higher grades, which, in turn, tends to encourage their self-concept as learners and their subsequent learning. Further research by Entwisle and her col-

leagues has continued to provide evidence that "the 'sorting and selecting' early in middle childhood sets the stage for placements that follow. . . . First grade track assignments predict test scores and dropout in middle school" (Entwisle, Alexander, & Olson, 2006, p. 192).

Thus, there is the possibility that effects that are small or moderate in magnitude can have practical importance because of the cumulative influence of small or moderate effects occurring across multiple aspects of development, and through altering the elementary school teacher's perceptions of the appropriate ability grouping for the child. For the learning community, it will be very 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.

Supporting the development of children learning English

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 Head Start 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 more than 60% in the 4-year-old sample. And, while about 23% of the fall child assessments were completed in Spanish for the 3-year-old sample (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 for 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 that of 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 the issue of which are the strongest measures for assessing Spanish-speaking children (Barrueco & Lopez, 2007), 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.

Juxtaposing Early Head Start and Head Start Impact Study Results

As noted in earlier sections of this article, it is exciting to be able to put side by side the first-year results of the Head Start Impact Study with the recently released findings from the follow-up of the Early Head Start Impact Study through the end of the prekindergarten year. While juxtaposing the findings in this way is informative, it is nevertheless important to acknowledge the methodological differences between the studies and the degree of implementation of the programs. The Head Start Impact Study has a sample that is representative of the program, whereas the Early Head Start Impact Study sample, while quite dispersed geographically, was not designed to be a representative sample of programs or families. The Head Start program reaches many more children than the Early Head Start program, and the latter was still fairly new at the time that evaluation study was initiated, with some sites still not fully implementing program standards. Yet both studies implemented experimental designs and focused on multiple domains of children's development.

The Early Head Start follow-up study indicates that the strongest pattern of positive out-

comes occurred when children participated in Early Head Start from birth to age 3, and then followed up with participation in a formal early care and education setting from age 3 to 5 (U.S. Department of Health and Human Services, April, 2006).

Putting these findings together with the pattern of positive impacts in the first year of the Head Start Impact Study opens up the lens to what precedes participation in Head Start. The questions posed to the learning community are thus: Would effects be stronger or more pervasive if children and families participated in Early Head Start from birth to 3 years and then Head Start from 3 to 5 years? 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?

At the other end of the age continuum, we are fortunate that the findings summarized in this article for the Head Start Impact Study are only the first round of results, and that there will be follow-up data as the children enter school and progress through third grade. Analyses of the follow-up data will make it possible to consider what happens when children in Head Start go on to experience a range of subsequent environments, and how experiences in the preschool years combine with those in the early years of elementary school to shape development.

CONCLUSION

In conclusion, while the introduction to this article identified the learning community as composed of Head Start program providers, policy makers, and researchers, 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, policy, and research. The Head Start Impact Study findings inform

and pose critical issues to a learning community with broad boundaries: the commu-

nity of those concerned with strengthening children's early development.

REFERENCES

- Abbott-Shim, M. S. (Chair). (2006, June). *Enhancing school readiness: The Head Start Quality Research Center studies (2001-2006)*. Poster symposium at Head Start's Eighth National Research Meeting, Washington, DC.
- Arnett, J. (1989). Caregivers in day-care centers: Does training matter? *Journal of Applied Developmental Psychology, 10*, 541-552.
- Arnold, D. H., Fisher, P. H., Doctoroff, G. L., & Dobbs, J. (2002). Accelerating math development in Head Start classrooms. *Journal of Educational Psychology, 94*, 762-770.
- Barrueco, S., & Lopez, M. L. (2007, March). *Measurement and young Spanish-speaking children: A psychometric analysis of published linguistic and literacy assessments*. Paper presented at the meetings of the Society for Research in Child Development, Boston, MA.
- Belsky, J., Vandell, D. L., Burchinal, M., Clarke-Stewart, K. A., McCartney, K., Tresch Owen, M., and The NICHD Early Child Care Research Network. (2007). Are there long-term effects of early child care? *Child Development, 78*, 681-701.
- Bryant, D. M., & Wesley, P. (2006, June). *The consultant's role in an early care and education quality enhancement intervention*. Paper presented at Head Start's Eighth National Research Meeting, Washington, DC.
- Campbell, F. A., & Ramey, C. (1995). Cognitive and school outcomes for high risk African American students at middle adolescence: Positive effects of early intervention. *American Educational Research Journal, 32*, 743-772.
- Campbell, F. A., Ramey, C. T., Pungello, E. P., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science, 6*, 42-57.
- Clark-Chiarelli, N., Dickinson, D. K., Bolte, G., & Buteau, E. (2004, June). *The impact of program delivered literacy in-service training (PD-LIT) on teacher practice and preschool children's achievement*. Presentation as part of poster symposium at Head Start's Sixth National Research Meeting, Washington, DC.
- Clark-Chiarelli, N., Grospen, J., & Bolte, B. (2006, June). *Program delivered professional development: Is it effective?* Presentation as part of poster symposium at Head Start's Eighth National Research Meeting, Washington, DC.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Earlbaum.
- Cook, R. (2006, June). *Plenary session on the Head Start Impact Study first year results*. Presentation at Head Start's Eighth National Research Meeting, Washington, DC.
- Entwisle, D. R. (1995). The role of schools in sustaining early childhood program benefits. *Future of Children, 5*, 133-144.
- Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2006). Educational tracking within and between schools. In A. C. Huston & M. N. Ripke (Eds.), *Middle childhood: Contexts of development* (pp. 173-197). Cambridge: Cambridge University Press.
- Fischel, J. E., & Bracken, S. S. (2006). *Curriculum approaches to enhance early literacy in Head Start*. Poster symposium at Head Start's Eighth National Research Meeting, Washington, DC.
- Fischel, J. E., Spira, E. G., Shaller, G. E., Fuchs-Eisenberg, A., Katz, S., & Storch, S. A. (2004, June). *The enhancement of emergent literacy skills in Head Start: Outcomes of classroom curriculum research*. Poster symposium at Head Start's Sixth National Research Meeting, Washington, DC.
- Ginsburg, H. P., Goldberg, R., Kaplan, J. C., Cordero, M. I., Eisenband, J. G., Galanter, M., et al. (2006). Helping early childhood educators to teach mathematics. In M. Zaslow & I. Martinez-Beck (Eds.), *Critical issues in early childhood professional development* (pp. 171-202). Baltimore: Brookes.
- Halle, T., & Hair, E. (2006, June). *The effect of type and extent of child care on low-income children's outcomes in kindergarten, grade 1 and grade 3*. Head Start's Eighth National Research Meeting, Washington, DC.
- Halle, T., & Vick, J. E. (2007). *Quality in early childhood care and education settings: A compendium of measures*. Washington, DC: Child Trends for the Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Hamilton, G., Freedman, S., Gennetian, L., Michalopoulos, C., Walter, J., Adams-Ciardullo, D., et al. (2001). *National Evaluation of Welfare-to-work Strategies: how effective are different welfare-to-work approaches? Five-year adult and child impacts for eleven programs*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families and Office of the Assistant Secretary for Planning and Evaluation; and U.S. Department of Education. Retrieved August 7, 2007, from the U.S. Department of Health and Human Services Web site: <http://aspe.hhs.gov/hsp/newws/index.htm>

- Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early Childhood Environment Rating Scale, Revised edition*. New York: Teachers College Press.
- Hyson, M., & Biggar, H. (2006). NAEYC's standards for early childhood professional preparation: Getting from here to there. In M. Zaslow & I. Martinez-Beck (Eds.), *Critical issues in early childhood professional development* (pp. 283-308). Baltimore: Brookes.
- Layzer, J. (2005, August 1). *Comments on the first year results of the Head Start Impact Study*. Presentation at the Brookings Institution Seminar on the National Head Start Impact Study, Washington, DC.
- Love, J. M., Harrison, L., Sagi-Schwartz, A., van Ijzendoorn, M. H., Ross, C., Ungerer, J. A., et al. (2003). Child care quality matters: How conclusions may vary with context. *Child Development, 74*, 1021-1033.
- Love, J. M., Kisker, E. E., Ross, C., Raikes, H., Constantine, J., Boller, K., et al. (2005). The effectiveness of Early Head Start for 3-year-old children and their parents. *Developmental Psychology, 41*, 885-901.
- Ludwig, J., & Phillips, D. (2007). The benefits and costs of Head Start. *Social Policy Report, 21*, 3-17.
- McCartney, K., & Rosenthal, R. (2000). Effect size, practical importance, and social policy for children. *Child Development, 71*, 173-180.
- McGroder, S. M., Zaslow, M., Moore, K. A., & Brooks, J. L. (2006). Mandatory welfare-to-work programs and preschool-age children: Do impacts persist into middle childhood? In A. C. Huston & M. N. Ripke (Eds.), *Middle childhood: Contexts of development* (pp. 350-369). Cambridge: Cambridge University Press.
- McGroder, S., Zaslow, M. J., Moore, K. M., & LeMestrel, S. (2000). *The National Evaluation of Welfare-to-Work Strategies: Impacts on young children and their families two years after enrollment: Findings from the Child Outcomes Study*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families and Office of the Assistant Secretary for Planning and Evaluation; and U.S. Department of Education. Retrieved August 7, 2007, from the U.S. Department of Health and Human Services Web site: <http://aspe.dhhs.gov/hsp/NEWS/child-outcomes/summary.htm>
- National Association for the Education of Young Children and the National Council for Teachers of Mathematics. (2002, April). *Position statement on early childhood mathematics: Promoting good beginnings*. Washington, DC: National Association for the Education of Young Children. Retrieved on August 8, 2007, from the NAEYC Web site: <http://www.naeyc.org/about/positions/pdf/psmath.pdf>
- National Early Literacy Panel. (2006). *Synthesizing the scientific research on development of early literacy in young children*. Washington, DC: National Institute for Literacy. Retrieved August 7, 2007, from the National Institute for Literacy Web site: <http://www.nifl.gov/partnershipforreading/family/ncfl/NELP2006Conference.pdf>
- National Institute of Child Health and Human Development Early Child Care Research Network. (2000). The relation of child care to cognitive and language development. *Child Development, 71*, 958-978.
- National Institute of Child Health and Human Development Early Child Care Research Network. (2002a). Early child care and children's development prior to school entry. *American Educational Research Journal, 39*, 133-164.
- National Institute of Child Health and Human Development Early Child Care Research Network. (2002b). Child care and children's peer interaction at 24 and 36 months. *Child Development, 72*, 1478-1500.
- Ramey, C. T., Campbell, F. A., Burchinal, M., Skinner, M. L., Gardner, D. M., & Ramey, S. L. (2000). Persistent effects of early intervention on high-risk children and their mothers. *Applied Developmental Science, 4*, 2-14.
- Roundtable on Measuring Quality in Early Childhood and School-Age Settings: At the Junction of Research, Policy and Practice (2006, December). *Meeting summary prepared by Child Trends*. Sponsored by the Child Care Bureau, U.S. Department of Health and Human Services, Washington, DC. Retrieved August 14, 2007, from the Child Care and Early Education Research Connections Web site: <http://www.researchconnections.org/location/ccrca12621>
- Shonkoff, J., & Phillips, D. A. (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press.
- Smith, S. (2004). *Pilot of a new classroom assessment instrument: Supports for social-emotional growth* [unpublished study]. New York: Child and Family Policy Center, Steinhart School of Education, Culture, and Human Development, New York University.
- Smith, M., Dickinson, D. K., Sangeorge, A., & Anastopoulos, L. (2002). *Early Language and Literacy Classroom Observation (ELLCO) toolkit*. Baltimore: Brookes.
- Sophian, C. (2004). Mathematics for the future: Developing a Head Start curriculum to support mathematics learning. *Early Childhood Research Quarterly, 19*, 59-81.
- Tout, K., Brooks, J., Zaslow, M., Redd, Z., Moore, K., McGarvey, A., et al. (2004). *Welfare reform and children: A synthesis of impacts in five states. The Project on State-Level Child Outcomes*. Washington, DC: Administration for Children and Families and the Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. Retrieved August 7, 2007, from the Administration for Children and Families Web site: http://www.acf.hhs.gov/programs/opre/welfare_

- employ/ch_outcomes/reports/welfare_reform_children/wel_ref_child.pdf
- U.S. Department of Health and Human Services, Administration for Children and Families. (2002). *Making a difference in the lives of infants and toddlers and their families: The impacts of Early Head Start*. Washington, DC: Author. Retrieved August 7, 2007, from the Mathematica Web site: <http://www.mathematica-mpr.com/PDFs/ehsfinalsumm.pdf>
- U.S. Department of Health and Human Services, Administration for Children and Families. (2003, May). *Head Start FACES 2000: A whole-child perspective on program performance*. Fourth Progress Report. Washington, DC: Author. Retrieved August 7, 2007, from the Administration for Children and Families Web site: http://www.acf.hhs.gov/programs/opre/hs/faces/reports/faces00_4thprogress/faces00_4thprogress.pdf
- U.S. Department of Health and Human Services, Administration for Children and Families (2005, May). *Head Start Impact Study: First year findings*. Washington, DC: Author. Retrieved August 7, 2007, from the Administration for Children and Families Web site: http://www.acf.hhs.gov/programs/opre/hs/impact_study/reports/first_yr_finds/first_yr_finds.pdf
- U.S. Department of Health and Human Services, Administration for Children and Families (2006, April). *Preliminary findings from the Early Head Start pre-kindergarten follow-up*. Washington, DC: Author. Retrieved August 7, 2007, from the Administration for Children and Families Web site: http://www.acf.hhs.gov/programs/opre/ehs/ehs_resrch/reports/prekindergarten_followup/prekindergarten_followup.html
- Wasik, B. (2006, June). *Implementation and evaluation of a collaborative language and literacy training and assessment program for Head Start*. Presentation at Head Start's Eighth National Research Meeting, Washington, DC.
- Westberg, L., & Molfese, V. (2005, May). *Overview of the work of the National Early Literacy Panel*. Presentation at the meeting of the Literacy Researchers of Color Consortium, Washington, DC. Retrieved August 7, 2007, from the National Institute for Literacy Web site: <http://www.nifl.gov/partnershipforreading/family/ncfl/NELP2006Conference.pdf>
- Yoshikawa, H. (2005). *Placing the first-year findings of the National Head Start Impact Study in context*. Brief prepared for the Society for Research in Child Development. Retrieved August 14, 2007, from the Society for Research in Child Development Web site: <http://srcd.org/documents/policy/Impactstudy.pdf>
- Zaslow, M. J., McGroder, S. M., & Moore, K. A. (2000). *Summary report: Findings from the child outcomes study of the National Evaluation of Welfare-to-work Strategies two years after enrollment*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families and Office of the Assistant Secretary for Planning and Evaluation; and U.S. Department of Education. Retrieved August 7, 2007, from the U.S. Department of Health and Human Services Web site: <http://aspe.dhhs.gov/hsp/NEWWS/child-outcomes/summary.htm>