The release of the very brief (kindergarten and first grade) longitudinal follow-up to the National Head Start Impact Study has renewed questioning about the effectiveness of this 45-year-old program that has served over 25 million poor children and their families. Head Start’s friends and foes alike do not appear to fully recognize what this study is and what it is not. Contrary to popular belief, the results were never meant to provide a definitive answer to the long-standing question of whether Head Start works. And while the random assignment design has been heralded as better than all Head Start research that has gone before, it turns out that this design actually confounded understanding of the study’s results.

Head Start’s many critics have interpreted this study as being a classic medical random assignment design in which the treatment group receives the medicine and the control group receives a placebo. This design characterized evaluations of older experimental programs such as the Perry Preschool that compared participating children’s progress to that of children who stayed home until they reached school age. These programs were compensatory interventions meant to offset the negative results of spending the critical early years in deprived circumstances. Today, relatively few poor children stay at home until they are old enough to enter kindergarten or first grade. The majority of states offer preschool classes to at least some at-risk children. Further, welfare assistance is not as easily available as it was decade or so ago. Poor mothers are in the workforce and their children are in substitute care. Head Start is no longer the only game in town.
If the National Impact Study is not a classic evaluation of Head Start vs. no intervention or an assumed negative environment, what exactly is it? The authors make clear that the study is a comparison of Head Start with a wide variety of other formal programs as well as parent care. Unfortunately, comparing Head Start with “everything else” can tell us very little. It would be fair to predict that Head Start’s impact on school readiness is probably superior to home rearing, as many studies have shown, but not markedly better than state preschool programs. This is an empirical issue worthy of further study. However, the Impact Study sheds little light on the matter. The value of the study would have been greatly enhanced if the control group was not treated as a hodge-podge of “everything else” but was divided into subgroups including parent care, state preschool programs, and child care. Instead, the evaluators erroneously assert that the evaluation addresses the “overall average impact of the program” (U. S. Department of Health and Human Services, 2010, p. xv). This would only be true if every single program (or non-program) the control children received had absolutely no impact. This assumption is erroneous on its face.

The Impact Study is obviously not an effectiveness study of Head Start but is rather a value-added comparison, as the author admits. In essence, the study is meant to assess whether Head Start has greater value than that of a variety of other treatments. Brooks-Gunn argues that this design of comparing one treatment to a collection of other treatments “is the likely explanation for the lack of sustained effects in the National Impact Study,” noting that this type of comparison “in general has not shown impacts” (2010). (See also Camilli, Vargas, Ryan, & Barnett, 2010, on this same point.) She assumes as I do that the 60% of the children in the control group who were in programs experienced adequate quality. The Abecedarian investigators reported that children in their control group who attended quality child care
programs showed better scores on preschool measures of intellectual development than those reared mostly at home (Burchinal, Lee, & Ramey, 1989). A model program that used a broader value-added design was the Chicago Child-Parent Centers, which makes its findings more impressive than those of the Perry Preschool, where the control group had no intervention at all.

Complicating the matter further is that the groups in the Impact Study were badly contaminated. The intent-to-treat design included a relatively large number of children in the Head Start group who never attended the program (17% of the 3 year olds and 14% of the 4 year olds). The intended design for the control group was that children could experience any intervention but Head Start in the first year (although the 3-year-olds controls could enroll in Head Start in the second year, which many did). However, many of their parents voted with their feet and enrolled their children in another Head Start center that was not part of the study. Thus 18% of the 3 year olds and 14% of the 4 year olds in the supposed control group actually attended Head Start. Those implementing the study attempted to correct for both the crossovers in the control group and the no-shows in the Head Start group through statistical procedures, never a totally satisfactory course of action. For example, parents motivated enough to hike across town to enroll their child in another Head Start center might be much more committed to their child’s education.

Another factor that may have contributed to group contamination has received no attention. The author served on the oversight committee for the National Impact Study, a position that gave him the opportunity to speak directly with the contractors. He learned that they were hampered by having fewer children on the waiting lists than had been anticipated. (Wait-listed children were assigned to the control group.) One obvious reason is that the Head Start recruiters were understandably reluctant to beat the bushes to recruit more families than
they knew they could serve. The contractors also complained that some Head Start directors were finding ways to circumvent the random assignment. We must remember that Head Start providers do not serve as scientists but real people with real needs. For decades Head Start has been charged with enrolling the children and families with the highest risks. This mandate may have lead to a higher proportion of children at high risk being in the treatment group than was planned and a resulting “less needy” control group who were placed on the wait lists.

Countering this is Russ Whitehurst’s view that only the better Head Start centers were used in the study, so the children who attended should have had more robust outcomes even if they were at higher risk (personal communication, April 8, 2010). He reasons that only Head Start centers with expected sizable waiting lists were used, and better centers have a greater likelihood of having long waiting lists.

Differences within and between the Head Start and control groups are just one reason why the longitudinal follow-up of the Impact Study should not be treated as a test of the value of Head Start. The program’s value should only be assessed against the goals that have been set for it. Since its inception Head Start’s purpose has been to prepare poor children for school. Over the years scientists, policymakers, and the public have developed unreasonable expectations that Head Start should raise IQ scores, lift children and families out of poverty, and close the achievement gap between poor and more affluent children. Congress tried to quell this practice in 1998 by mandating the singular goal of improved school readiness. Measured against this outcome, Head Start is certainly a success. Although the initial findings of the Impact Study were not as robust as hoped, they clearly indicated that by the end of the Head Start experience children were more ready for school entry than those in the control group (Ludwig & Phillips, 2007; Yoshikawa, 2005).
Of course, although there are positive post-Head Start effects, they are not nearly of the magnitude found for such programs as the Chicago Child-Parent Centers, the New Jersey Abbott Centers, and the Oklahoma state prekindergarten. It makes sense that the greater the immediate post-preschool effects, the greater the size of sustained effects. We should not ignore the repeated finding that the benefits of any preschool intervention are never completely maintained once a child has entered formal schooling. Barnett (2010) has reasoned this is not so much a “fade-out” phenomenon as it is a “catch-up” phenomenon in which the control group children catch up to the intervention group once they enter school and are exposed to academic training. Brooks-Gunn (2010) presents confirmatory evidence on this point. Given the widely found but modest impact immediately after the Head Start year, the standard diminution of this benefit as children progress through school generates the prediction that the kindergarten and first-grade results of the National Impact Study will be positive but small in size. This is exactly what was found.

There is another obvious reason why the Head Start children in the Impact Study displayed few sustained effects during kindergarten and first grade. Even the strongest optimist would not expect the benefits of preschool to be permanent. Head Start can only prepare children for school. Once the schools take over it is their responsibility to keep the momentum going. There is considerable evidence that following Head Start poor children go on to poor quality schools (Lee, Loeb, & Lubeck, 1998). Kate Walsh, president of the National Council on Teacher Quality, called it “a scandal of monumental proportions that there were two distinct school systems in the U.S., one for the middle class and one for the poor” (quoted in Thomas & Wingert, 2010, p. 26). There is a body of evidence that the benefits that accrue to Head Start children are indeed more lasting when they attend high-quality elementary schools (Magnuson,
Ruhm, & Waldfogel, 2007; Reynolds & Hayakawa, 2010). Examining the same issue, Holod, Gardner, and Brooks-Gunn (2010) found sustained effects through the third grade with the effects being more pronounced for poor children who attended more affluent schools. (A few Head Start children in the National Impact Study did go on to better schools, and their data should be examined.) Findings from the Chicago Child-Parent Centers also provide convincing evidence of enhanced preschool effects when treatment is followed by good quality schooling (Reynolds & Hayakawa, 2010). Evidence that poor schools attenuate the gains accrued from preschool intervention was provided by Zigler, Pfannenstiel, and Seitz (2008). In this study poor children whose parents had received 3 years of home visitation and who then attended a 2-year preschool program were found to be equivalent to middle-class children on a comprehensive school readiness measure. However, 3 years after school entry the poor children had fallen behind. It would be unfair to hold Head Start responsible when its graduates lose their advantage once they attend failing schools. It is also unreasonable to expect that a brief preschool experience will have more power over children’s academic fate than their experiences in elementary schools, which have them a lot longer than Head Start does.

A related issue concerns the length of treatment. The Impact Study assessed outcomes after less than 9 months of Head Start enrollment (U. S. Department of Health and Human Services, 2010). Reynolds and Hayakawa (2010) pointed out that no one should expect enduring effects from a program that is less than 9 to 12 months in length. After the initial year of the Impact Study, all the 3 year olds in the Impact Study were given access to Head Start. Thus in the 3-year-old cohort half of the control group attended Head Start when they turned 4 years old, and over 70% of the children in the Head Start group returned for a second year. Yet the investigators did not assess the value of having 2 years of Head Start. Instead they assessed the
value of having *access* to the program at a younger age. This research question somewhat defies interpretation, but at least the data on 2 years of attendance exist and can be mined in the future.

**Beyond the Impact Study**

The National Impact Study was widely heralded as the best longitudinal study of the effects of Head Start to date. It is certainly better than much earlier work, but, as we have seen, it is far from a perfect assessment. And even if it was, it would be wrong to treat the results in isolation. Barnett wisely reasons that “New studies don’t simply obviate everything that has gone before” (2010). He correctly argues that the National Impact Study must be interpreted within the context of the entire 45 years of research that have now been conducted on preschool intervention.

Countering the weak effects found in the longitudinal follow-up of the National Impact Study, much evidence exists of the benefits of Head Start. Indeed, some of this evidence is provided by the original findings of the Impact Study itself, when the children were tested immediately following Head Start. These results are harmonious with Currie’s (2001) description of the literature as clearly showing that the school readiness achieved by Head Start children is greater than that achieved by children attending more typical child care centers or family day care homes. At a congressional Science and Public Policy briefing, Steve Barnett (2002) also summarized the data and concluded that Head Start children have higher achievement test scores and that the program has favorable long-term effects on grade repetition, special education, and high school graduation. Consistent with Barnett’s conclusions, Garces, Thomas, and Currie (2002) found that Head Start graduates not only had less grade retention and greater educational achievements, but they exhibited less criminal behavior when they were young adults. Specifically, Head Start was associated with a reduction in the chance of being
arrested by approximately 12 percentage points. This results in huge dollar savings. A great
deal of the cost benefits reported for model programs such as the Perry Preschool were due to
less criminal justice and victim costs. There is enough evidence to date to indicate that Head
Start too is a cost-effective program. On the basis of the findings of the National Impact Study
immediately after Head Start, Ludwig and Phillips concluded that the small positive effects
found at the end of Head Start would be “large enough to generate long-term money value
benefits that outweigh program costs” (2007, p. 6).

Further evidence that Head Start is a sound program when assessed by its dollar benefits
vs. its dollar costs came in a report by the Harvard economist David Deming. He concluded that
Head Start’s benefits were “about 80% as large as model programs such as the Perry Preschool
(2009, p. 111). (In a personal communication, Barnett, April 6 2010, has questioned Deming’s
conclusions.) The point is that Head Start produces benefits that more than cover its costs. It
may not have huge returns like the Perry Preschool and the Chicago Child-Parent Centers, but it
costs considerably less. A reasonable conclusion is that Head Start is not as good a program as
the models, nor can it be at its current level of funding. However, it still produces a fair bang for
the buck.

This conclusion too has some support from the Impact Study. A broad consensus among
scholars now exists that high-quality preschool interventions result in improved performance
both in school and later in life (National Forum on Early Childhood Policy and Programs, 2010;
makes clear that Head Start is simply not a high-quality program, but this same evidence also
indicates that neither is it a low-quality program. The original findings of the study show that
Head Start is superior to the alternatives available to poor children in America. For example, the
gap in reading skills between Head Start attendees and the national norm was only half as great as that found for the control group. Barnett too notes that “While the Head Start program is not of the quality of the model programs” it is better than many other programs poor children attend (2010).

One choice that is increasingly available to poor children is state prekindergarten programs. Barnett and the author agree that Head Start is superior to some state programs but inferior to others, e.g., Oklahoma and New Jersey. This calls into doubt the recommendation of Haskins and Sawhill (2010) who used the National Impact Study to argue that Head Start should be transferred to the states (an effort that was pursued by the George W. Bush administration). Another weakness in this proposal is that roughly some dozen states have never taken the initiative to mount preschool programs for at-risk children. This lack of commitment does not hold promise that they will suddenly roll out programs better than Head Start should devolution take place.

The critics have also ignored the one area where the Impact Study follow-up discovered clear benefits of Head Start. Children in the program had markedly better dental health and less robust but important other health benefits. The nation has never fully appreciated the benefits accruing to children just from the health component of Head Start. For example, Nisbett (2010) reported that during the early school years graduates of the Head Start program die at dramatically lower rates than comparison children. This is consistent with the findings of Ludwig and Miller (2007), who discovered lower mortality rates for Head Start children ages five to nine from causes that could be ameliorated through the program’s health services. Love, Tarullo, Raikes, and Chazan-Cohen (2006) also emphasized the poor health of entering Head Start children and the program’s value in improving their health status. One aspect of health is
nutrition, and Head Start’s nutrition component has clearly lead to less obesity in participating children (Frisvold, 2007). This finding led the noted authority on the current epidemic of obesity, Kelly Brownell, to comment, “It looks like being in Head Start might be a more powerful means of reducing obesity than any other program specifically designed for that purpose.” (personal communication, May 7, 2008). Obesity was not examined in the Impact Study.

Also overlooked by critics is the fact that approximately 12% of the children enrolled in Head Start are children with disabilities who often have special health needs. The Harvard group’s evaluation of the National Impact Study notes that the subgroups of special needs children and children living in rural areas had more robust outcomes than the average for the total group (National Forum, 2010). Many other scholars have now documented the many accomplishments of Head Start in regard to physical health, dental health, and nutrition (Hale, Seitz, & Zigler, 1990; O’Brien, Connell, & Griffin, 2004; Zigler, Piotrkowski, & Collins, 1994). It is sad but somewhat traditional that the Impact Study’s findings of health benefits were overshadowed by the disappointing smaller effects on cognitive skills.

The Real Lessons of the National Impact Study

The longitudinal findings of the Impact Study read against the total literature on early intervention provide real direction that decision makers must take if Head Start is to fulfill its potential. The critics of Head Start who have focused on the study’s findings of weak effects have chosen to ignore one of the major purposes of the project. Quoted in the follow-up report is the directive from the advisory committee that developed the blueprint for the study: “The research and findings should be used in combination with the rest of the Head Start research

The “rest of the Head Start research effort” shows that quality matters. If policymakers want larger and more sustained effects than those found in the Impact Study, they must improve quality so that children leave Head Start at the school readiness levels attained by graduates of the famous models. The Obama administration is already moving in this direction with a comprehensive roadmap containing a number of specific actions directed at improving school readiness and promoting long-term success of graduating Head Start children (Administration for Children and Families, 2010).

It is now abundantly clear that the single most important factor in children’s classroom performance is the quality of the lead teacher and his/her daily teaching practices. Much attention is given to this in the current quality improvement roadmap. Unfortunately, the present stewards of Head Start are burdened by history. From the very beginning, Head Start’s education component was weak (Barnett, 2002; Omwake, 1997; Zigler & Styfco, 2010). Part of the reason is that as a War on Poverty program, Head Start was designed to offer opportunities to poor adults. Many were hired as teachers even though they did not have any level of professional education and had no experience. Policymakers got used to the cheap labor and never provided the funds to hire more well-equipped staff. This is not to imply that Head Start has no highly qualified teachers. It has many dedicated professionals who resist the lure of better paying jobs in the public schools because they truly want to help poor children and their families get ready for school.

Congress is certainly aware of the need to improve the quality of teaching in Head Start. The 2007 reauthorization mandates that by 2013, 50% of lead teachers must have a BA degree.
But why only 50%? Every lead teacher in Head Start should have a BA, and their sorry salaries must give way to salaries that match those of teachers in public schools. Only then will Head Start centers be able to compete in the current market for trained early childhood professionals.

The value of such a policy is affirmed by data from the Oklahoma universal preschool (Gormley, Phillips, & Gayer, 2008). Head Start is part of this program, so all the lead teachers must have BA degrees just like those in the rest of the public school system. The performance of the Oklahoma Head Start children was markedly better than that found for Head Start nationwide. Another way to enhance the effectiveness of teachers concerns group size. When Head Start began in 1965, we were proud of the 5:1 ratio of students to teachers. This is now 10:1. We have much evidence that indicates that a Head Start classroom should consist of approximately 17 children with one BA-level teacher trained in early childhood education and one assistant teacher who holds either an associate’s degree or a CDA certificate.

There is no free lunch here. Quality improvement costs money, and Congress must decide whether Head Start should be a truly model program and “the pride of the nation,” as it was once called, or a less than optimal program which it is now. In addition to quality issues, Head Start has many other problems that only Congress can resolve by legislative action. The most glaring is the fact that after 45 years of operation, Head Start serves only about half of eligible children. This situation will worsen now that Congress has mandated that entry into Head Start will be permitted up to 133% of the poverty level rather than at or below poverty level. This change was a long time coming, but it will increase the size of the eligible population. Unfortunately, the mandate to serve the near-poor, like the one calling for BA-level teachers, is empty because it is largely unfunded.
Other changes are likewise long overdue. The evidence is clear that the more intense an intervention is, the stronger and more lasting the benefits. Head Start would certainly be more effective if it was a 2-year program, which it still is in a few sites, rather than a single academic-year program. Yet this would be a hard change to make because it is difficult to argue that we should give some poor children 2 years of Head Start while so many others are receiving none.

The new realities of the American family suggest another structural change for Head Start. With the end of welfare Head Start parents must work. Head Start centers must move to full-day, full-year programs that provide parents the child care they need while at the same time providing their children with a longer period of preschool education. Head Start has been slowly evolving in this direction since the Clinton administration, and this movement should be accelerated.

Head Start pioneered parent involvement, and this component must be beefed-up. The program has always invited parents to be part of their child’s education because they have the most influence over their child’s development. In this age of the working parent, Head Start must adjust its parent involvement activities to account for this new reality. The Chicago Child-Parent Centers also operate on the premise that parent involvement is important to promoting the child’s school readiness and has in place a modern parent involvement component that could easily be copied by Head Start. Another way to strengthen the parent component is to increase the number of home visits from the current two a year—a far cry from the weekly visits that helped produce the impressive outcomes of the Perry Preschool project.

The Obama administration should be commended for viewing Head Start as one segment of a birth-to-eight-system (a view the author has long espoused, e.g., Zigler & Styfco, 1998). We now know just how terribly ravaging growing up in poverty is to children’s development. To
offset the huge negative effects we must get away from 1-year inoculation thinking. A range of interventions is needed to alter the life chances of poor children. We should begin with a birth-to-three intervention (like Early Head Start), dovetailed with a 2-year preschool intervention, which in turn is dovetailed with a program through the first three grades of elementary school. The Department of Health and Human Services must work closely with the Department of Education to make such a spectrum of interventions possible.

The National Impact Study was originally conceptualized to be a catalyst for moving Head Start in directions that would be beneficial to the children and parents the program serves. The naysayers focus on the lack of robust outcomes in the array of findings. This array, however, holds many insights into what Head Start needs to do better to achieve the desired results. Since the study has already prompted a quality improvement roadmap, we can only conclude that the National Impact Study has achieved its purpose.
References


U. S. Department of Health and Human Services, Administration for Children and Families.  


