Social Policy Report
Giving Child and Youth Development Knowledge Away
Volume XXI, Number 3 2007

The Benefits and Costs of Head Start

Jens Ludwig
University of Chicago

Deborah Phillips
Georgetown University

Abstract

We review what is known about Head Start's impacts on children and argue that the program is likely to generate benefits to participants and society as a whole that are larger than program costs. Our conclusions differ from those in some previous reviews because we use a more appropriate standard to judge program effectiveness (benefit-cost analysis), draw on a body of new evidence for Head Start's long-term effects on early cohorts of participating children, and discuss why common interpretations of a recent randomized experimental evaluation of Head Start's short-term impacts may be overly pessimistic. Estimating the long-term benefits of Head Start for recent participants necessarily requires a number of assumptions. But we believe there is a plausible case that short-term effects on achievement scores of .1 to .2 standard deviations might be large enough for Head Start to pass a benefit-cost test. Data from the experiment imply that Head Start enrollment – as distinct from assignment to the experimental treatment group – usually generates impacts of at least this magnitude. While, in principle, there could be more beneficial ways of deploying Head Start resources, the benefits of such changes remain uncertain and there is some downside risk. There is a growing scientific consensus that a variety of early childhood interventions generate benefits in excess of costs at current levels of spending, which suggests the value of increased spending in this area. However there remains considerable uncertainty about what form any additional investment should take. Additional government funding to support rigorous research to identify the relative strengths of Head Start and its alternatives, as well as the critical "active ingredients" in these programs that most effectively produce short- and long-term developmental benefits, would be a particularly high value-added activity.

Social Policy Report

Editor
Lonnie Sherrod, Ph.D.
sherrod@fordham.edu

Associate Editor
Jeanne Brooks-Gunn, Ph.D.
brooks-gunn@columbia.edu

Director of SRCD Office for Policy and Communications
Mary Ann McCabe, Ph.D.
mmccabe@srcd.org

Managing Editor
Amy D. Glaspie

From the Editor

Lonnie and I are pleased to introduce the latest *Social Policy Report*, "The benefits and costs of Head Start", authored by Jens Ludwig, University of Chicago and Deborah Phillips, Georgetown University. This excellent article continues our series on early childhood development, education and policy. Past reports include "PK-3: An aligned and coordinated approach to education for children 3 to 8 years old" (Bogard & Takanishi, 2005); "Putting the child back into child care: Combing care and education for children ages 3-5" (Brauner, Gordic, & Zigler, 2004); "Kindergarten: An overlooked educational policy priority" (Vecchiotti, 2003); "Do you believe in magic?: What we can expect from early childhood intervention programs" (Brooks-Gunn, 2003); "Emotions matter: Making the case for the role of young children's emotional development for early school readiness" (Raver, 2002); "At what age should children enter kindergarten?: A question for policy makers and parents (Stipek, 2002); and "Parental leave policies: An essential ingredient in early childhood education and care policies" (Kamerman, 2000).

Ludwig and Phillips tackle the difficult problem of estimating the probable long-term effects of Head Start programs today. Our current policies are based on the fact that a handful of early childhood education programs have been found to be cost-effective over the long-run. These estimates are not based on the programs that the vast majority of young children attend—that is, the federally-funded Head Start and the primarily state-funded pre-Kindergarten programs; these programs are believed to vary in quality much more than the relatively small, intensive experimental programs with well-specified curricula and professional training on which we have based our early educational policies . Nonetheless, Ludwig and Phillips maintain that the modest short-term effects that have been reported for Head Start and pre-K programs could result in long-term effects for these youngsters, and the authors carefully outline their assumptions underlying this argument.

Thomas Cook and W. Steven Barnett, also experts on this topic, each provide a commentary, the first focusing on the plausibility of the authors' assumptions and the second on the possibility of the federal and state programs having the features necessary for sustained effects for children. Each is hopeful about the benefits of these Federal and state programs, but more cautiously than Ludwig and Phillips. As Head Start is part of the annual federal budget and as pre-K funds are often up for re-authorization or expansion by each state, this *Social Policy Report* provides a blueprint for the benefits that we might realistically expect from both types of programs. Hence it should prove helpful to these policy debates, in that itprovides a basis from which to consider the recent evaluations of Head Start and pre-K programs. Lonnie and I hope that this *Report* insures that available research enters decisions on the fate of these programs which are so important to our children and their future.

Jeanne Brooks-Gunn, Ph.D., Associate Editor
Columbia University

The Benefits and Costs of Head Start

Jens Ludwig
University of Chicago

Deborah Phillips
Georgetown University

From its inception in 1965, Head Start has served as both social intervention and national research laboratory (Phillips & White, 2004). Decades of research have generated substantial knowledge about what Head Start provides and accomplishes for young children growing up in poverty. While Head Start may be the single most heavily researched program in the country, there remains considerable debate about the program's effectiveness. Policymakers understandably want firm evidence about the value of a program that serves almost 1 million of our nation's most vulnerable young children each year at an annual cost of about $7 billion. While science deals with probabilities and operates outside of the immediate need to make decisions, policymakers must make firm and costly choices under conditions of uncertainty (Shonkoff, 2000).

In this essay we seek to help narrow the range of uncertainty about the relative costs and benefits of the Head Start program. This goal seems particularly urgent in light of the impending reauthorization of Head Start by the U.S. Congress, and is the purpose of this report.

Much of the debate about Head Start stems from confusion about how to judge the magnitude of program impacts. Besharov (2005) uses the scale of the social problem being addressed – in this case the test score gap between rich and poor children, or minority and white children – as a benchmark for Head Start's effectiveness. A common alternative is to compare effect sizes from Head Start with the scale suggested by Cohen (1977) about what constitutes a "large" versus "small" impact. We argue that the most appropriate standard for judging Head Start's effectiveness is benefit-cost analysis. Policy interventions should be held accountable for generating net benefits, not some arbitrary benchmark for what constitutes a "large" benefit, much less the requirement that the program generates miraculous benefits and totally eliminates a complicated social problem (see also Duncan and Magnuson, in press; McCartney and Rosenthal, 2000).

Assuming that readers are persuaded that benefit-cost analysis is the correct way to judge Head Start's effectiveness, two practical hurdles remain – estimating the program's benefits, and estimating program costs. A growing body of research provides at least suggestive evidence that Head Start as the program operated through the 1980s may pass a benefit-cost test. One advantage of studying cohorts of program participants from several decades ago is that we are able to follow their outcomes into adolescence and adulthood to examine whether program impacts persist over the long term. The drawback is that the data available to study children from the 1960s, 1970s and 1980s are limited in important ways. Moreover Head Start and its alternatives have been changing over time, so extrapolating from Head Start's impacts on poor children from several decades ago to the current program's impacts on children is challenging.

A recent government-funded randomized Head Start experiment provides rigorous evidence for the program's short-term impacts. But in the absence of time travel there is no way to estimate directly Head Start's very long-term impacts on today's cohorts of participating children. As a consequence, a variety of out-of-sample projections must be made about how short-term impacts for today's children will translate into long-term outcomes as they grow up. This exercise necessarily requires us to make a number of un-testable assumptions. Wherever possible we do our best to generate estimates for current Head Start's long-term benefits using as many different approaches as possible in order to assess the sensitivity of our results to the specific assumptions that we impose.

With these important caveats in mind, we believe there is a plausible case to be made that Head Start as the program operates today may also generate benefits in excess of program costs. Moreover there is some reason to believe that the ratio of benefits to costs for Head Start (as with many other early childhood interventions) may compare favorably with most other educational interventions (see also Harris, 2007). More difficult to determine with currently available evidence is where best to invest new public dollars across different types of early childhood interventions. We take up this issue in the final section of this report, focusing on the new landscape of state pre-K programs.

Evidence on Head Start's Long-Term Impacts

While researchers have been studying Head Start for over 40 years, only in recent years have social scientists made much headway in identifying the causal impacts of the program on participating children. There is now an accumulating body of evidence on Head Start's long-term impacts that seems to suggest the program probably passed a benefit-cost test for those children who participated during the program's first few decades (see Currie & Thomas, 1995, Currie, 2001, Garces, Thomas, & Currie, 2002, Ludwig & Miller, 2007).

Economists Eliana Garces, Duncan Thomas and Janet Currie evaluate Head Start by comparing the experiences of siblings who did and did not participate in the program.

The analytic sample consists of children who would have participated in Head Start in 1980 or earlier. Data consist of retrospective self reports of Head Start participation by people who have reached adulthood. While people may misremember or misreport Head Start participation, if misremembering is random then the result will be simply to lead the study to understate Head Start's impacts (i.e. attenuate the impact estimate). These sorts of within-family across-sibling comparisons help to eliminate the confounding influence of unmeasured family attributes that are common to all children within the home (but not, of course, unshared family inputs), but at the cost of substantially reducing sample size. While their study represents an important improvement over previous non-experimental studies some other limitations of this type of sibling-comparison research design remain[1].

Garces, Thomas and Currie (2002) report that non-Hispanic white children who were in Head Start are about 22 percentage points more likely to complete high school than their siblings who were in some other form of preschool, and about 19 percentage points more likely to attend some college. These impact estimates are equal to around one-quarter and one-half of the "control mean." For African-Americans the estimated Head Start impact on schooling attainment is small and not statistically significant, but for this group Head Start relative to other preschool experience is estimated to reduce the chances of being arrested and charged with a crime by around 12 percentage points, which, as with the schooling effect for whites, is a very large effect.[2]

Ludwig and Miller [2007] use a different research design to overcome the selection bias problems in evaluating the long-term effects of Head Start and generate qualitatively similar findings for schooling attainment, although unlike Garces et al. (2002) they find evidence for impacts for blacks, as well as whites. Their design exploits a discontinuity in Head Start funding across counties generated by the way that the program was launched in 1965. Specifically, the Office of Economic Opportunity (OEO) provided technical grant-writing assistance for Head Start funding to the 300 counties with the highest 1960 poverty rates in the country, but not to other counties. The result is that Head Start participation and funding rates are 50 to 100% higher in the counties with poverty rates that just barely put them into the group of the 300 poorest counties compared to those counties with poverty rates just below this threshold. So long as other determinants of children's outcomes vary smoothly by the 1960 poverty rate across these counties, any discontinuities (or "jumps") in outcomes for those children who grew up in counties just above versus below the county poverty-rate cutoff for grant-writing assistance can be attributed to the effects of the extra Head Start funding.

The result is that Head Start participation and funding rates are 50 to 100% higher in the counties with poverty rates that just barely put them into the group of the 300 poorest counties compared to those counties with poverty rates just below this threshold.

Using this regression discontinuity design, Ludwig and Miller find that a 50-100% increase in Head Start funding is associated with an increase in schooling attainment of about one-half year, and an increase in the likelihood of attending some college of about 15% of the control mean. Importantly, the estimated effects of extra Head Start funding on educational attainment are found for both blacks and whites. These estimates are calculated for children who would have participated in Head Start during the 1960s or 1970s, and cannot be calculated for more recent cohorts of program participants since the Head Start funding discontinuity across counties at the heart of this research design seems to have dissipated over time. However these schooling estimates do have the limitation of relying on data from the decennial census, which identifies the county in which census respondents are living during adulthood, rather than when they were at Head Start age. As a result these estimates could be subject to some error if there is systematic selective migration across counties, although available data do not seem to provide much support for substantial migration bias (Ludwig & Miller, 2007).

These impact estimates taken at face value would suggest that Head Start as it operated in the 1960s through the 1980s generated benefits in excess of program costs, with a benefit-cost ratio that might be at least as large as the 7-to-1 figure often cited for model early childhood programs such as Perry Preschool. Currie [2001] notes that the short-term benefits of Head Start to parents in the form of

1 There necessarily remains some uncertainty about why some children within a family but not others participate in Head Start. For example, sibling comparisons might overstate (or understate) Head Start's impacts if parents enroll their more (or less) able children to participate in the program. Moreover, this approach might understate Head Start's impacts if there are positive spillover effects of participating in the program on other members of the family, since in this case the control group for the analysis (i.e. siblings who do not enroll in Head Start themselves) will be partially treated (i.e. benefit to some degree from having a sibling participate in Head Start).

2 The share of all children ever booked or charged with a crime in their data is 9.7% for the full sample and 10% for the sibling sample. These figures do not imply that Head Start achieves more than a 100% reduction in crime for program participants, since the right comparison for the estimated Head Start effect on African-American participants is the average arrest rate for the siblings of these children, which does not seem to be reported in the study.

high-quality child care together with medium-term benefits from reductions in special education placements and grade retention might together offset between 40 and 60 percent of the program's costs.[3] Ludwig and Miller's [2007] estimates seem to imply that each extra dollar of Head Start funding in a county generates benefits from reductions in child mortality and increases in schooling attainment that easily outweigh the extra program spending.[4] In addition Frisvold [2007] provides some evidence that Head Start might reduce childhood obesity.

These findings would appear to counter one commonly-held view that only very intensive, tightly controlled, and expensive early childhood programs are capable of generating lasting benefits to poor children. What remains unclear is how Head Start might affect the life chances of low-income children *today*. Head Start's impacts on children may change over time both because the program itself evolves, and, importantly, because the types of developmental environments – at home and in early childhood programs -- that children would experience if they are not in Head Start also change as more mothers enter the labor force and the range of other local, state and federal programs for young children expands (see, for example, Hill, Brooks-Gunn, & Waldfogel, 2003). Whether the program's net impact on participating children should be larger or smaller for more recent cohorts compared to earlier cohorts of children depends on whether Head Start is improving more or less quickly than the environments that low-income children would have experienced absent Head Start. More generally this highlights a generic challenge to understanding the long-term impacts of contemporaneous government programs: we can only estimate long-term impacts for people who participated in the program a long time ago.

> These impact estimates would suggest that Head Start, as it operated in the 1960s through the 1980s, generated benefits in excess of program costs. Would such benefits be seen for Head Start in the 21st century?

Short-Term Benchmarks for Long-Term Success

Our best estimate is that Head Start currently costs nearly $9,000 per child.[5] How large would Head Start's short-term impacts need to be for us to believe that the program's long-term benefits justify program expenditures? There is no way to answer this question directly, since today's Head Start children are – obviously – still quite young. And in fact there is no entirely satisfactory way of answering this question at all, since any effort at addressing this issue necessarily requires extrapolating several estimates out of sample and imposing a number of additional un-testable assumptions.

We try to answer this question in two ways, first by examining the short-term impacts that have been found for studies of other early childhood interventions where there is also evidence for long-term benefits in excess of program costs, and, second, by estimating directly the dollar value of a standard deviation increase in early childhood test scores. While as noted above both approaches are subject to considerable uncertainty given important limitations with the available evidence, we believe there is a plausible case to be made that positive impacts on achievement test scores on the order of .1 to .2 standard deviations (and perhaps even much smaller

3 While even today a large share of all Head Start participants receive services four days a week for just part of the day and part of the calendar year, even this coverage may support part-time maternal work. Moreover even part-time Head Start coverage may be combined with other sources of care to reduce the out-of-pocket expenditures that low income working mothers pay for child care.

4 Ludwig and Miller [2007] estimate the impact of an additional $400 per four year old in Head Start funding in a county. The dollar value of the decline in child mortality is equal to around $120 per four year old in the county. They also estimate an increase in schooling attainment of around one-half year per child. Card [1999] suggests an extra year of schooling increases earnings by 5 to 10 percent. We conservatively assume the extra $400 in Head Start funding raises lifetime earnings by 2 percent per child, which Krueger [2003] shows is worth at least $15,000 in present value using a 3 present discount rate (even assuming no productivity growth over time). The benefits would be even larger if we accounted for the fact that increased schooling also seems to reduce involvement with crime [Lochner and Morretti, 2004], and that the costs of crime to society are enormous – perhaps as much as $2 trillion per year [Ludwig, 2006].

5 For 2006 Head Start's federal funding per child for children's programs is around $6,976 (Personal communication, Jens Ludwig with Craig Turner, May 21, 2007). This figure is slightly lower than the ratio of total federal spending on Head Start to the number of Head Start participants because the total Head Start federal spending figure includes costs for training and teacher assistance, research, and IT support (www.nhsa.org/download/advocacy/fact/HSBasics.pdf). Local Head Start providers are required to provide the equivalent of 20% of their federal grant in either in-kind or cash assistance, which sometimes can come from state funding to these local Head Start grantees. Many centers fulfill this requirement through the provision of in-kind resources such as parent volunteering to the local provider, subsidized classroom space (from local public schools or churches), or subsidized services. In any case multiplying federal funding per child by 1.2 yields a figure on the order of $8,400 per child. In addition the U.S. Department of Agriculture (USDA) provides funding to help give Head Start children meals and snacks, which is estimated to cost perhaps $500 per year. Head Start children might have received some of these nutrition subsidies from the USDA even if they were enrolled in child care or early childhood programs other than Head Start. Nonetheless we conservatively count these USDA subsidies as part of the program costs, and together with the local grantee match and federal funding add up to nearly $9,000 per child. This figure captures the mix of program durations across the country (ranging from 4 days per week for 3 or so hours per day up to 5 days per week, more than 6 hours per day, for 200 or more days per year). Available data make it quite difficult to derive the average cost per child from part-day programs versus full-day programs because many Head Start programs that provide expanded service coverage mix together different streams of funding.

than that) would be large enough to generate long-term dollar-value benefits that outweigh program costs.

Short-term Impacts of Yesteryear's Head Start and Perry Preschool Programs

The findings from Garces et al. (2002) and Ludwig and Miller (2007) provide at least suggestive evidence that the Head Start program of the 1960's to 1980's generated long-term benefits that were larger than program costs. If the short-term impacts of today's Head Start were about as large as the short-term impacts of yesterday's program, and if the latter passes a benefit-cost test, there would be some reason to believe that the same might be true of the current program.

Using the same sibling-difference design as in Garces et al. [2002], Currie and Thomas (1995) studied children who would have been in Head Start in the 1980s or earlier and found that Head Start participation increased scores on the Peabody Picture Vocabulary Test (PPVT) vocabulary test by around .25 standard deviations in the short term for both white and African-American children. These impacts persisted for whites, but faded out within three or four years for blacks.[6] Head Start's impacts on Peabody Individual Achievement Test (PIAT) math scores might be around half as large and were not statistically significant [p. 345, fn 10].[7] Ludwig and Miller (2007) found that a 50-100% increase in Head Start funding does not lead to statistically significant increases in 8th grade student achievement test scores in either math or reading, although they cannot rule out impacts smaller than about .2 standard deviations. Unfortunately not much is currently known about Head Start's causal effects on short-term *non*-cognitive outcomes for earlier cohorts of program participants.[8]

> ...the literature as a whole is consistent with the idea that there are multiple pathways to long-term success.

While there remains some debate about the relative importance of different early childhood cognitive or non-cognitive skills in predicting subsequent outcomes (see Duncan et al., 2005, Hinshaw, 1992, Jimerson, Egeland & Teo, 1999; Miles & Stipek, 2006; Tremblay et al., 1992), the literature as a whole is consistent with the idea that there are multiple pathways to long-term success. There is some evidence from Currie and Thomas (1995) that Head Start affects children's non-cognitive as well as cognitive outcomes, in the form of fairly sizable reductions in the risk of grade retention. Head Start impacts on short-term non-cognitive outcomes might be at least as important as those on cognitive outcomes in understanding how and why the program generates lasting benefits to participants. But unfortunately research in this area considers a wide range of different non-cognitive outcomes, which are more difficult to compare across studies compared to the results of standardized achievement tests. For this reason one might wish to interpret short-term test scores as a proxy for the bundle of early skills that promote long term outcomes. Under this interpretation the previous research on earlier Head Start cohorts suggests that short-term impacts of around .25 standard deviations for vocabulary and perhaps .1 for math might be large enough to generate long-term benefits in excess of program costs.

We can then look at the short- versus long-term impacts of the widely-cited Perry Preschool program, which provided poor 3- and 4-year old children with two years of services at a total per-child cost of about twice that of Head Start.[9] At the end of the second year of services, Perry had increased PPVT vocabulary scores by around .91 standard deviations and scores on a test of nonverbal intellectual performance (the Leiter International Performance test) by around .77 standard deviations [Schweinhart et al., 2005, p. 61]. By age 14, impacts on reading and math scores had faded to just over .3 standard deviations, but large long-term impacts were found for schooling, crime and other outcomes measured through age 40 [Schweinhart et al., 2005].

The dollar value of Perry Preschool's long-term benefits (in present dollars) range from nearly $100,000 calculated using a 7 percent discount rate to nearly $270,000 using a 3 percent discount rate [Belfield et al., 2006, p. 180-1]. By "discount rate" we essentially mean the opportunity cost of receiving the benefits from this social program sometime

6 Currie and Thomas [1995, Table 6] use a sibling-difference research design and estimate a short-term effect of Head Start on PPVT test scores of nearly 7 percentile points in the national distribution for both blacks and whites. The standard deviation of percentile ranking scores (i.e. a uniform distribution with values between 1 and 100) will be around 29 points, implying short-term effect sizes in the Currie and Thomas study of around one-quarter of a standard deviation.

7 Currie and Thomas [1995], p. 345, footnote 10, note the PIAT math results are not statistically significant, but that version of the study does not report the math point estimates themselves. However an earlier version of the study, Currie and Thomas [1993], reports results for PIAT math, PIAT reading and PPVT scores but not results interacted with age, so we cannot recover short- versus long-term effects. However the overall impacts for whites for PIAT math scores are about half as large as the PPVT results, and PIAT reading scores are about 15% of the PPVT impacts.

8 Currie and Thomas [1995, Table 4] do find some evidence that Head Start might reduce grade retention for white children who participated in the program in the 1980s or earlier.

9 Currie [2001] cites Perry costs of $12,884 per child in 1999 dollars.

in the future rather than today. For example suppose the government could invest in some interest-bearing asset that would yield a 7 percent return per year. If we were considering a program to improve the earnings of low-income people that cost $100, we would need the increased earnings that result to be equal to at least $107. Absent this, society could invest that $100 and instead give program participants the accrued principal and interest one year from now to make them better off. Put differently, $1,000 received off into the future is worth less than $1,000 received today. The higher the return to alternative investments – that is, the higher the opportunity cost of money, i.e. the discount rate – the lower the present value of $1,000 received in the future rather than today.[10]

Next, we take the leap of extrapolating to the short-term Head Start data reported above by (a) assuming that short-term test score impacts are proportional to the dollar value of long-term program benefits and (b) using the conservative 7 percent discount rate, which implies that Head Start's short-term impacts would need to be at most around 9 percent as large ($9,000 / $100,000) as those of Perry Preschool to generate benefits that are large enough to outweigh Head Start's costs of around $9,000 per child. The resulting impact estimates of .08 and .07 standard deviations for vocabulary and nonverbal performance, respectively, are well within the range of the .10 to .25 standard deviation estimates reported above. Of course, long-term gains may not be proportional to short-term impacts, there are obvious differences in the samples of children that participated in the Perry Preschool and Head Start programs, and the long-term benefits that accrue to children in early childhood programs could be different across birth cohorts because of changes over time in things like labor market conditions, social program generosity or incarceration policies. Nonetheless, at a minimum, the Perry Preschool data raise the possibility that "small" short-term impacts might be sufficient for a program with the costs of Head Start to pass a benefit-cost test.

The Value of Increasing Early Childhood Test Scores

Another way to think about how large Head Start's short-term impacts would need to be in order for the program to pass a benefit-cost test is to measure directly the value of a 1 standard deviation increase in early childhood test scores. Because few studies have followed people from early childhood all the way through adulthood, this exercise is, as with our previous estimation exercise, also subject to some uncertainty. In fact, learning more about how short-term program impacts on children's cognitive (and non-cognitive) outcomes translate into long-term changes in other behavioral outcomes of interest represents in our view one of the most important priorities for future research to support stronger benefit-cost analyses of early childhood interventions.

In any case, based on currently available evidence from the British National Child Development Study (NCDS), which includes achievement test scores measured at age 7 and earnings measured at age 33 for a sample of people born in the U.K. in 1958, Krueger has estimated that an increase in early childhood test scores in either reading or math of 1 standard deviation might plausibly be associated with higher lifetime earnings of about 8 percent.[11] If Krueger's argument is correct, then the short-term impacts on reading or math that would be needed to generate $9,000 in benefits from increased future earnings would be on the order of .09 (using a 3 percent discount rate and assuming no productivity growth).[12] This suggests that short-term effect sizes of .10 to .25 might be more than enough for Head Start to pass a benefit-cost test.

There is to date no entirely satisfactory way of determining how early test score impacts relate to longer life outcomes, especially for current cohorts of young children who would experience these benefits off into the (unknown, and unknowable) future. But the two different approaches used here both suggest that short-term impacts that would be considered quite small by the usual standards of education research – on the order of .1 standard deviations or so – could potentially generate long-term benefits that would at least equal Head Start's cost per participant (around $9,000). Given the uncertainties with these calculations, a more conservative approach would be to require that Head Start improve short-term test scores by .1 to .2 standard deviations in order to believe that the program might plausibly generate long-term benefits that could be large enough to justify the costs.

How Large Are Head Start's Current Short-Term Impacts?

The best evidence currently available on Head Start as it operates today comes from a recent randomized

10 For an excellent introduction to discounting and other related issues see Boardman, Greenberg, Veining and Weimer (1996).

11 Krueger [2003] notes that Currie and Thomas' [1999] analyses of these data imply that a 1 standard deviation increase in test scores increases lifetime earnings by around 8 percent. This impact is smaller than what has been estimated for a 1 standard deviation increase in test scores measured during adolescence for more recent US samples, which typically suggest earnings gains of around 20 percent. The difference is presumably due as Krueger notes to some combination of differences in the time period studied, the US vs UK labor markets, the fact that Currie and Thomas control for both reading and math scores simultaneously while most US studies examine one type of test score at a time in their effects on earnings.

12 Krueger [2003] reports increased lifetime earnings from a .2 standard deviation increase in test scores using a 3 percent discount rate and assuming no productivity growth of $15,174 in 1998 dollars, equal to around $18,800 in current dollars. So the effect size required to generate $7,000 in benefits is equal to ($7,000 / $18,800)*.2 = .37*.2 = .07.

experimental evaluation of Head Start's impacts measured within one year of random assignment (or, on average, about 9 months after enrollment). This evaluation was mandated by the federal government and carried out by Westat for the U.S. Department of Health and Human Services (Puma et al., 2005). The experimental design 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). Importantly, then, 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 (Zaslow, 2006).

Intent-To-Treat Effects Vs. Effects of Head Start Participation

The results of the Head Start National Impact Study (NIS) have been characterized as both "disappointingly small" (Besharov, 2005, p. 1) and as "consistently positive" and "impressive" (Yoshikawa, 2005). Much of the public discussion of these findings fails to recognize that the main results, particularly those in the executive summary to the several-hundred-page report, are not *intended* to reflect the effects of actual Head Start *participation* or "treatment" in the language of traditional experimental design. The executive summary and most of the tables in the body of the report itself focus on the causal effects of offering children the *chance* to participate in Head Start by assigning them to the Head Start experimental group – that is, the *intent-to-treat* impact. These results are often discussed as if they represent the effects of Head Start participation. They do not.

> The Head Start Impact Study examines over a hundred programs that vary in quality, children served, and implementation, rather than a small and tightly controlled demonstration.

In practice, not everyone who is offered the chance to participate in Head Start will actually enroll, and some who are assigned to the non-Head Start control group will find their way into a Head Start program, perhaps in another locale. It would be neither feasible nor ethical to prevent "control" families from seeking out other Head Start programs or to force families to participate in Head Start if they decide that it will not meet their own or their children's needs or better alternative opportunities present themselves. If some people assigned to the experimental treatment group do not participate in the program, and, relatedly, if some people assigned to the control group enroll in Head Start on their own, then the effects of Head Start participation (the effect of treatment on the treated) can be different – sometimes quite different – from the effects of treatment-group assignment.

In the Head Start experimental data, we see that around 86% of 4 year olds assigned to the experimental treatment group enrolled in Head Start, while 18% of 4 year olds assigned to the control group wound up in Head Start on their own [p. 3-7, Puma et al., 2005].[13] Indeed, as the NIS report states, these crossovers made it more difficult to find impacts.

The problems of drawing inferences about Head Start participation from the effects of treatment-group assignment can be easily seen by imagining an example in which everyone assigned to the treatment group participates in Head Start but, because of their own efforts, so does everyone in the control group. If the average quality of the Head Start programs experienced by children in the treatment and control groups were the same, the effects of treatment group assignment (the intent-to-treat estimate) would be equal to exactly zero. It would obviously be incorrect to infer from these estimates that Head Start does nothing to improve the life chances of participating children. The central point is that if Head Start participation rates are less than 100% among children assigned to the treatment group or greater than 0% among those in the control group, or both, then the effects of actual Head Start enrollment (the effect of treatment on the treated) will be larger than the estimated effect of being assigned to the treatment group (the intent-to-treat effect). More than 20 years ago, Howard Bloom [1984] proposed a method for translating intent-to-treat (ITT) effects into estimates for the effects of treatment on the treated (TOT). He noted that under some conditions we can learn about the effects of treatment participation – in this case, Head Start enrollment – by scaling differences in the treatment and control groups in average outcomes by the difference in the treatment and control groups in treatment participation rates.[14] This leads the TOT impact estimates (and standard errors) to be larger than those from the ITT estimation.

The Bloom procedure makes several assumptions:

13 The figures for 3 year olds assigned to the treatment and control groups equal 89% and 21%, respectively.

14 That is, under Bloom's procedure the TOT impact is equal to the difference in the average outcome of interest for children assigned to the treatment versus control group (the ITT impact) divided by the difference in program enrollment rates between the treatment and control group. This is numerically equivalent to estimating a two-stage least squares model (with no other covariates included in the model) where the endogenous explanatory variable of interest is Head Start program enrollment and the instrumental variable is equal to an indicator for assignment to the treatment group.

(1) that random assignment is in fact random, and that treatment group assignment has no effect on children who do not participate in Head Start [15]; (2) that everyone who would participate in Head Start if assigned to the control group would also participate if they had been assigned to the treatment group instead; and (3) that the average quality of the Head Start programs attended by children assigned to the treatment versus control groups is comparable. The Bloom TOT procedure that we use differs from the approach Westat used to calculate TOT estimates for their appendix tables to the NIS report, in that our calculations take into account both that some treatment group children did not participate in Head Start *and* that some children assigned to the control group received Head Start services anyway.[16] Note that if the assumptions mentioned above are met the Bloom procedure for calculating TOT estimates fully preserves the strength of the study's experimental design. The numerator in Bloom's calculation compares average test scores or other outcomes of interest for all children assigned to the treatment group with all children assigned to the control group, while the denominator compares the Head Start enrollment rate for all children assigned to the treatment group with the enrollment rate for all children assigned to the control group.

Why focus on the effects of actually participating in Head Start rather than the intent-to-treat estimates? One answer is that the effect sizes for the Head Start experiment's intent-to-treat estimates are often compared to estimates from Perry Preschool, the North Carolina Abedarian program and the results of more recent evaluations of universal state pre-K programs, all of which de facto estimate treatment effects given that all treated children attended the programs and all control children did not. This sort of apples (TOT)-to-oranges (ITT) comparison will understate the relative effectiveness of Head Start.

A more important reason for focusing on estimates for the effects of actually participating in Head Start (treatment on the treated) is to avoid confusion in conducting a benefit-cost analysis of Head Start. In public discussions about Head Start's costs, the focus is always on the costs per actual enrollee. The benefit measure that should be compared with this cost is then the dollar value of the benefits per enrollee – that is, the dollar-value of the gains from actually participating in Head Start.

> In practice, not everyone who is offered the chance to participate in Head Start will actually enroll and some who are assigned to the non-Head Start control group will find their way into a Head Start program,...

Head Start's Short-Term Impacts

In Table 1, we show the ITT impacts (regression-adjusted point estimates and standard errors that are converted into effect size terms) for each of the cognitive outcome domains reported in the Executive Summary of Westat's report for the first-year findings of the Head Start experiment [Puma et al., 2005]. [17] Table 1 also presents our own estimates for the effects of actually participating in Head Start (the effects of treatment on the treated) derived using Bloom's approach together with information about Head Start enrollment rates in the experiment's treatment and control groups. In the Head Start experiment, the difference in Head Start participation rates between the treatment and control groups is around 68 percentage points and so, using the Bloom procedure, we would estimate that the effects of Head Start enrollment on children are about 1.5 times as large as the intent-to-treat effects that are commonly misinterpreted to represent the effects of Head Start participation.[18] These results are best interpreted as providing a range within which the "true" effects of Head Start likely fall.

Table 1 shows that, at least for cognitive skills, all of the Head Start impact estimates point in the direction consistent with beneficial program impacts, although many of these

15 Stated differently, the latent propensity to participate in Head Start *if* assigned to the treatment group is assumed to be equivalent for children who were, in fact, assigned to the treatment and control groups. This should be true if random assignment was in fact random, since the propensity to participate in Head Start – like all other baseline characteristics – will be equally distributed between treatment and control groups (subject to sampling error).

16 The Westat NIS report describes the Bloom [1984] procedure for handling "no shows" in the treatment group, but does not use this procedure to handle the problem of control group members who wind up in Head Start on their own [p. 4-29, 4-35]. Instead the report seems to drop control group families who wind up in Head Start on their own and then re-weight the remaining control group members; see pp. 4-35,6. The report mentions the Bloom [1984] approach we use to calculate TOT impacts accounting for compliance rates in both the treatment and control groups on p. 4-36, but notes only that Westat will explore how findings from this procedure compare to their default procedure in future reports. As Westat notes, the TOT procedure that they actually employ in the study is non-experimental and so susceptible to selection bias, unlike the Bloom procedure we use, which tries to preserve the strength of the experimental design and can provide unbiased estimates for the effects of enrolling in Head Start if the assumptions outlined above are met.

17 Table 1 presents Westat's own preferred regression-adjusted point estimates and standard errors, based on Westat's examination of whether there is any evidence of program gains between the beginning of the school year and when the fall outcome measures are collected.

18 If we instead adjusted only for the fact that some but not all of those assigned to the experimental group participated in Head Start (i.e. ignored the fact that the control group received some Head Start on their own, or, put differently, assumed control group Head Start enrollment rates were zero), then since 86% of the experimental group 4 year olds participated in Head Start, the TOT estimate calculated using this procedure would be 1 / (.86 – 0) = 1.16 times the ITT estimate.

Table 1: Intent-to-treat (ITT) Effect Sizes from the National Head Start Impact Study and Estimated Effects of Treatment on the Treated (TOT)

Outcome	3 year olds ITT	3 year olds TOT	4 year olds ITT	4 year olds TOT
Woodock-Johnson letter identification	.235* (.074)	.346* (.109)	.215* (.099)	.319* (.147)
Letter naming	.196* (.080)	.288* (.117)	.243* (.085)	.359* (.126)
McCarthy draw-a-design	.134* (.051)	.197* (.075)	.111 (.067)	.164 (.100)
Woodcock-Johnson spelling	.090 (.066)	.132 (.096)	.161* (.065)	.239* (.097)
PPVT vocabulary	.120* (.052)	.17* (.077)	.051 (.052)	.075 (.076)
Color naming	.098* (.043)	.144* (.064)	.108 (.071)	.159 (.107)
Parent-reported literacy skills	.340* (.066)	.499* (.097)	.293* (.075)	.435* (.112)
Oral comprehension	.025 (.062)	.036 (.091)	-.058 (.052)	-.086 (.077)
Woodcock-Johnson applied problems	.124 (.083)	.182 (.122)	.100 (.070)	.147 (.103)

First and third columns reproduce ITT impact estimates for all cognitive outcomes reported in Westat's Executive Summary of the first year findings report from the National Head Start Impact Study, reported as effect sizes, i.e. program impacts divided by the control group standard deviation (Puma et al., 2005). Standard errors are shown in parentheses also in effect size terms; these were not included in the Westat report but were generously shared with us by Ronna Cook of Westat. Second and fourth columns are our own estimates for the effects of treatment on the treated (TOT) derived using the approach of Bloom (1984), which divides the ITT point estimates and standard errors by the treatment-control difference in Head Start enrollment rates. For 3 year olds the adjustment is to divide ITT by (.894 - .213) = .681, for 4 year olds adjustment is to divide ITT by (.856 - .181) = .675 (see Exhibit 3.3, Puma et al., 2005, p. 3-7). * = Statistically significant at the 5 percent cutoff.

point estimates are not statistically significant and in general the point estimates are larger (both absolutely and in relation to their standard errors) for 3 year olds than 4 year olds. For vocabulary, pre-reading and pre-writing skills, Head Start's TOT (the effects of treatment on the treated) effects for 3-year olds range from .15 to .35 standard deviations, while for 4 year olds the impacts on the PPVT are one-third to one-half as large and even smaller for pre-reading and pre-writing. Parent-reported literacy skills show much more pronounced Head Start impacts, equal to .5 and .4 standard deviations for 3 and 4 year olds, respectively. There are reasons to believe that the results from direct student assessments in this outcome domain may be more reliable than those derived from parent reports.[19]

Given the findings by Greg Duncan and his colleagues (in press) that early math scores are the strongest predictor of subsequent achievement test scores, one particular concern with the Head Start experiment results has seen that the impact estimates on early math scores (measured by the Woodcock-Johnson applied problems test) are not statistically significant. Head Start's impact on this test

19 Rock and Stenner [2005, p. 21] note that for the Early Childhood Longitudinal Study of the Kindergarten Class of 1998-99 (ECLS-K) parent reports of children's social competence and skills have not proven reliable, with "the main concern [being] that parents often have little basis for determining whether behavior is age appropriate." Analogous concerns could in principle apply to parent reports about their children's literacy skills.

equals .18 and .15 standard deviations for 3 and 4 year olds, respectively, although we have demonstrated elsewhere that if Westat had analyzed the experimental data pooling 3 and 4 year olds together the impact estimates for these early math scores would have been statistically significant [Ludwig and Phillips, 2007]. Duncan's study also finds that attention skills are important in predicting future test scores. The closest measure to this in the HSNIS is a variable for hyperactive behavior, where we see a Head Start impact of -.26 standard deviations for 3 year olds but a zero point estimate for 4 year olds.

The Question of Relative Effectiveness: Head Start and Pre-K

The fact that the current incarnation of Head Start seems to pass a benefit-cost test does not rule out the possibility that there could be even more cost-effective ways of deploying Head Start resources. One possibility that has figured prominently in debates about Head Start would involve dedicating a larger share of its resources to making the program more academically oriented, with likely trade-offs affecting the program's provision of health, nutrition, and social services to disadvantaged children, or embarking on a wholesale shift of public dollars from Head Start to state pre-K programs. The assumption, based on impressively large impact estimates (ranging from .26 to .80 for academic outcomes) emerging from evaluations of state pre-k programs (see Barnett, et al., 2005; Gormley et al., 2005), is that focusing a greater share of program funds and children's time on academic instruction will generate stronger achievement outcomes.

> The range of treatment effects from the Head Start Impact Study range from .15 to .35 standard deviations for 3-year-olds.

Four points are important in this context. First, the existing evaluations of contemporary state pre-K programs, while a major improvement over prior research in this area, are nonetheless all based on the regression-discontinuity design that may be susceptible to bias of unknown sign and magnitude (see Gormley & Gayer, 2005; for a discussion of the use of regression-discontinuity methodology in pre-k evaluation research). The discontinuity is introduced by the strict birthday cut-offs for pre-K entry used by the participating states. One identifying assumption here is that the selection process of children into pre-K is "smooth" around the cut-off (that is, does not change dramatically for children with birthdays on either side of the enrollment date-of-birth cutoff), but this need not be the case because there is a discrete change at the birthday threshold in terms of the choice set that families face in making this decision. The randomized trial used to evaluate Head Start is far less susceptible to these biases.

Second, the pre-K evaluations that have been done to date focus on those states that are leaders in this area. The experiences of pre-K programs in these states may or may not reflect the average pre-K effect we would observe if we made a wholesale shift of resources from Head Start to Pre-K. We do know from analyses of the Early Head Start Impact Study focused on the implementation of the Early Head Start Performance standards that those programs that implemented them fully had stronger effects on children (Love et al., 2005). Similar subgroup analyses have not yet been performed on the Head Start Impact Study, but we know from both the early intervention and child care literatures that variation in quality and context matter for the delivery and impacts of early childhood programs. At issue, however, is the fact that we presently lack a rigorous direct comparison of the developmental impacts of state pre-K and Head Start programs for comparably low-income children.

Third, among other differences with Head Start, the Oklahoma pre-K program is universal, highly accessible, and free for all 4-year olds in the state (Gormley et al., 2006). About two-thirds of four-year olds in the state are enrolled. This program, as well as the universal program in West Virginia were included as part of Barnett and colleagues' (2005) analysis of state pre-K programs. If there are positive spillover effects from attending school with more affluent or higher-achieving children, then "peer effects" could account for part of the difference in impacts between pre-K and Head Start. At a minimum, it is important to be aware that comparisons between state pre-K and Head Start impacts are, to some extent, comparing universal and targeted programs.

Fourth, the recent Head Start experimental evaluation provides rigorous information about the short-term impacts of Head Start as it has operated since the program's inception, namely as a comprehensive programs focused on nutrition, physical and mental health, parenting, and social services, as well as education. The long-term impacts from earlier Head Start cohorts summarized above also derive from this comprehensive approach to early intervention. These impacts extend to non-academic, as well as academic, outcomes. Indeed, the major share of the total dollar-value of the benefits reported for comprehensive early intervention programs derives from reductions in crime (Belfield et al., 2006), whose developmental pathways have both cognitive, especially language-communication capacities, and social-emotional roots (Bierman & Erath, 2006; McCord &

Commentary
Benefit-Cost Analysis of Early Childhood Programs

W. Steven Barnett
Rutgers University

Like many other analyses of the benefits and costs of public early care and education (ECE), this report relied on one of three studies that constitute a kind of Rosetta stone for the economics of ECE. These are the Perry Preschool (Barnett, 1996; Belfield, Nores, Barnett, & Schweinhart, 2005), Abecedarian (Barnett & Masse, 2007), and Chicago Child-Parent Center (Temple & Reynolds, 2007) studies. These three are unique in providing comprehensive estimates of costs and benefits based on follow-up from the preschool years into adulthood. Basic design characteristics and findings of these studies together with estimated costs and benefits are reported in Table 1. Methodologies were highly similar so that estimates are comparable across studies except as noted in Table 1.

Although these three studies are useful individually, they are of greater value when considered together. No one should expect any public program to produce the same results as any one of the studies. To borrow a phrase from the EPA, for any particular public ECE program "your mileage may vary." In general, variations in the population served, program design, and the neighborhood and broader social context can be expected to affect costs and benefits. Insights into how "mileage" varies with population, program, and context can be gained from comparisons among these studies and with other studies in the larger literature. Several salient examples are considered here.

All three programs served disadvantaged children who were primarily (or entirely) African-American, though there is some variation in degree of disadvantage. As a rule of thumb, one might expect similar programs implemented for broader populations to produce smaller benefits for less disadvantaged populations. There is likely to be some rough correspondence between the incidence of the problems ameliorated by ECE (e.g., special education, high school drop-out, and crime) and the economic benefits produced. There is some evidence that this occurs (Barnett & Belfield, 2006). However, larger benefits might be expected for some children not included in these studies, particularly children from non-English speaking backgrounds (Gormley, Gayer, Phillips, & Dawson, 2005).

All three programs were intensive compared to the ECE available to most American children, including typical Head Start and state pre-K. They had well-paid, highly qualified teachers with strong supervision. Staffing ranged from the Perry Preschool's one teacher for every 6 children to Chicago's teacher and aide for each 16 children. The Abecedarian program provided child care in full-day, year-round services from the first year of life to age 5. The other two programs offered half-days over up to two school years. The program differences are evident in costs. All three cost more than typical child care and many public pre-K programs. Chicago was less expensive than Head Start and some state pre-K programs.

In essence, Chicago was a less intensive replication of the Perry half-day Pre-K approach. As a result, Chicago cost much less. Chicago also yielded all of the same types of effects, but each is smaller, resulting in smaller economic benefits. While differences in the population or context could account for some of the differences in outcomes, these other differences are relatively small. Overall, the pattern is highly suggestive of a dose response relationship between intensity in the classroom and benefits from child gains.

The Abecedarian program provided over 5 times as many hours per year as the half-day school-year programs, and more years of service. Its cost is correspondingly high. Yet, the reduced child care costs and increased maternal earnings together more than offset the additional cost. Moreover, the maternal earnings benefit estimate includes only gains after children entered school, the result of more persistent labor force participation during the preschool years. The immediate effect on earnings from free child care birth to five was not measured, but would only add to estimated benefits. Thus, the high cost of birth to five high-quality child care due to hours and duration turns out to be misleading. The extra time is basically self-financing, at least for a population where employment is significantly constrained by

the affordability of quality child care. This has implications for Head Start, which is often half-day and typically on a school-year schedule.

Finally, differences in crime benefits across the three studies raise perplexing issues. The Perry program had large benefits from crime reduction. As expected, Chicago had smaller benefits. Abecedarian had none. Differences in population and neighborhoods could contribute to the results. However, program differences may have played a role. A curriculum comparison study involving the Perry Preschool found social and emotional development highly sensitive to differences among curricula (Schweinhart, Weikart, & Larner, 1986). There were early indications that Abecedarian had negative impacts on aggression (Haskins, 1985). Other research suggests that Abecedarian's early start and long hours might be implicated (Belsky, et. al., 2007). This could imply a tradeoff between child care benefits and some child development benefits. Given the potential magnitude of these benefits, research on how to secure both child care and socio-emotional development benefits should have a high priority.

References

Barnett, W. S. (1996). *Lives in the balance: Age-27 benefit-cost analysis of the High/Scope Perry Preschool Program* (Monographs of the High/Scope Educational Research Foundation, 11). Ypsilanti, MI: High/Scope Press.

Barnett, W. S., & Belfield, C. R. (2006). Early childhood development and social mobility. *Future of Children*, *16*(2), 73-98.

Barnett, W. S. & Masse, L. N. (2007). Early childhood program design and economic returns: Comparative benefit-cost analysis of the Abecedarian program and policy implications. *Economics of Education Review*, *26*(1), 113-125.

Belfield, C.R., Nores, M., Barnett, W.S., & Schweinhart, L. (2006). The High/Scope Perry Preschool Program: Cost-benefit analysis using data from the age-40 follow-up. *Journal of Human Resources 41*(1): 162–190.

Belsky, J., Vandell, D.L., Burchinal, M., Clarke-Stewart, K.A., McCartney, K., Owen, M.T., & the NICHD Early Child Care Research Network. (2007). Are there long-term effects of early child care? *Child Development*, *78*(2), 681–701.

Gormley, W.T., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-k on cognitive development. *Developmental Psychology, 41*(6), 872-884.

Haskins, R. (1985). Public school aggression among children with varying day-care experience. *Child Development 56*(3), 689–703.

Schweinhart, L. J., Weikart, D. P., & Larner, M. B. (1986). Consequences of the three preschool curriculum models through age 15. *Early Childhood Research Quarterly, 1*(1), 15-45.

Temple, J.A. & Reynolds, A.J. (2007). Benefits and costs of investments in preschool education: Evidence from the Child-Parent Centers and related programs. *Economics of Education Review*, *26*(1), 126-144.

Table 1

Three Comprehensive Benefit-Cost Analyses

	Carolina Abecedarian	Chicago Child-Parent Centers	High/Scope Perry Preschool
Year began	1972	1983	1962
Location	Chapel Hill, NC	Chicago, IL	Ypsilanti, MI
Sample size	111	1,539	123
Research design	Randomized trial	Matched neighborhoods	Randomized trial
Ages	6 weeks to age 5	Ages 3-4	Ages 3-4
Program schedule	Full-day, year round	Half-day, school year	Half-day, school year
Selected Findings			
Special education	25% v. 48%	14% v. 25%	37% v. 50%
Retained in grade	31% v. 55%	23% v. 38%	35% v. 40%
High school graduation	67% v. 51%	62% v. 51%	65% v. 45%
Ever arrested as juvenile	Not Measured	17% v. 25%	16% v. 25%
Ever arrested as young adult	45% v. 41% (age 19-24)	Not Measured	25% v. 40% (ages 19-24)
Adult Smoker	39% v. 55% (age 21)	Not Measured	45% v. 56% (age 27)
Costs and Benefits (2006 dollars, discounted at 3%)			
Cost	$ 70,697	$ 8,224	$ 17,599
Child Care	30,753	2,037	1,051
Maternal Earnings	76,547	0	0
K-12 Cost Savings	9,841	5,989	9,787
Post-Secondary Ed. Cost	- 9,053	- 685	- 1,497
Abuse & Neglect Cost Savings	Not Measured	329	Not Measured
Crime Cost Savings	0	41,100	198,981
Welfare Cost Savings	218	Not Measured	885
Health Cost Savings	19,804	Not Measured	Not Included
Earnings	41,801	34,123	74,878
Second Generation Earnings	6,373	Not Included	Not Included
Total Benefits	$176,284	$ 83,511	$ 284,086
B-C Ratio	2.5	10.1	16.1

Commentary
The Warrant for Universal Pre-K: Can Several Thin Reeds make a Strong Policy Boat?

Thomas D. Cook
Northwestern University

Vivian C. Wong
Northwestern University

The universal pre-K movement seems to be winning its political campaign, in part thanks to social science. The dominant and empirically supported theory in education is that engaged time on task raises individual achievement, aggregate human capital and national productivity. To increase time on task, we cannot easily extend the school day or the school year, though attempts at this are being made. Nor can we easily induce students to do more homework, or get teachers to deepen their students' engagement in classroom learning, or get struggling students to respond better to "College for All" rhetoric. Of the few remaining alternatives for increasing engaged time on task, one is for children to begin their school career earlier.

Empirical findings seem to support this option. Neurological results indicating greater brain plasticity in younger years suggest that the pre-K years deserve an especially high priority (Cunha, Heckman, Lochner, & Masterov, 2006; Shonkoff & Phillips, 2000). In addition, evaluations of many different pre-K programs have shown short-term cognitive gains—even in random assignment studies (Weikart, Bond, & McNeil, 1978, Campbell & Ramey, 1995; and see Barnett, 1995; Currie, 2001; and Heckman & Masterov, 2005 for reviews). Some studies have even indicated social and economic benefits in adulthood (Schweinhart et al., 2005; Campbell, Ramey, & Miller-Johnson, 2002; Reynolds et al. 2001). Pre-K would seem to be a robustly effective intervention whose long-term financial benefits even out-weigh its costs (Belfield, Nores, Barnett, & Schweinhart, 2006).

But the studies indicating positive pre-K effects are not strong when examined individually. The Perry Preschool Project (1978) involves a very small and local sample exposed to an unusually expensive intervention evaluated according to control group criteria that could not be reproduced today. Moreover, most of the program's financial benefits are due to a few incarcerations registered during the current 20-year pro-imprisonment policy that may or may not continue into the future (Barnett, 1996). Reynolds et al.'s Chicago study (2001) depends on an opaque matching procedure and on data analyses (Heckman-type selection models and propensity scores) that have routinely failed to recreate similar effect sizes to an experiment on the same topic. This implies the possibility of a selection confound not fully controlled. The Abcedarian Project (Campbell et al., 2002) also involves a very local intervention that was even more intensive and expensive than Perry Preschool and, while cognitive gains in the early 20's were indicated, there was no clear evidence of reduced incarceration or improvements in the other adult outcomes assessed in Perry Preschool. The national Head Start evaluation (Puma et al., 2005) has a strong sampling and random assignment design, and short-term effects are evident in some domains. But they are spotty even in treatment-on-treated analyses, and we have no idea how the effects will hold up across elementary school let alone into adulthood. Fortunately, we have a long-term study of Head Start; but as the program was 40 years ago and not as it is today. Moreover, no long-term effects were observed for test scores, graduation rates or college enrollments, though these were not as good as the tests for mortality. Short-term positive results have also been claimed for Early Head Start (Love et al., 2005), but only after heroic analytic effort. Finally, regression-discontinuity results show clearly that five state programs have raised achievement (Wong, Cook, Barnett, & Jung, 2007; Barnett, Lamy, Jung, Wong & Cook, 2007). But the five states have better than average pre-K programs, effects were stronger for alphabet learning than for more general pre-reading or mathematical skills, and long-term effects cannot be ascertained yet.

These findings are all the more limited because of a temporal mis-match built into almost all the long-term benefit-cost calculations now available. We are most interested in the long-term results of current programs implemented in the immediate future; but it is self-evident that such results cannot be directly observed. Instead, an indirect case has to

be cobbled together from long-term studies implemented in a past that does not match even today, let alone any realistically imaginable future. All pre-K policy has to be based on extrapolative leaps of faith from data as well as on the data themselves, on educational and human development theory, and on political realities.

Fortunately, the existing theory and findings are at least consistent, leading us to revise our priors and believe that short-term cognitive effects of national pre-K are very likely and that effects into adulthood are plausible. But we are not yet sure that these various thin reeds can be woven together into a truly sturdy pre-K boat capable of weathering most future storms. Indeed, one such storm is already on the horizon. Latino children are currently under-represented in pre-K (Rumberger & Tran, 2006) and would doubtless remain so under a universal pre-K program. Since Latino children already do very poorly throughout their school careers (U.S. Department of Education, 2003), the implication is that universal but voluntary pre-K may cause them to fall even further behind other groups. This would not be due just to lower enrollments; it would also occur if elementary school teachers raise their standards to accommodate the more numerous and better trained pre-K graduates they encounter who come from disproportionately non-Latino groups. To hope that these teachers will not raise elementary school standards because of universal pre-K is perverse, for this would reduce the benefits to other students and the nation at large. Latino access to pre-K presents a serious problem that may get worse. While enrollment campaigns targeted at Latino families may reduce the problem, they are not likely to achieve what a mandatory pre-K program would. But mandatory pre-K opens up a large can of cacophonously strident political worms that current advocates of pre-K would doubtless prefer to avoid.

References

Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *The Future of Children*, 5(3), 25-50.

Barnett, W. S. (1996). *Lives in the balance: Age-27 benefit-cost analysis of the High/Scope Perry Preschool Program*. Ypsilanti, MI: High/Scope Press.

Barnett, W. S., Lamy, C., Jung, K., Wong, V. C., & Cook, T. D. (2007). Effects of five state pre-kindergarten programs on early learning: National Institute for Early Education Research.

Belfield, C. R., Nores, M., Barnett, W. S., & Schweinhart, L. J. (2006). The High/Scope Perry Preschool Program: Cost-benefit analysis using data from the age-40 followup. *Journal of Human Resources*, 41(1), 162-190.

Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early Childhood Education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science*, 6(1), 42-57.

Cunha, F., Heckman, J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence of life cycle skill formation. In F. Hanushek & F. Welch (Eds.), *Handbook of the Economics of Education*. North Holland, Amsterdam: Elsevier.

Currie, J. (2001). Early Childhood Education Programs. *Journal of Economic Perspectives*, 15(2), 213-238.

Heckman, J. J., & Masterov, D. V. (2005). *The productivity argument for investing in young children*. Unpublished manuscript, Chicago.

Love, J. M., Kisker, E. E., Ross, C., Constantine, J., Boller, K., Chazan-Cohen, R., et al. (2005). The effectiveness of Early Head Start for 3-year-old children and their parents: Lessons for policy and programs. *Developmental Psychology*, 41(6), 885-901.

Puma, M., Bell, S., Cook, R., Heid, C., & Lopez, M. (2005). *Head Start Impact Study: First Year Findings*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families.

Reynolds, A. J., Temple, J. A., Tobertson, D. L., & Mann, E. A. (2001). Long-term effects of an early childhood intervention on educational achievement and juvenile arrest. *JAMA*, 285(18), 2339-1026.

Rumberger, R., & Tran, L. (2006). *Preschool participation and the cognitive and social development of language minority students*. Los Angeles, CA: University of California, Los Angeles.

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2005). *The High/Scope Perry Preschool study through age 40*. Ypsilanti, MI: High/Scope Press.

Shonkoff, J. P., & Phillips, D. (2000). *From Neurons to Neighborhoods: The Science of Early Child Development*. Washington, DC: National Academies Press.

U.S. Department of Education (2003). Status and trends in the education of Hispanics. National Center for Education Statistics.

Weikart, D. P., Bond, J. T., & McNeil, J. T. (1978). *The Ypsilanti Perry Preschool Project: Preschool years and longitudinal results through fourth grade*. Yipsilanti, MI: High/Scope Press.

Wong, V. C., Cook, T. D., Barnett, W. S., & Jung, K. (2007). An effectiveness-based evaluation of five state preschool programs using regression-discontinuity: Northwestern University and National Institute for Early Education Research.

Tremblay, 1992; Moffitt, 1993; National Research Council/ Institute of Medicine, 2001).

Indeed, some have pointed to the fact that early childhood programs like Head Start achieve long-term behavioral impacts despite "fade out" of initial achievement test score gains and speculated that lasting program impacts on non-cognitive skills might be the key drivers of long-term program impacts on outcomes such as school completion or employment (Carniero & Heckman, 2003). Of course, short-term boosts in academic skills may also be the key mechanism. Most developmentalists would argue that cognitive, emotional, and social capabilities are inextricably intertwined through-out the life course and that adult outcomes arise from complex interactions among these domains of development (Shonkoff & Phillips, 2000). Unfortunately, the recent pre-k evaluations have not yet reported findings for social-emotional or other non-academic outcomes, and we do not know if a comprehensive as compared to a more academically-oriented program provides a better early setting for fostering a broad array of outcomes.

In light of these issues, it would be risky to shift priorities substantially within Head Start in a manner that would erode its delivery of comprehensive services or to shift Head Start dollars to state pre-k programs given uncertain benefits and some downside risk. We do not, however, mean to claim that Head Start is a perfect program that cannot be improved. Low-income preschoolers are clearly capable of larger learning gains than are presently being produced by Head Start, for example. We don't know how best to achieve these gains in the context of Head Start, although plausible candidate possibilities include the use of college-educated teachers who are paid on the usual public school salary scale, focused professional development, full-day exposure to proven curricula and instructional strategies, identification and provision of extra help for students who lag behind, effective parent engagement, and support/leadership from program/school administrators (Gormley et al., 2005; Sawhill, 2006). Efforts to identify the active ingredients of pre-K success are also in their infancy.

Conclusions

There is an accumulating body of suggestive evidence that Head Start is capable of generating long-term benefits and passes a benefit-cost test, at least for children who participated during the first few decades of the program. For today's Head Start, we have rigorous evidence of short-term impacts from a recent experimental evaluation. There is obviously no direct way to empirically identify the long-term benefits of Head Start on children who are still in their early elementary school years. We instead use several different methods for estimating how short-term experimental impacts might translate into long-term outcomes. Each of these estimation approaches is imperfect, requires extrapolating out of sample, and necessarily imposes a number of un-testable assumptions. However, as Head Start re-authorization looms others have been making their own judgments about the long-term effectiveness of Head Start. We believe that our essay is a useful addition to this debate by noting that benefit-cost comparisons are a more useful standard for judging the program than other benchmarks that are regularly invoked, providing new estimates for the effects of actually participating in Head Start based on data from the recent randomized experimental Head Start study, and presenting suggestive evidence indicating that – despite its limitations – makes the general point that even impacts that are "small" by the usual standards of education or developmental research could potentially generate lifetime benefits that are large in relation to program costs.

Specifically, our calculations with their caveats in mind suggest that with a cost of $9,000 per child short-term effect sizes of .1 or .2 are likely sufficient to generate benefits in excess of costs in both the short- and long-term. The estimated effects of Head Start enrollment on children – the effects of treatment on the treated – implied by the recent experimental study of the program typically exceed this threshold.

The evidence available for a variety of early childhood interventions – ranging from relatively low-cost large-scale programs like Head Start and the Chicago Child-Parent Centers to small, very intensive randomized model experimental programs like Perry Preschool and Abecedarian – all seems to point in the general direction of lasting program benefits that on the margin are in excess of program costs (Shonkoff and Phillips, 2000, Carniero and Heckman, 2003, Belfield et al., 2006, Knudsen et al., 2006). The usual efficiency standard in public economics is, under the assumption of declining marginal benefits from expanding government programs, to invest up to the point where the marginal dollar invested generates exactly one dollar more in program benefits. By this standard there is an efficiency case to be made for substantially expanding existing investments in early childhood education.

What remains unclear is exactly what form these investments should take. The current policy landscape includes a variety of proposals on this point, which include suggestions to expand state universal pre-K programs as well as to initiate more intensive and expensive efforts that seek to "scale-up" what are believed to be the active ingredients in Perry Preschool or Abecedarian (Ludwig and Sawhill, 2007; Duncan, Ludwig and Magnuson, 2007). Perhaps the most efficient use of additional government resources at this point would be to invest more in the "R&D" necessary to make informed judgments about how best to expand different early childhood

programs and coordinate these expansions with both existing programs and elementary school curricula. In our view the key questions for expanding early childhood education are how, how much, and how soon, rather than if. Relatively modest additional investments in randomized experimentation can help shed light on these questions, which presumably should appeal to both political progressives who are eager to improve the life chances of disadvantaged children as well as those who are generally skeptical of government interventions and so eager to see evidence of efficient and practical implementation before lending their support to new public programs.

Endnotes

The authors are grateful for the support provided by the Buffett Early Childhood Fund and the McCormick Tribune Foundation to the National Forum on Early Childhood Program Evaluation through the National Scientific Council on the Developing Child. Thanks to Kathryn Clabby for outstanding research assistance. Very helpful comments were provided by Jeanne Brooks-Gunn, Philip Cook, Thomas Cook, Janet Currie, William Dickens, Greg Duncan, Dave Frisvold, Katherine Magnuson, Gillian Najarian, Matthew Neidell, Helen Raikes, Lonnie Sherrod, Jack Shonkoff, Craig Turner, Hiro Yoshikawa, Marty Zaslow, and three anonymous referees. Special thanks to Ronna Cook at Westat for making available additional information about the first-year randomized Head Start evaluation and to Amy Lowenstein for assistance with references. Any errors and all opinions are of course our own.

References

Barnett, W. S., Lamy, C., & Jung, K. (2005). *The effects of state prekindergarten programs on young children's school readiness in five states*. Retrieved April 7, 2007, from the National Institute for Early Education Research Web site: http://nieer.org/resources/research/multistate/fullreport.pdf

Belfield, C. R., Nores, M., Barnett, S. W., & Schweinhart, L. J. (2006). The High/Scope Perry Preschool Program: Cost-benefit analysis using data from the age-40 follow-up. *Journal of Human Resources*, 41(1), 162-190.

Besharov, D. J. (2005, October). Head Start's broken promise. *On the Issues*. Retrieved April 7, 2007, from the American Enterprise Institute Web site: http://www.aei.org/publications/pubID.23373/pub_detail.asp

Bierman, K., & Erath, S. A. (2006). Promoting social competence in early childhood: Classroom curricula and social skills coaching programs. In K. McCartney & D. Phillips (Eds.), *The handbook of early childhood development* (pp. 595-615). Malden, MA: Blackwell.

Bloom, H. S. (1984). Accounting for no-shows in experimental evaluation designs. *Evaluation Review*, 8(2), 225-246.

Boardman, A. E., Greenberg, D. H., Veining, A. R., & Weimer, D. L. (1996). *Cost-benefit analysis: Concepts and practice*. Toronto, Canada: Prentice Hall.

Card, D. (2001). Estimating the returns to schooling: Progress on some persistent econometric problems. *Econometrica*, 69(5), 1127-1160.

Carneiro, P., & Heckman, J. J. (2003). Human capital policy. In J. J. Heckman & A. B. Krueger (Eds.), *Inequality in America: What role for human capital policies?* (pp. 77-240). Cambridge, MA: MIT Press.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Currie, J. (2001). Early childhood education programs. *Journal of Economic Perspectives*, 15(2), 213-238.

Currie, J., & Thomas, D. (1993). *Does Head Start make a difference?* (NBER Working Paper No. 4406). Retrieved April 7, 2007, from the National Bureau of Economic Research Web site: http://papers.nber.org/papers/w4406.v5.pdf

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

Currie, J., & Thomas, D. (1999). *Early test scores, socioeconomic status, and future outcomes* (NBER Working Paper No. 6943). Retrieved April 7, 2007, from the National Bureau of Economic Research Web site: http://papers.nber.org/papers/w6943.pdf

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (in press). School readiness and later achievement. *Developmental Psychology*.

Duncan, G. J., & Magnuson, K. A. (in press). Penny wise and effect size foolish. *Child Development Perspectives*.

Duncan, G.J., Ludwig, J., & Magnuson, K.A. (in press). Reducing poverty through preschool interventions. *The Future of Children*.

Frisvold, D. (2007, January). *Head Start participation and childhood obesity*. Paper presented at the Allied Social Science Association Meetings, Chicago, IL.

Garces, E., Thomas, D., & Currie, J. (2002). Longer-term effects of Head Start. *American Economic Review*, 92(4), 999-1012.

Gormley, W., & Gayer, T. (2005). Promoting school readiness in Oklahoma: An evaluation of Tulsa's pre-K program. *Journal of Human Resources*, 40(3), 533-558.

Gormley, W. T., Gayer, T., Phillips, D., & Dawson, B. (2005).

The effects of universal pre-K on cognitive development. *Developmental Psychology*, 41(6), 872-884.

Harris, D. N. (2007). *New benchmarks for interpreting effect sizes: Combining effects with costs* (Working Paper). Madison: University of Wisconsin at Madison.

Hill, J., Brooks-Gunn, J., & Waldfogel, J. (2003). Sustained effects of high participation in an early intervention for low-birth-weight premature infants. *Developmental Psychology*, 39, 730-744.

Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin*, 111, 127-154.

Jimerson, S., Egeland, B., & Teo, A. (1999). A longitudinal study of achievement trajectories: Factors associated with change. *Journal of Educational Psychology*, 91(1) 116-126.

Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences*, 103, 10155-10162.

Krueger, A. B. (2003, February). Economic considerations and class size. *Economic Journal*, 113, 34-63.

Lochner, L., & Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American Economic Review*, 94(1), 155-189.

Love, J. M., Kisker, E. E., Ross, C., Constantine, J., Boller, K., Chazan-Cohen, R., et al. (2005). The effectiveness of Early Head Start for 3-year-old children and their parents: Lessons for policy and programs. *Developmental Psychology*, 41(6), 885-901.

Ludwig, J. (2006, September). *The costs of crime*. Testimony to the U.S. Senate Judiciary Committee, Washington, DC.

Ludwig, J., & Miller, D. L. (2007). Does Head Start improve children's life chances? Evidence from a regression-discontinuity design. *Quarterly Journal of Economics*, 122(1), 159-208.

Ludwig, J., & Phillips, D. A. (2007, March). *The benefits and costs of Head Start* (NBER Working Paper No. 12973). Retrieved April 7, 2007, from the National Bureau of Economic Research Web site: http://www.nber.org/papers/w12973

Ludwig, J., & Sawhill, I. (2007, February). *Success by ten: Intervening early, often, and effectively* (Hamilton Project Discussion Paper 2007-02). Retrieved April 7, 2007, from the Brookings Institution Web site: http://www3.brookings.edu/views/papers/200702ludwig-sawhill.pdf

McCartney, K., & Rosenthal, R. (2000). Effect size, practical importance, and social policy for children. *Child Development*, 71, 173-180.

McCord, J., & Tremblay, R. E. (Eds.). (1992). *Preventing antisocial behavior: Intervention from birth through adolescence*. New York: Guilford.

Miles, S. B., & Stipek, D. (2006). Contemporaneous and longitudinal associations between social behavior and literacy achievement in a sample of low-income elementary school children. *Child Development*, 77(1), 103-117.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, 100, 674-701.

National Research Council/Institute of Medicine. (2001). *Juvenile crime, juvenile justice*. Washington, DC: National Academies Press.

Phillips, D., & White, S. H. (2004). New possibilities for research on Head Start. In E. Zigler & S. J. Styfco (Eds.), *The Head Start debates* (pp. 263-278). Baltimore, MD: Brookes Publishing.

Puma, M., Bell, S., Cook, R., Heid, C., Lopez, M., Zill, N., et al. (2005). *Head Start Impact Study: First year findings*. Washington, DC: U.S. Department of Health and Human Services.

Rock, D. A., & Stenner, A. J. (2005). Assessment issues in the testing of children at school entry. *The Future of Children*, 15(1), 15-34.

Sawhill, I., & McLanahan, S. (2006, Fall). Introducing the issue. *The Future of Children*, 16(2), 3-18.

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2005). *Lifetime effects: The High/Scope Perry Preschool Study through age 40*. Ypsilanti, MI: High/Scope Press.

Shonkoff, J. (2000). Science, policy, and practice: Three cultures in search of a shared mission. *Child Development*, 71(1), 181-187.

Shonkoff, J., & Phillips, D. (Eds.). (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press.

Tremblay, R. E., Masse, B., Perron, D., Leblanc, M., Schwartzman, A. E., & Ledingham, J. E. (1992). Early disruptive behavior, poor school achievement, delinquent behavior, and delinquent personality: Longitudinal analyses. *Journal of Consulting and Clinical Psychology*, 60(1), 64-72.

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 April 7, 2007, from http://srcd.org/documents/policy/Impactstudy.pdf

Zaslow, M. (2006, June). *Issues for the learning community from the first year results of the Head Start Impact Study*. Paper presented at Head Start's Eighth National Research Conference, Washington, DC. Retrieved April 7, 2007, from the Child Trends Web site: http://www.childtrends.org/Files//Child_Trends-2006_06_27_SP_HeadStartPlenary.pdf

About the Authors

Jens Ludwig is Professor of Social Service Administration, Law, and Public Policy at the University of Chicago, a Non-Resident Senior Fellow at the Brookings Institution, and Faculty Research Fellow of the National Bureau of Economic Research (NBER). His research interests are in the areas of urban poverty, education, crime, and housing. He is currently serving as the Project Director for the NBER's long-term evaluation of the HUD Moving to Opportunity (MTO) residential mobility experiment, which provides low-income public housing families the opportunity to relocate to private-market housing in less disadvantaged neighborhoods. He is the co-author with Duke University professor Philip J. Cook of Gun Violence: The Real Costs (2000, Oxford University Press), and co-editor with Cook of Evaluating Gun Policy (2003, Brookings Institution Press). Ludwig holds a doctorate in economics from Duke. Prior to coming to the University of Chicago he was Professor of Public Policy at Georgetown University. He is an elected member of the Association for Public Policy Analysis and Management's (APPAM) policy council (board of directors). In 2006 he was awarded APPAM's David Kershaw Prize for distinguished contributions to public policy by the age of 40.

Deborah A. Phillips is Professor in the Department of Psychology at Georgetown University and Affiliated Faculty of the Georgetown Public Policy Institute. Her main interests are in early child development, environmental influences on development, child poverty, and the intersection of child development and public policy. She is currently involved in the National Institute of Child Health and Human Development's multi-site, longitudinal study of the developmental effects of child care and continues her research on the child care workforce. She is also interested in the care of young children with disabilities. Professor Phillips has served on numerous task forces and advisory groups that address child and family policy issues, including the Task Force on Meeting the Needs of Young Children of the Carnegie Corporation of New York and the Research Task Force of the U.S. Department of Health and Human Services' Advisory Committee on Head Start Quality and Expansion. Prior to joining the Department of Psychology, Professor Phillips directed the Board on Children, Youth, and Families at the National Academy of Sciences and has held positions in the U.S. Congress, Yale University's Bush Center in Child Development and Social Policy, and the National Association for the Education of Young Children. Her co-edited volume, From Neurons to Neighborhoods: The Science of Early Child Development, was published in 2000 by the National Academy Press. She earned her doctorate in psychology from Yale University.

W. Steven Barnett is Board of Governors Professor at Rutgers—The State University of New Jersey where he is also Co-Director of the National Institute for Early Education Research (NIEER). His primary research interests are in education economics and policy and the economics of human ecology with a particular emphasis on early childhood. He has conducted benefit-cost analyses of the Perry Preschool and Abecedarian programs. With Co-Director Ellen Frede and other colleagues at NIEER he has developed a series of studies to systematically investigate the effects of variations in preschool education program design including: curriculum, length of day, teacher qualifications and training and age at start. Recent publications include the fourth annual surveys of state Pre-K program standards and policies, *The State of Preschool: 2006 State Preschool Yearbook,* and "Benefits and Costs of Quality Early Childhood Education in the Spring 2007 issue of *The Children's Legal Rights Journal (CLRJ).* He is also co-author of "Two-way and Monolingual-English Immersion in Preschool Education: An Experimental Comparison, forthcoming in *Early Childhood Research Quarterly.* Professor Barnett earned his Ph.D. in economics at the University of Michigan.

Thomas D. Cook is Professor of Sociology, Psychology and Social Policy Joan and Serepta Harrison Chair in Ethics and Justice Northwestern University, Evanston, Ill. 60208, USA. He is interested in social science research methodology, program evaluation, whole school reform, and contextual factors that influence adolescent development, particularly for urban minorities. He received the Myrdal Prize for Science from the Evaluation Research Societies in 1982, the Donald Campbell Prize for Innovative Methodology from the Policy Sciences Organization in 1988, and the Distinguished Scientist Award of Division 5 of the American Psychological Association in 1997. He is a trustee of the Russell Sage Foundation and a member of its Committee of the Future of Work. Cook was elected to the American Academy of Arts and Sciences in April 2000 and was inducted as the Margaret Mead Fellow of the American Academy of Political and Social Sciences in April 2003.

Social Policy Report is a quarterly publication of the Society for Research in Child Development. The *Report* provides a forum for scholarly reviews and discussions of developmental research and its implications for the policies affecting children.

Electronic access to the *Social Policy Report* is available at the *Report*'s website:
http://www.srcd.org/spr.html

Purpose

Social Policy Report (ISSN 1075-7031) is published four times a year by the Society for Research in Child Development. Its purpose is twofold: (1) to provide policymakers with objective reviews of research findings on topics of current national interest, and (2) to inform the SRCD membership about current policy issues relating to children and about the state of relevant research.

Content

The *Report* provides a forum for scholarly reviews and discussions of developmental research and its implications for policies affecting children. The Society recognizes that few policy issues are noncontroversial, that authors may well have a "point of view," but the *Report* is not intended to be a vehicle for authors to advocate particular positions on issues. Presentations should be balanced, accurate, and inclusive. The publication nonetheless includes the disclaimer that the views expressed do not necessarily reflect those of the Society or the editors.

Procedures for Submission and Manuscript Preparation

Articles originate from a variety of sources. Some are solicited, but authors interested in submitting a manuscript are urged to propose timely topics to the editors. Manuscripts vary in length ranging from 20 to 30 pages of double-spaced text (approximately 8,000 to 14,000 words) plus references. Authors are asked to submit manuscripts electronically, if possible, but hard copy may be submitted with disk. Manuscripts should adhere to APA style and include text, references, and a brief biographical statement limited to the author's current position and special activities related to the topic. (See page 2, this issue, for the editors' email addresses.)

Three or four reviews are obtained from academic or policy specialists with relevant expertise and different perspectives. Authors then make revisions based on these reviews and the editors' queries, working closely with the editors to arrive at the final form for publication.

The Committee on Child Development, Public Policy, and Public Information, which founded the *Report*, serves as an advisory body to all activities related to its publication.