

December 21, 2010

Board of Advisors

Robert Boruch
University of Pennsylvania

Jonathan Crane
Coalition for Evidence-Based
Policy

David Ellwood
Harvard University

Judith Gueron
MDRC

Ron Haskins
Brookings Institution

Blair Hull
Matlock Capital

Robert Hoyt
Jennison Associates

David Kessler
Former FDA Commissioner

Jerry Lee
Jerry Lee Foundation

Dan Levy
Harvard University

Diane Ravitch
New York University

Howard Rolston
Abt Associates
Brookings Institution

Isabel Sawhill
Brookings Institution

Martin Seligman
University of Pennsylvania

Robert Solow
Massachusetts Institute of
Technology

Nicholas Zill
Westat, Inc.

President

Jon Baron
jbaron@coalition4evidence.org
202-683-8049

Jon Baron
jbaron@coalition4evidence.org
202-683-8049

1725 I Street, NW
Suite 300
Washington, DC 20006
www.coalition4evidence.org

Yvette Sanchez Fuentes
Director, Office of Head Start
HHS Administration for Children and Families
Washington, D.C. 20447

Re: Proposed Head Start Designation Renewal System (Docket ID ACF-2010-0003, RIN 0970-AC44)

Dear Ms. Fuentes:

The Coalition for Evidence-Based Policy – a nonprofit, nonpartisan organization – strongly supports HHS’s proposed system for designation renewal of Head Start grantees. We also have suggestions for the proposed criteria to identify low-performing grantees that would be subject to recompetition, as outlined in section 1307.3 of the *Federal Register* notice. Our specific comments are as follows.

1. We believe the Head Start Impact Study, showing that the program produces few positive effects at the end of first grade, provides a strong basis for program reform/recompetition.

As you are aware, HHS's Head Start Impact Study found that the program, after producing some initial gains during preschool, had almost no effect on children's cognitive, social-emotional, or health outcomes at the end of first grade. Because this study was a large, well-conducted randomized evaluation with a reasonably long-term follow-up, we believe it provides convincing evidence that the program's overall impact on children's lives falls short of what was intended.

We recognize, however, that the \$7 billion Head Start program is actually a broad funding stream that funds a wide range of Head Start agencies and centers that use different classroom curricula, personnel, and approaches. Although the impact study found that the program as a whole has few positive, sustained effects, certain Head Start centers and approaches may be highly effective. This pattern was found, for example, in the Institute of Education Sciences' recent randomized evaluation of 36 charter middle schools across 15 states: although the schools' average effect on student achievement over two years was approximately zero, the effect varied greatly across schools (schools' effect sizes ranged from -0.43 to +0.33 in reading, and from -0.78 to +0.65 in math).¹ Similarly, studies that have randomly assigned students or classrooms to teachers in grades K-5 have found large variations across teachers in their effect on student achievement.²

A similar pattern of varying effects may well apply to Head Start agencies and centers. If so, a recompetition that successfully reallocates funding from low-impact to high-impact agencies and centers could help transform a program that currently has few effects into a potent force for educational achievement of low-income children.

¹ Gleason, Philip et. al. *The Evaluation of Charter School Impacts: Final Report*, conducted by Mathematica for the Institute of Education Sciences, June 2010.

² Kane, Thomas J. Douglas O. Staiger. "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation," *NBER Working Paper 14607*, December 2008. Raj Chetty et. al. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence From Project STAR," *NBER Working Paper 16381*, September 2010.

2. However, we urge HHS to learn from Job Corps and other programs: typical methods to identify low-performing grantees are often invalid and unlikely to lead to program improvement.

The Department of Labor’s Job Corps and the Job Training Partnership Act (JTPA) programs both have used sophisticated performance measurement systems to identify high- and low-performing awardee centers, to use the results to reward and sanction centers, and in some cases to make award renewal decisions. As examples, performance measures used in one or both of these systems include service quality as determined through periodic reviews of center operations by regional office monitors, and program participants’ outcomes (e.g., employment, earnings) statistically-adjusted to account for differences in their pre-program characteristics and local labor market conditions.

Both of these programs were evaluated in large, well-conducted randomized evaluations, which in addition to measuring overall program impacts, provided measures of impact for each of the awardee centers participating in the study. Careful analyses by Schochet and Burghardt (2008)³ and Barnow (2000)⁴ comparing these center-level impacts to centers’ ratings under the performance measurement systems found little or no correlation between the impacts and ratings. The authors of the Job Corps analysis summarize their findings as follows:

The present analysis examines whether impacts . . . on key outcomes are associated with measured center performance level. Did higher-performing centers produce larger impacts than lower-performing centers? . . . [W]e find that the answer is no. Impacts on key outcome measures are not associated with measured center performance level. Thus, the performance measurement system does not seem to be achieving the goal of ranking centers on the basis of their ability to improve student outcomes relative to what these outcomes would have been in the absence of Job Corps.

The analysis found that a main reason Job Corps’ performance measurement system did not correctly predict impacts is that the econometric methods it used to “level the playing field” across centers did not effectively do so, and that centers admitting more capable participants therefore received erroneous higher performance ratings.

3. Thus, our main suggestion is that HHS, where possible, use data from the Head Start Impact Study to identify performance measures that correlate with agency and center-level impacts – and then use these measures to help identify the low-performing grantees subject to recompetition.

Specifically we suggest, as did the authors of the Head Start Impact Study, that an analysis of the study data be undertaken to “determine what types of programs, centers, classrooms . . . relate to more positive impacts for children and families.”⁵ Such an analysis – which has been done as part of other large impact evaluations – is possible because (i) the study measured a sizable number of center characteristics, such as center size, director qualifications, teacher training, and classroom quality as measured by the Early Childhood Environmental Rating Scale-Revised; and (ii) study data could be used to measure agency and center-level impacts (albeit with less precision than one might like, given the small samples in each center).

³ Schochet, Peter Z. and John A. Burghardt. “Do Job Corps Performance Measures Track Program Impacts?” *Journal of Policy Analysis and Management*, vol. 27, no. 3, pp. 556-576 (2008).

⁴ Barnow, Burt S. “Exploring the Relationship between Performance Management and Program Impact: A Case Study of the Job Training Partnership Act,” *Journal of Policy Analysis and Management*, vol. 19, no. 1, pp. 118-141 (2000).

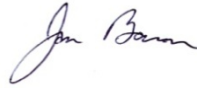
⁵ U.S. Department of Health and Human Services, Administration for Children and Families, *Head Start Impact Study: Final Report*, January 2010, p. xxxix.

This analysis could thus develop an initial set of performance measures that are validated against experimental impacts. Because some of the correlations between these measures and impacts could be spurious (due to testing for a large number of correlations), we also recommend that the validity of these measures be re-tested in future randomized evaluations involving Head Start centers.

To summarize, we are highly supportive of HHS's proposal to re-compete low-performing Head Start grantees, and urge the Department to identify such grantees, where possible, with performance measures that are validated against program impacts estimated from the Head Start Impact Study.

Thank you for the opportunity to provide input into this important program reform. We wish to make clear that our organization is nonprofit and nonpartisan, is not affiliated with any program or program model, and has no financial interest in the ideas we are proposing.

Sincerely,

A handwritten signature in cursive script that reads "Jon Baron".

Jon Baron, President