Limitations of Head Start Data Sources

A key step in understanding the validity of a program effectiveness evaluation is an assessment of the design and data limitations inherent in program studies or data sources. Design and data limitations influence which research questions can be asked and assessed; for example, small sample sizes in a study can hamper analyses about subgroups of children. The research design or analytical methods of an impact study can also influence the quality of the findings and the resulting interpretation. While there is a wealth of information collected on Head Start, there are still important constraints that should be considered when interpreting results from specific studies and drawing conclusions based on the overall body of evidence. Below are examples of the limitations inherent in selected important sources of Head Start data sources and studies.

Source	Limitations
Program Child Outcome Data	Individual Head Start programs are required to collect and use child outcome data to track student progress and improve program performance. This information could be used to implement data-driven improvements, especially at the classroom and center level. However, although grantees comply with these data collection requirements, many may lack the data capacity (e.g. research and evaluation staff, well developed IT systems) needed to analyze and make practical use of the information they collect • While programs may use child outcome data to tailor services for particular children, many struggle with how to analyze aggregate trends at the center or program level and effectively use this data improve program performance.[1] • Head Start does not require a consistent data collection tool across programs for child outcome or program quality assessment, leading to incompatible child outcome data across programs. The underdeveloped data collection capacity of individual Head Start programs impedes the program as a whole and outside researchers from evaluating program effectiveness in improving child school readiness outcomes.
Program Information Report (PIR)	The PIR reports administrative data on program capacity; it is not designed or intended to provide data about Head Start effectiveness in attaining school readiness goals. However, because it is the only annual census of all Head Start programs, it is tempting to try and use this data to evaluate program effectiveness. Researchers should be aware of limitations in the PIR data to address program effectiveness: • PIR data is collected at the program level, not at the center, classroom or individual child level. Therefore, it does not allow for an assessment of child-level outcomes or center/classroom level conditions, which constitute an important component of children's immediate environment. • There are many important measures omitted from the PIR. For example, PIR data include very few child outcome or school readiness measures and no precise measures of a child's exposure to Head Start, that is, how much time they were enrolled in Head Start. The latter measure is important because length of time in a Head Start program influences child outcomes. • Definitions have changed over time for certain measures in the PIR, such as racial and ethnic

	categories. These changes can impede attempts to analyze child enrollment trends across the years.
Family and Child Experiences Survey (FACES)	The Head Start Family and Child Experiences Survey (FACES) provides a detailed picture of Head Start programs, classrooms, teachers, and participants. The survey is administered to a nationally representative sample of three- and four-year old children entering Head Start for the first time and follows them through the end of kindergarten. Although FACES provides a wealth of information on Head Start children, families and programs across many years, there are also limitations to this data. The examples of limitations listed below are specifically with regards to FACES 2009, but many also apply to all FACES cohorts. • FACES surveys preschool-aged children in Head Start, but it does not include a comparison group of children attending non Head-Start early education centers or home-based programs. As a result, the data cannot be used to compare service components or service dosages experienced by Head Start children versus other preschool-aged children. • American Indian/Alaska Native and Migrant and Seasonal Head Start programs as well as Head Start programs in U.S. territories were not included in the study sample, so the study findings do not address children in these specific programs. • FACES is limited to those children who are newly entering into Head Start. Therefore, at the beginning of the study, four-year olds who had already completed a year of Head Start were systematically excluded from the study sample. • FACES 2009 collects data at three points: (1) at baseline, (2) after the first year of Head Start, and (3) after a second year of Head Start for four-year olds. However, because there is not a fourth data collection point at the *beginning* of the second year of Head Start for four-year olds, there is no clean measure of progress across a second year of Head Start that does not include "summer loss," the well documented decline in cognitive outcomes that occurs in the summer months when children are not exposed to school.[2] • There are small sample sizes for certain populations such as Asians or particular groups of Latinos. • Limited information on the development of certain groups of children: there is not much information on the cognitive/language development of children who speak neither Spanish nor English.[3]
Head Start Impact Study (HSIS)	The Head Start Impact Study was a randomized controlled trial to assess the impact of an offer of Head Start services (treatment) compared to no offer of Head Start services (control) that began in 2002. Head Start eligible children were randomly assigned to the treatment or control group and followed through the end of third grade. Although HSIS has a rigorous experimental design, including random assignment, it also has limitations, as follow: • HSIS was not accompanied by an implementation study or survey questions on more specific program content to capture design variation. Therefore, researchers do not know how well the intervention was carried out, or if it was carried out differently across sites. To be able to fully interpret the results of the study, a thorough understanding of children's experiences in the treatment and control groups is imperative. A new study, conducted by the Secondary Analysis of Variation in Impacts of Head Start Center will partially fill this gap by investigating how Head Start impacts were influenced by the

Head Start Impact Study (HSIS), continued	quantity and quality of treatment, as well as control children's early education experiences.[4] • The impact study did not compare the results of low versus high performing centers as a subgroup analysis, which would be key to answering policy questions about the level of quality needed to produce significant impacts on school readiness. • HSIS does not indicate how Head Start compares to other preschool programs, like state pre-kindergarten or private preschools. This is a policy relevant limitation, as the lack of direct comparisons complicates efforts to identify the most effective programs. • HSIS collected summary measures of elementary school characteristics, classroom activities and structural quality measures, and teacher characteristics. However, a more detailed child-specific elementary school implementation analysis may be needed to determine if children in the control group who on average were less school ready at kindergarten than the treatment group received more intensive services to help them catch up.[5] • HSIS does not follow children long enough to know if Head Start generates more benefits than costs. Impact evaluations of other preschool programs such as Abecedarian and Perry Preschool followed children for upwards of 20 years. • Due to ethical issues, children in the control group of the three year old cohort who were not selected for participation in Head Start during their first year of preschool were allowed to participate in Head Start for their second year. This policy makes it difficult to interpret the effects of second year of Head Start on the three year old cohort, since some three year olds attended Head Start for only their first year, some attended during their second year only, some attended both years, and some never attended the program. This mixing of control and treatment groups complicates the analysis and interpretation of Head Start impacts. • A high proportion of the control group was in similar settings (center-based care) as the treatment group (as opposed to having no exposure to center-based care), making it harder to detect Head Start impacts. • American Indian/Alaska Native and Migrant and Seasonal Head Start programs were not included in the study sample, so the study findings do not address children attending programs. • The HSIS reports do not present disparities in school readiness outcomes in a format that policymakers can readily understand. Ideally, the reports should present child school readiness outcomes/scores at baseline (to see where children started) and each follow up period (to see where they ended). They should report the same school readiness outcomes for the national child population and non-poor child population so policymakers can see how Head Start children fare relative to other children across time. They should also present a gap analysis by subgroup to show whether school readiness gaps between different subgroups of children were widening or narrowing. The data for a gap analysis were collected, yet have not been estimated in a digestible format as part of the main HSIS reports.
Head Start Research-Based, Developmentally Informed (REDI)	The Head Start REDI program was designed to promote children's social-emotional, language and emergent literacy skills by integrating a research-based enrichment curriculum and teacher professional development in Head Start programs that used either High/Scope curriculum or Creative Curriculum. From 2002 to 2004, the REDI project conducted a randomized controlled trial to assess the school readiness impacts of four-year

Intervention	olds participating in the REDI program compared to four-year olds participating in "usual" Head Start programs with the basic High/Scope or Creative Curriculum. Although the REDI evaluation has many strengths, including its experimental design and an associated implementation study, it also has limitations, as follow: • Although the REDI intervention included an implementation study of treatment (e.g., enrichment curriculum) classrooms, it did not document the experience of the control group. In other words, no measures were collected on how Head Start programs implemented their basic (e.g., non-enrichment) program curricula, which could influence the interpretation of impact results.[6] As an illustrative and hypothetical example, if control group teachers did not implement the basic program curriculum with fidelity (e.g. they added their own enhancements), then the interpretation of REDI impacts could not be considered the effects of the enhancement above and beyond the basic curriculum, because the latter was not carried out as intended. • The evaluation used very brief assessments of discrete skills as a proxy for school readiness because of the limited attention span of young children. The ability of these brief assessments to predict future cognitive and behavioral skills is unknown. • It took eight weeks to assess all children at the beginning of the study because researchers needed to work around classroom schedules and give children time to grow accustomed to the classroom setting. Researchers also had to begin post-intervention assessments eight weeks before the end of the program year. As a result, child impacts were assessed based on an average of only 25 weeks out of a 35 week program. • Teacher turnover and reassignment resulted in baseline data that was available only from 68% of teachers. • REDI intervention tested both an enriched curriculum as well as increased professional development and mentoring for teachers. Therefore, it is not possible to tease apart the effects of each of these components.[7]
Head Start Hip-Hop to Health Jr. Intervention	Hip-Hop to Health Jr. was a randomized controlled trial from 1999 to 2003 that assessed the impact of Hip-Hop to Health Jr, an obesity prevention program for minority preschoolers, on health outcomes of Head Start three- to five-year olds. Twenty-four Head Start programs serving primarily black or Latino communities in Chicago, IL were randomly assigned to an intervention group which experienced the Hip-Hop to Health Jr program, or a comparison group which received a general health intervention. Limitations associated with impact evaluation of the Head Start Hip-Hop to Health Jr. intervention include: • The results may not be generalizable to other race/ethnic populations. • Specially trained early childhood educators carried out the intervention, so the results by not be easily replicated by regular classroom teachers. • The study did not measure other types of health outcomes beyond body weight and height. Other child health outcomes measured by biomarkers or medical records, such as stress or other diagnoses, could have given a fuller picture of intervention effects. • The intervention health curriculum was conducted 3 times a week, while the control health curriculum

	was conducted once a week. Therefore, it is possible that impacts are due to dosage rather than differences in actual curricula.[8]
I Am Moving, I Am Learning – Head Start Implementation Evaluation Project	I Am Moving, I Am Learning (IM/IL) is an obesity prevention approach intended to increase and improve Head Start children's daily physical activity and healthy food choices by training Head Start educators on how to integrate obesity prevention activities into their existing practices. In 2007, the IM/IL Implementation Evaluation Project was conducted to evaluate the degree to which Head Start grantees who had received IM/IL training were implementing the IM/IL approach in their programs. Limitations include: • The Head Start programs that participated in the IM/IL Implementation Evaluation Project were not a random sample. In addition, the evaluation was voluntary, so Head Start programs were not required to participate. This may raise concerns of selection bias when interpreting results. • The study sample consists of of 17 programs in Virginia and West Virginia, so the results may not be generalizable to other Head Start programs implementing IM/IL. • The study did not include an outcomes evaluation component as it was designed to describe the implementation of IM/IL in Head Start programs, not to evaluate the impact of IM/IL on child health outcomes. • "Training of trainers" strategies used to teach Head Start staff to implement IM/IL have changed over the course of the study.[9]

Sources & notes:

[1] Office of Head Start, personal communication, August 23, 2013.

[2] For a review of the research on summer loss, see Sloan McCombs, J., Augustine, C.H., Schwartz, H.L., Bodilly, S.J., McInnis, B., Lichter, D.S., & Brown Cross, A. (2011). *Making summer count: How summer programs can boost children's learning*. Santa Monica, CA: The RAND Corporation. Retrieved from http://www.rand.org/content/dam/rand/pubs/monographs/2011/RAND_MG1120.pdf.

[3] West, J. (2013, July). *Coordinated data analysis: Maximizing early care and education data. Head Start Family and Child Experiences Survey* [Presentation slides]. Child Care & Early Education Research Connections.

[4] Secondary Analysis of Variation in Impacts of Head Start Center. Steinhardt School of Culture, Education, and Human Development, New York University. Retrieved from http://steinhardt.nyu.edu/ihdsc/research/headstart/.

[5] Yoshikawa, H., Weiland, C., Brooks-Gunn, J. Burchinal, M.R., Espinosa, L.M., Gormley, W.T.,…, Zaslow, M.J. (2013). *Investing in our future: The evidence base on preschool education*. Ann Arbor, MI: Society for Research in Child Development. Retrieved from http://www.srcd.org/sites/default/files/documents/washington/mb_2013_10_16_investing_in_children.pdf.

[6] Bierman, K.L., Domitrovich, C.E., Nix, R.L., Gest, S.D., Welsh, J.A., Greenberg, M.T.,…, Gill, S. (2008). Promoting academic and social-emotional school readiness: The Head Start REDI program. *Child Development, 79*(6), 1802-1817.

[7] Ibid; Domitrovich, C.E., Gest, S.D., Gill, S., Jones, D. & Sanford DeRousie, R. (2009). Individual factors associated with professional development training outcomes of the Head Start REDI Program. *Early Education and Development, 20*(3), 402-430.

[8] Fitzgibbon, M. L., Stolley, M. R., Schiffer, L., VanHorn, L., KauferChristoffel, K. & Dyer, A. R. (2005). Two-year follow-up results for Hip-Hop to Health Jr.: A randomized controlled trial overweight prevention in preschool minority children. *Journal of Pediatrics, 145*(5), 618-625.

[9] Fox, M.K., Hallgren, K., Boller, K. & Turner, A. (2010). *Efforts to meet children's physical activity and nutritional needs: Findings from the I Am Moving, I Am Learning implementation evaluation*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.