Title: The Role of Classroom Quality in Explaining Head Start Impacts

Authors and Affiliations:

Maia C. Connors
New York University
Department of Applied Psychology
IES-Predoctoral Interdisciplinary Research Training Fellow

Allison H. Friedman-Krauss
New York University
Department of Applied Psychology
IES-Predoctoral Interdisciplinary Research Training Fellow

Pamela A. Morris
New York University
Department of Applied Psychology

Lindsay C. Page
University of Pittsburgh
School of Education

Avi Feller
Harvard University
Department of Statistics

Abstract Body

Background / Context:

As a result of the 1998 reauthorization of Head Start, the Department of Health and Human Services conducted a national evaluation of the Head Start program (the Head Start Impact Study, HSIS; Puma et al., 2010). The goal of Head Start is to improve the school readiness skills of low-income children in the United States, and the HSIS found that it does indeed have modest impacts on children's learning. However, questions still remain with regard to the source of those impacts—how much of the effects of the assignment to Head Start are due to the entry into Head Start, the higher quality of Head Start relative to the counterfactual experiences, and/or all the other services that Head Start provides. Although results from the HSIS show that randomization to Head Start led to improvements in the quality of care children received, the original study does not explicitly explore classroom quality as a mechanism for explaining those impacts on children's learning. Thus, a particularly important part of evaluating the effectiveness of Head Start is understanding whether the program improves the quality of participants' early learning environments, and if that improvement in quality contributes to Head Starts' impacts on children's school readiness skills.

There is a substantial body of experimental and correlational research that has found associations between the quality of children's early childhood classrooms and their subsequent academic success (Burchinal, Kainz, & Cai, 2011; Pianta, Barnett, Burchinal, & Thornburg, 2009; Raver et al., 2008; 2011; Zaslow et al., 2010). Nonetheless, despite a wealth of research on how quality and quantity of care are related to *outcomes* for children, there is little rigorous research that make these linkages directly to program *impacts*. Furthermore, most prior research does not fully address issues of selection bias caused by differential care use by families. As such, the current work is important not only in "explaining" the impacts of the HSIS; it also will provide much-needed causal evidence about the effects of high quality care arrangements on children's developmental outcomes leveraging the random-assignment nature of the HSIS.

Purpose / Objective / Research Question / Focus of Study:

This study seeks to answer the question: Are impacts on Head Start classroom quality associated with impacts of Head Start on children's learning and development? This study employs a variety of descriptive and quasi-experimental methods to explore the role of classroom quality as a mediator or mechanism of Head Start impacts.

Setting:

The HSIS was designed to be nationally representative of 3- and 4-year-olds attending Head Start programs in the United States and included children in 22 states. Observations of classroom quality occurred in the child's primary care setting, including Head Start centers, other public and private center-based care facilities, and family child care homes. Direct assessments of children's cognitive skills occurred in the child's Head Start center or home.

Population / Participants / Subjects:

This research uses data from the HSIS and includes 4,440 3- and 4-year-old children who were randomly assigned off a waitlist to either receive an invitation to participate in Head Start services or to the control group. Children initially applied to 351 Head Start programs across 81 Head Start grantees. A total of 2,644 children were randomized to receive Head Start services and 1,796 were randomized to the control group. Following randomization, children enrolled in a total of 1,632 classrooms across 930 Head Start and non-Head Start centers.

Intervention / Program / Practice:

Children were randomly assigned to receive Head Start services or to a control group. The control group did not have access to Head Start; instead, some children in the control group enrolled in other center-based or family child care programs while others stayed at home with a parent, relative, or other caregiver (collectively referred to as parental care). As Head Start is based on a "whole child" model, children randomly assigned to the Head Start group had access to a set of comprehensive services including preschool education, medical, dental, and mental health care, nutrition services, and parental involvement activities.

Research Design:

Random assignment occurred prior to the beginning of the 2002-03 school year. Children were randomly assigned to Head Start within center groups rather than individual centers because of the small size of many centers: small centers were combined with nearby centers into 202 center groups. Data collection began during the fall of 2002, and classroom quality was measured during the winter and spring of 2003. Direct assessments of children's cognitive skills occurred in the fall of 2002 and spring of 2003.

Data Collection and Analysis:

Measures of children's cognitive skills include early receptive language (Peabody Picture Vocabulary Test; Dunn, Dunn, & Dunn, 1997), math skills (Woodcock Johnson III Applied Problems), and early literacy (Woodcock Johnson III Letter-Word Identification; Woodcock, McGrew, & Mather, 2001). Classroom quality was assessed using three tools: The Early Childhood Environment Rating Scale (ECERS-R; Harms, Clifford, & Cryer, 1998), The Family Day Care Rating Scale (FDCRS; Harms & Clifford, 1989), and the Arnett Caregiver Interaction Scale (CIS; Arnett, 1989). The ECERS-R and FDCRS are analogous observational tools that are used to measure quality in classrooms in center-based early childhood programs and family child care settings, respectively. Items and subscales assess the quality of space, materials, and experiences including language interactions between teachers and children. The CIS is an observational tool that focuses on the quality of interactions and relationships between teachers and children, and was used in both center-based and family child care programs. Trained independent researchers completed all observations in Head Start and non-Head Start classrooms as well as family child care homes. To improve measurement of quality, the current study utilizes three construct-specific measures of classroom quality created through exploratory and confirmatory factor analysis of items across these tools: Materials & Space for Learning, Positive Teacher-Child Interactions, and Negative Teacher-Child Interactions; all factors range from 0 to 1 (Connors, Friedman-Krauss, Jones, Morris, & Yudron, 2013).

Observational classroom quality data is missing for three different groups: 814 children who were exclusively in parental care, 601 children who were in formal care but whose classroom was not observed, and 601 children who were missing data on their type of child care setting as well as a classroom observation. Missing classroom quality data is a serious threat to our ability to understand 1) if random assignment to Head Start resulted in impacts on classroom quality, 2) if classroom quality predicts variation in impacts of Head Start random assignment, and 3) if classroom quality is a mechanism through which Head Start impacts children's outcomes.

In order to handle these three types of missing data, initial analyses explore the impacts of treatment random assignment on classroom quality using multinomial logit models to estimate the joint impact of randomization to Head Start on children's movement into formal care, classrooms that were observed, and higher quality care:

$$Pr(Y_{quality}) = B_0 + B_1 \text{Treatment} + B_2 \sum \text{center groups} + e$$

To facilitate these analyses, the three measures of classroom quality were dichotomized into high and low quality using the rough equivalent of high quality as defined by the HSIS (i.e. 5 out of 7 on the ECERS-R and 3 out of 4 on the CIS), and as indicated by the literature on early childhood classroom quality thresholds (Zaslow et al., 2010).

Future analyses will build on these initial findings and extend them using quasi-experimental methods that can begin to answer the causal question of whether impacts on classroom quality explains impacts of Head Start on children's learning. Quantitative methods for modeling mediational processes are an exciting and active area of exploration in the recent methodological literature (e.g. Rubin, 2004; Bloom, 2006; Gallop et al., 2009; VanderWeele & Vansteelandt, 2009; Bullock, Green, & Ha, 2010; Imai, Keele, & Tingley, 2010; Page, 2011a, 2011b). Nevertheless, the field of causal mediation analysis is still very much in its infancy. Therefore, we will capitalize on several analytic approaches—such as OLS regression, instrumental variables (IV) estimation, and, to the extent possible, principal stratification—to "surround" our substantive question (i.e. triangulate the findings) regarding the causal drivers of Head Start impacts. This multi-pronged analytic approach will help us understand how early childhood program quality effects the impact of Head Start on children's learning. A benefit of the principal stratification approach is that it allows us to formally evaluate the tenability of the assumptions underpinning IV estimation. However, this approach increases the "cost" of modeling and makes additional assumptions. In doing this set of work, we aim to (1) answer critical substantive questions about Head Start as well as (2) learn valuable methodological lessons about how to conduct meditational analysis in experimental studies.

Findings / Results:

As shown in table 1, there is substantial variation in quality of Head Start classrooms, with scores ranging from below .27 to above .91 on all three measures of classroom quality (please insert table 1 here).

Results of the multinomial logit models indicate that random assignment to Head Start is indeed associated with increases in all three classroom quality factors (please insert figures 1-3 here).

For example, 44% of children in Head Start compared to 6% of children in the control group are predicted to be in classrooms characterized by high quality Materials & Space, a difference of 38 percentage points. Similarly, 88% of children in Head Start compared to 36% of children in the control group are predicted to be in classrooms characterized by high quality Positive Teacher-Child Interactions, a difference of 52 percentage points. In addition, these analyses indicate that random assignment to Head Start is associated with increases in access to formal care (50% of children in the control group are predicted to be in parent care compared to only 11% of the treatment group, a difference of 39 percentage points) and increases in the ability of researchers to observe the quality that care (nearly twice as many children in the control group were in formal care that is missing a quality observation compared to children in the treatment group), an issue that we will attend to carefully in addressing our question of interest.

Further analyses will focus on the "second stage" of this analysis, examining how these impacts on quality are associated with impacts on outcomes for children using such approaches as IV and principal stratification approaches, as discussed above.

Conclusions:

Our preliminary findings show that random assignment to Head Start is associated with children's entry into care, as well as the quality of the care arrangements that they receive. Future analyses will address critical questions about the extent to which these impacts on quality are associated with impacts on child outcomes, using a variety of analytic approaches. In presenting these results, we will discuss the strengths and assumptions underlying these approaches in assessing the causal effects of quality of care on outcomes for children.

The study we propose is uniquely positioned to inform Head Start programming. Our emphasis on the predictors and mechanisms of impacts will inform questions of investment in and implementation of various features of Head Start, including both structural and process aspects of program quality. The HSIS data and these analyses provide a means to learn more about which Head Start classrooms and centers are most effective at supporting children's development. With new methodological advances in estimating causal effects in randomized trials, we are particularly well suited to take on these challenging questions. Armed with the knowledge that this paper will produce, policymakers and practitioners can make concrete improvements in aspects of their programs that are likely to make a substantial difference in outcomes for children.

Appendices

Appendix A. References

Arnett, J. (1989). *Arnett Caregiver Interaction Scale*. Retrieved from Jaeger, E. & Funk, S. (2001). *The Philadelphia Child Care Study: An Examination of Quality in Selected Early Education and Care Settings*. Philadelphia, PA: Saint Joseph's University.

Bloom, H. S. (2006). *Learning More from Social Experiments: Evolving Analytic Approaches*. New York: Russell Sage Foundation.

Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (Don't expect an easy answer). *Journal of Personality and Social Psychology*, *98*, 550-558.

Burchinal, M. R., Kainz, K., & Cai, Y. (2011). How well are our measures of quality predicting to child outcomes: A meta-analysis and coordinated analyses of data from large scale studies of early childhood settings. In M. Zaslow, I. Martinez-Beck, K. Tout & T. Halle (Eds.), *Measuring quality in early childhood settings*. Baltimore: Brookes Publishing.

Connors, M.C., Friedman-Krauss, A.H., Jones, S.M., Morris, P.A., & Yudron, M. (2013, November). Refining early measures of classroom quality. In P. Morris (Chair) *Moderators, mechanisms, methods and measurement in the Head Start Impact Study: Findings from the Secondary Analysis of Variation in Impacts of Head Start (SAVI) Center*. Symposium to be presented at the Association for Public Policy Analysis and Management Fall Research Conference, Washington, D.C.

Dunn, L.M., Dunn, L.L., and Dunn, D.M. (1997). *Peabody picture and vocabulary test, third edition (PPVT)*. Circle Pines, MN: American Guidance Service.

Gallop, R., Small, D. S., Lin, J. Y., Elliott, M. R., Joffee, M., & Ten Have, T. R. (2009). Mediation analysis with principal stratification. *Statistics in Medicine*, *28*, 1108-1130.

Harms, T., & Clifford, R. M. (1989). *Family day care rating scale (FDCRS)*. New York, NY: Teachers College Press.

Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early Childhood Environment Rating Scale (Revised Edition)*. New York, NY: Teachers College Press.

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*, 309-334.

Page, L. C. (2011a). Understanding the impact of attending a career academy high school on long-term labor-market outcomes: Bayesian estimation of causal effects in a random assignment study with partial compliance over time. Harvard University: Unpublished Paper.

Page, L. C. (2011b). Comparing methodological approaches for conducting mediation analysis: The benefits of principal stratification. Harvard University: Unpublished Paper.

Pianta, R. C., Barnett, W. S., Burchinal, M. R., & Thornburg, K. R. (2009). The effects of preschool education: What we know, how public policy is or is not aligned with the evidence base, and what we need to know. *Psychological Science in the Public Interest*, *10*, 49–88.

Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., ... & Spier, E. (2010). *Head Start Impact Study*. *Final Report*. Washington, D.C.: Administration for Children & Families.

Raver, C. C., Jones, S. M., Li-Grining, C., Metzger, M., Champion, K. M., Sardin, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. Early Childhood Research Quarterly, 23, 10-26. doi:10.1016/j.ecrqesq.2007.09.001.

Raver, C. C., Jones, S. M., Li-Grining, C. P., Zhai, F., Bub, K, & Pressler, E. (2011). CSRP's impact on low-income preschoolers' pre-academic skills: Self-regulation and teacher-student relationships as two mediating mechanisms. Child Development, 82(1), 362–378. doi:10.1111/j.1467-8624.2010.01561.x

Rubin, D. B. (2004). Direct and indirect causal effects via potential outcomes. *Scandinavian Journal of Statistics*, *31*, 161-170.

VanderWeele, T. J. & Vansteelandt, S. (2009). Conceptual issues concerning mediation, interventions and composition. *Statistics and its Interface*, *2*, 457-468.

Woodcock, R.W., McGrew, K.S., and Mather, N. (2001). *Woodcock-Johnson III tests of achievement*. Itasca, IL: Riverside Publishing.

Zaslow, M., Anderson, R., Redd, Z., Wessel, J., Tarullo, L., & Burchinal, M. (2010). *Quality Dosage, Thresholds, and Features in Early Childhood Settings: A Review of the Literature, OPRE 2011-5*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Appendix B. Tables and Figures

Table 1. *Descriptive statistics of three measures of the quality of Head Start classrooms*

	N	Mean	S.D.	Min	Max
Materials & Space for Learning	914	0.712	0.138	0.189	1.000
Positive Teacher-Child Interactions	914	0.812	0.139	0.263	1.000
Negative Teacher-Child Interactions	912	0.288	0.078	0.250	0.917
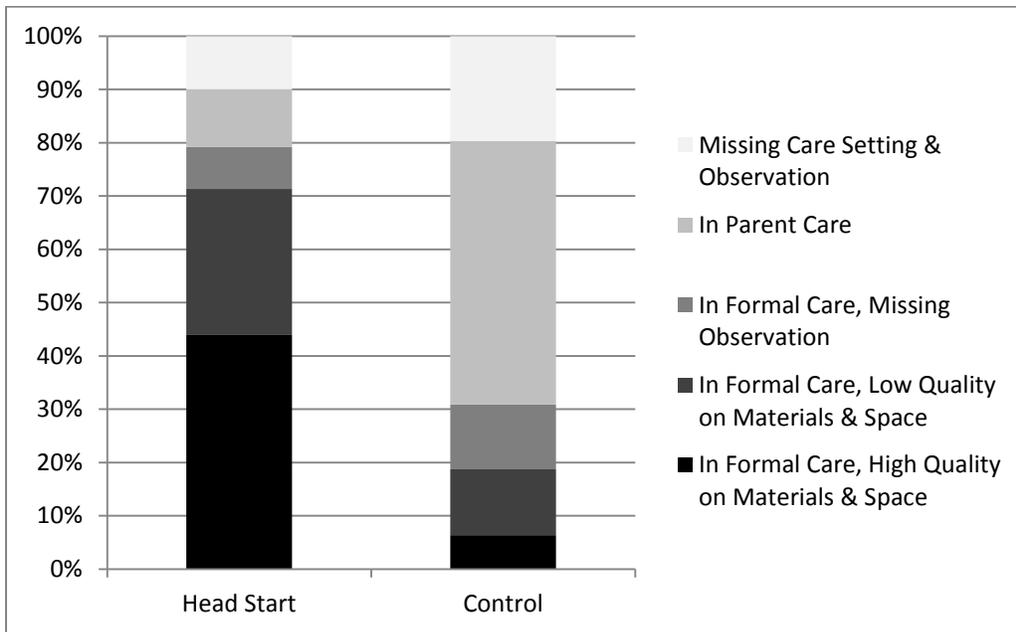
Figure 1. Impacts of random assignement to Head Start on use of formal care and the Materials & Space for Learning in children's classrooms.
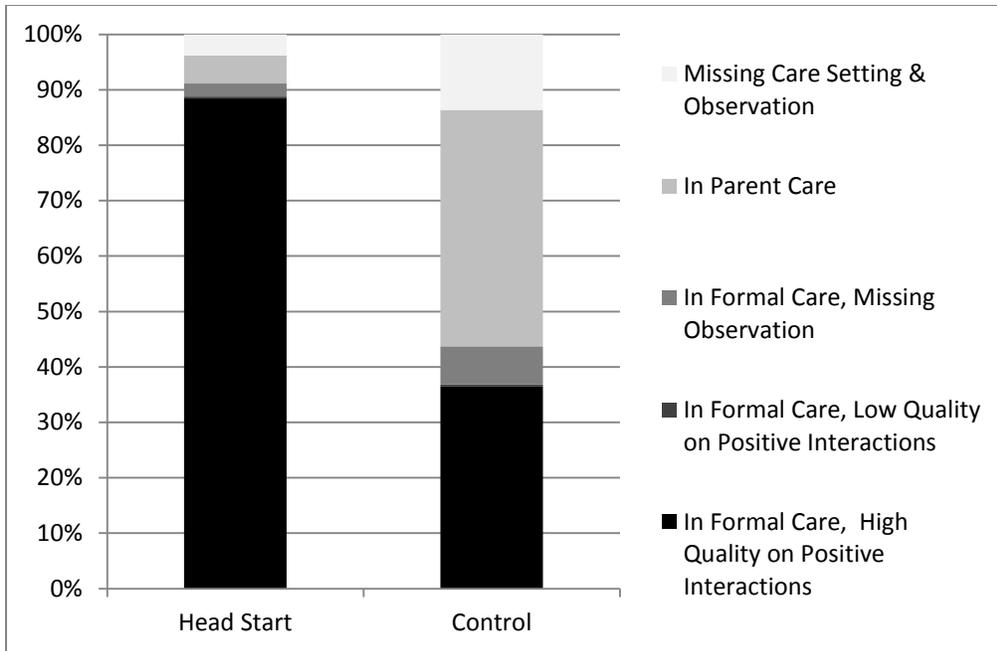
Figure 2. Impacts of random assignement to Head Start on use of formal care and Positive Teacher-Child Intearctions in children's classrooms.
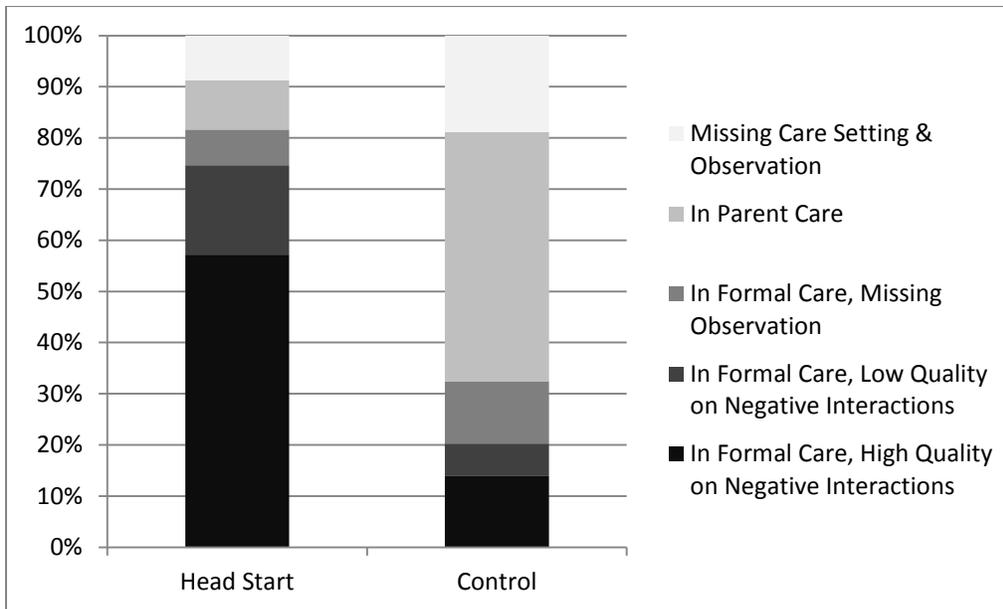
Figure 3. Impacts of random assignement to Head Start on use of formal care and Negative Teacher-Child Intearctions in children's classrooms.