Analyzing Nonresponse Bias Among Kindergarten Teachers in the Head Start Family and Child Experiences Survey (FACES)

Barbara Lepidus Carlson and Jerry West, Mathematica Policy Research

Key Words: Nonresponse Bias, Kindergarten Teachers, Head Start, FACES

Abstract: FACES follows children from their first year of Head Start through their kindergarten year. Child-level data are collected via child assessments, parent interviews, and teacher child reports. Teacher surveys are used to obtain information about FACES children's classrooms and teachers in Head Start and in kindergarten. During the one or two years that the sampled children are in a Head Start program, response rates among teachers are high; however, when the sampled children disperse to many kindergarten schools and classes, response rates are lower. This paper (1) describes methods used to increase response rates, such as using prior information on where children will attend kindergarten, and (2) explores whether nonresponse bias exists when data from the kindergarten teacher survey are used together with data from other sources. We do this by examining a set of sociodemographic characteristics of the children and schools associated with responding and nonresponding teachers.

Introduction/Background. The Head Start Family and Child Experiences Survey, or FACES, is a repeated longitudinal study of Head Start program quality and child outcomes carried out by Mathematica Policy Research for the U.S. Administration for Children and Families. It collects comprehensive data on the development of Head Start children, their families, and their educational experiences. FACES was first conducted in 1997, and a new cohort is started every three years, the most recent in 2009. The data referenced in this paper are from the FACES 2006 study, which began in fall 2006.

For each FACES cohort study, a new probability sample of children entering Head Start is selected using a multistage sampling design. Two age-based cohorts of children—3-year-olds and 4-year-olds—are sampled from each selected Head Start program[1] as children are beginning their first year of Head Start. Four-year-olds who were in Head Start the prior program year are not eligible for FACES. Each cohort is followed through one or two years of program participation, with a final followup in the spring of their kindergarten year (Table 1).

Table 1. Data Collection Schedule for 3- and 4-Year Olds

Cohort	Fall 2006	Spring 2007	Spring 2008	Spring 2009
3-year-olds	Head Start Year 1	Head Start Year 1	Head Start Year 2	Kindergarten
4-year-olds	Head Start Year 1	Head Start Year 1	Kindergarten	

To be eligible for the study, a sampled child had to be in his or her first year of Head Start and be one or two years away from kindergarten based on date of birth and local kindergarten cutoff dates. A child two years away from kindergarten was classified as a 3-year-old. A handful of children classified as 3-year-olds went to kindergarten the

[1] "Program" refers to a Head Start grantee or delegate agency. It can have more than one center at which it provides services.

following year, and a few children classified as 4-year-olds stayed in Head Start an extra year. But most children followed the expected pattern. Because children enter kindergarten in different years, the kindergarten round of data was collected over a two-year period (spring 2008 or spring 2009).

Only children for whom parental consent was obtained were included in the data collection. To remain eligible for kindergarten data collection, a child had to have remained in the Head Start program through the prior spring and then be attending kindergarten the following school year. Children who dropped out of Head Start to attend a state-funded pre-kindergarten program or who dropped out of preschool altogether were considered out of scope.

At each data collection point, including kindergarten, we administered a child assessment battery and conducted a parent interview for each sampled child. During the Head Start period, we also obtained teacher child reports from the child's lead teacher about the child's social and emotional development at each data collection point. We interviewed each child's teacher about herself and about her class at each point. The teacher child report, or TCR, is conducted as a web survey with a hard-copy option. The teacher survey is conducted using computer-assisted personal interviewing (CAPI). In kindergarten, we attempted to obtain a TCR and teacher survey from each sampled child's kindergarten teacher, but the teacher survey switched to a web mode rather than CAPI.

Motivation. When children move from Head Start to kindergarten, the data collection changes from a program-based approach to a primarily home-based approach, because the child is no longer in a Head Start program. The children disperse to many different kindergarten classrooms and schools, and the kindergarten teacher is our main source of information about the child's kindergarten experience. We combine kindergarten teacher–provided data—about the child, teacher, class, and school—with the parent- and child-provided data and attach to each child's record for analysis purposes.

Identifying and collecting data from kindergarten teachers pose challenges, and, as Table 2 shows, their participation rates decline relative to the high rates that we achieved for Head Start teachers. The rates shown in the table are unweighted child-level completion rates among those children known to be in Head Start or kindergarten.

Table 2. Teacher Survey and Teacher Child Report Completion Rates

Head Start or Kindergarten	Data Collection Point	Children in Head Start or Kindergarten	Teacher Child Report (TCR) Completion Rate (Percent)	Teacher Survey Completion Rate (Percent)
Head Start	Fall 2006	3,315	95.2	99.5
	Spring 2007	2,914	95.5	99.9
	Spring 2008	1,211	94.1	97.2
Kindergarten	Spring 2008	1,015	64.0	64.3
	Spring 2009	1,089	71.8	73.4

The fact that we had lower response rates for teacher-provided data in kindergarten than we had in the Head Start years raises questions about the potential for nonresponse bias in our estimates of child outcomes in kindergarten. The motivation for this paper is to explore this potential bias using the data we have available on all children, regardless of

whether or not their kindergarten teacher responded. We have no data on the kindergarten teachers who did not respond, so our analysis focuses on the available characteristics of the children, their families, the schools they attend, and their Head Start programs.

One way to quantify nonresponse bias is the following formula: $Bias_{NR} = \bar{Y}_R - \bar{Y}_{NR} \cdot n_{NR} / n_R + n_{NR}$. That is, it is the difference between the respondent and nonrespondent means times the nonresponse rate. If the nonresponse rate is low or the difference between the two means is small, then the bias is small. Because we generally do not know the mean for the nonrespondents, bias cannot be directly computed.

Although nonresponse bias cannot be measured directly, we do have baseline data on all children and can obtain publicly available data about the schools at which the children are attending kindergarten, if the school is known. By examining whether the children whose kindergarten teachers responded appear to fit a different profile than those whose teachers did not, we can make assumptions about the potential for nonresponse bias in our key outcome measures, such as children's school readiness skills, the types of kindergarten programs the children attend, and their classroom characteristics.

Sources of Nonresponse. Identifying and obtaining responses from children's kindergarten teachers require several steps, and we can lose kindergarten teacher responses at any of these (Table 3). Among the children who were in Head Start the spring prior to kindergarten and are in kindergarten during the following school year, the following steps are necessary to obtain data from the kindergarten teacher:

1. Determine from the parent or the Head Start program if the child is in kindergarten.
2. Determine from the parent or the Head Start program the name of the child's kindergarten school and teacher.
3. Notify the school district and obtain permission to contact the school and teacher. Note that we do not require active consent from the school district.
4. Contact the school to verify the child's enrollment and the teacher's name and address.
5. Contact the teacher to gain cooperation.
6. Have the teacher complete the TCR and teacher survey using the web or paper instrument.

If we do not determine whether the child is in kindergarten or obtain the child's kindergarten school and teacher name (steps 1 and 2), then we cannot attempt the kindergarten teacher interview. If the district, school, or teacher refuses participation, we consider the case attempted but not completed.

We ask the parent about kindergarten (whether the child is attending, school name, and teacher name) during the parent interview that we conduct in the spring of the last Head Start year and also in the spring of the kindergarten year. Sometimes the Head Start program provides this information as well.

Table 3 shows the outcomes of the different steps required to secure a completed teacher survey and/or TCR for the spring 2008 and spring 2009 kindergarten data collection. From this table we see that the percentage in the "school or teacher not known" category

is quite different between the two years, resulting in a higher overall completion rate in the second year. The rates in the table (and those that follow) are at the child level, not the teacher level, although the two are essentially the same in kindergarten, as it would be unusual for two or more of our sampled children to end up in the same kindergarten class. We include in this table those children with unknown kindergarten status, but we have excluded those known not to be in kindergarten. In this paper, a teacher's response is considered to be complete if she provided either the TCR or the teacher survey. A noncomplete teacher response can be either an active or passive refusal.

Table 3. Distribution of Final Child-Level Status Codes for Kindergarten Teacher Responses

		Spring 2008 (Percent)	Spring 2009 (Percent)
Not attempted	Not known if in kindergarten	7.6	8.1
	School or teacher not known	9.6	2.9
Attempted but not completed	District or school refused	1.4	2.8
	Teacher noncomplete	20.9	18.6
Completed	Teacher complete	60.6	67.7
	Total	100.0	100.0
Children in Head Start during the prior spring		1,098.0	1,185
Total children with teacher complete		665.0	802

Maximizing Response Rates. To maximize kindergarten teacher response, we offered to pay each teacher $25 for completing the teacher survey, and $5 or $7 for each completed TCR, with teachers being paid the higher rate for doing the TCR online. In spring 2008, the first year we did kindergarten data collection, we did not use the information about where a child would attend kindergarten from the prior spring to "jump start" this process. When preparing for spring 2009 data collection, we used the spring 2008 data to evaluate the quality of the information that parents had provided in the spring prior to their child entering kindergarten. How often did the child go to the school that the parent had indicated the prior spring? Based on this evaluation, we modified the approach for the spring 2009 collection in two ways. First, using the information from the parent interview, we began contacting schools to verify kindergarten enrollment and to identify teachers in November of the kindergarten year, instead of January/February as we had previously done. If we did not have this information from parents or if the information proved to be inaccurate, we contacted the child's Head Start program to ascertain where the child was enrolled in kindergarten. Second, we revised the look-up table for school names and addresses embedded in the kindergarten parent interview, improving its efficiency and accuracy, which reduced the number of unknown kindergarten school and teacher names. If we compare the rates in Table 3 for spring 2008 and spring 2009, we see that these were quite effective changes.

The questions examined in the remainder of this paper include the following: What characteristics are associated with a child's kindergarten teacher not responding? Do response patterns differ by subgroup? Do children with responding teachers differ from those with nonresponding teachers? This last question is not our main concern, however. It is not uncommon for respondents to differ from nonrespondents. The key question is whether the characteristics of children whose teachers responded differ from those of the full sample in any meaningful way; that is, is there any evidence of the potential for bias

in our key outcome measures when we look at observed characteristics that we think may be associated with those outcomes?

Results. We first compared response rates for selected subpopulations to see if any patterns emerged. Response rates here are unweighted child-level completion rates among those who were eligible (or possibly eligible)[2] for data collection in kindergarten. Children who were known not to be in kindergarten were ineligible and excluded from the denominator. We examined response in three ways (Table 4):

1. Number of attempted divided by all eligible (or possibly eligible)
2. Number completed divided by the number of attempted
3. Number completed divided by all eligible (or possibly eligible)

Table 4. Kindergarten Teacher Survey/TCR Completion Rates

		Attempted/ Eligible (Percent)	Completed/ Attempted (Percent)	Completed/ Eligible (Percent)
Overall		86.1	74.7	64.3
Pre-K Year	2007	82.9	73.1	60.6
	2008	89.0	76.0	67.7
Census Region	Northeast	88.4	78.5	69.4
	Midwest	88.2	80.0	70.6
	South	84.3	72.7	61.3
	West	84.8	69.0	58.5
Index of Family Risks[3]	0	93.3	79.6	74.3
	1–2	87.2	74.7	65.3
	3–4	82.0	72.4	59.4
Child Race/ Ethnicity	White Non-Hispanic	88.6	83.4	73.8
	Black Non-Hispanic	84.8	72.4	61.4
	Hispanic	85.0	71.1	60.5

From Table 4, we can see that the three types of completion rates were higher than the overall rates for children who were enrolled in Head Start in spring 2008 (primarily children who entered Head Start and were sampled at age 3), for children attending Head Start programs in the Northeast and Midwest, for those with fewer family risk factors, and for White non-Hispanic children. But this does not necessarily indicate that we have bias. To get a better indication of bias, we turned to a comparison between the profiles of the respondents and the nonrespondents. We first examined the following three types of comparisons:

1. Attempted versus not attempted
2. Completed versus attempted but not completed
3. Completed versus not attempted or not completed

[2] Children whose kindergarten status was unknown are referred to here as "possibly eligible."
[3] This index counts the number of the following risks that applied to the family: mother with less than a high school diploma, single parent, living below poverty level, and non-English speaking.

For each of these, we looked at the unweighted distributions for a number of observed characteristics (listed below) that were possibly associated with our key child outcome measures to assess the potential for nonresponse bias (we were fortunate to have data on all children at baseline, whether or not their kindergarten teacher responded). We looked no further at any characteristics when we found no evidence of a statistically significant difference (chi-square or t-test p-value <.05) between respondents and nonrespondents for any of the three definitions of response listed above. Although the unweighted distributions are not shown here in a table, these nonsignificant characteristics are indicated by *italicized type* in the list that follows.

- Child's Head Start program and center
  - census region
  - urbanicity
  - program size
  - *center size*
  - racial/ethnic minority enrollment
  - language minority enrollment
  - disability enrollment
  - whether part of a public school

- Child's baseline Head Start teacher
  - education, race/ethnicity
  - *years of teaching*
  - years of teaching Head Start

- Child's baseline Head Start classroom
  - *Half day or full day or home visits*
  - size
  - *child-to-teacher ratio*

- Child characteristics at baseline
  - age cohort
  - race/ethnicity
  - *how often read to*
  - *health status*
  - *BMI*
  - attending child care
  - vocabulary score
  - social skills score
  - behavior problems

- Parent/family characteristics at baseline
  - home language
  - family structure
  - marital status
  - highest education
  - income
  - poverty level
  - index of family risks
  - *mother's age at first birth*
  - *parental depression*

- Other child and parent/family characteristics
  - vocabulary score at end of Head Start
  - *social skills scores at end of Head Start*
  - *mobility*

- Characteristics of kindergarten school and district[4]
  - *public versus private*
  - *magnet/charter versus regular*
  - *highest grade in school*
  - *Title I eligibility*
  - *school size*
  - percentage of minority enrollment
  - percentage of students eligible for free or reduced-price lunch

Among those characteristics that had significantly different unweighted distributions, we then looked at the weighted distributions of the respondents and compared them to those of the *respondents and nonrespondents combined*. We used the child-level pre-kindergarten weight for these weighted distributions. This weight sums to the population of children eligible for the kindergarten data collection as of the spring of the prior program year (that is, children still enrolled in Head Start that prior spring).

If the difference in the weighted distributions was big enough to be meaningful, then it might indicate the possibility of nonresponse bias. This is our key question and is not simply a matter of whether or not the difference is statistically significant. If we statistically test for the difference between the respondents and the respondents plus nonrespondents, we must control for their lack of independence. Once we control for the covariance between the two groups in this statistical comparison, we are in essence comparing the respondents to the nonrespondents, which is not our main concern here. As discussed above, one can have significant differences between the two groups but not have nonresponse bias among respondents that is of a consequential magnitude when the nonresponse rate is low.

Although the threshold for being "big enough" is arbitrary, we attempted to quantify the size of the difference using the following three criteria for judging whether or not we had evidence of nonresponse bias:

1. Absolute difference: $\left|\bar{Y} - \bar{Y}_R\right| > .05$?
2. Relative difference: $\left|\bar{Y} - \bar{Y}_R\right| / \bar{Y} > .05$?
3. Effect size: $\left|\bar{Y} - \bar{Y}_R\right| / \hat{\sigma}_Y > .20$?

We found that none of the differences met the first or third criteria, but some met the second one—the relative difference being larger than five percentage points (Table 5). In this table, we highlight in red font those relative differences that were larger than .05 for one of the three types of response comparisons.

[4]From the Common Core of Data (CCD) and Private School Survey (PSS). These are available only for children whose kindergarten school was identified.

Table 5: Relative Differences Between Weighted Distributions of Respondents and the Full Sample

Variable	Value	All Eligible	All Attempted	All Completes	Attempted-Eligibles/ Eligibles	Completed-Attempted/ Attempted	Completed-Eligibles/ Eligibles
Census Region	Northeast	13.5	14.1	14.8	0.05	0.05	0.10
	Midwest	25.6	26.4	27.9	0.03	0.06	0.09
	South	33.1	32.5	32.5	0.02	0.00	0.02
	West	27.9	27.1	24.9	0.03	0.08	0.11
Urban Program		74.1	73.2	70.3	0.01	0.04	0.05
Program >40% Black or >40% Hispanic		66.0	63.8	62.2	0.03	0.02	0.06
Public School Program		26.8	27.1	29.3	0.01	0.08	0.09
Program has ≤50% English Home Language		27.1	26.4	24.3	0.02	0.08	0.10
Program Enrollment		1292.0	1267.0	1197.0	0.02	0.06	0.07
Percentage Disabled in Program		12.6	12.9	13.3	0.02	0.03	0.06
Teacher Education< Bachelor's degree		56.7	56.4	55.1	0.00	0.02	0.03
Teacher Race	White NH	44.2	46.3	47.8	0.05	0.03	0.08
	Black NH	29.5	28.2	26.9	0.04	0.05	0.09
	Hispanic	21.8	20.8	20.4	0.05	0.02	0.06
	Other	4.6	4.7	4.9	0.02	0.05	0.06
Class Size		17.6	17.6	17.6	0.00	0.00	0.00
Years Teaching Head Start		8.5	8.4	8.4	0.01	0.01	0.02
Pre-K Year	2007	50.9	49.3	47.7	0.03	0.03	0.06
	2008	49.2	50.7	52.3	0.03	0.03	0.06
Child's Race	White NH	24.2	25.3	27.3	0.05	0.08	0.13
	Black NH	30.6	29.8	28.9	0.03	0.03	0.06

Table 5 *(continued)*

Variable	Value	All Eligible	All Attempted	All Completes	Attempted-Eligibles/ Eligibles	Completed-Attempted/ Attempted	Completed-Eligibles/ Eligibles
	Hispanic	35.8	35.3	34.6	0.01	0.02	0.03
	Other	9.4	9.5	9.3	0.01	0.03	0.02
Non-English Home Language		29.5	28.2	27.1	0.05	0.04	0.08
Family Structure	Mother + father	48.3	48.2	50.0	0.00	0.04	0.04
	Mother only	42.6	42.7	40.9	0.00	0.04	0.04
	Other	9.1	9.1	9.1	0.00	0.00	0.00
Parents' Marital Status	Married	35.0	35.2	36.1	0.00	0.03	0.03
	Not married	18.6	18.7	19.6	0.00	0.05	0.05
	Single parent	46.4	46.2	44.4	0.00	0.04	0.04
Parent Highest Education	< HS Dip/GED	32.9	31.7	31.2	0.04	0.01	0.05
	HS Dip/GED	33.6	33.3	33.0	0.01	0.01	0.02
	Some college	25.5	26.8	27.7	0.05	0.03	0.08
	Bachelor's +	8.0	8.3	8.2	0.03	0.01	0.02
Household Income	≤$15,000	36.8	35.9	34.5	0.02	0.04	0.06
	$15,001-$30,000	43.0	42.9	44.0	0.00	0.03	0.02
	>$30,000	20.2	21.3	21.6	0.05	0.02	0.07
Income as Percentage of Poverty	≤100%	58.4	57.3	56.5	0.02	0.01	0.03
	101-185%	29.8	30.5	31.1	0.02	0.02	0.05
	186%+	11.8	12.3	12.4	0.04	0.01	0.05
Family Risk Index	0 risks	12.6	13.8	15.0	0.09	0.09	0.19

Table 5 *(continued)*

Variable	Value	All Eligible	All Attempted	All Completes	Attempted-Eligibles/ Eligibles	Completed-Attempted/ Attempted	Completed-Eligibles/ Eligibles
	1-2 risks	62.5	62.7	62.0	0.00	0.01	0.01
	3-4 risks	24.9	23.5	23.0	0.06	0.02	0.08
Any Child Care		34.8	36.6	36.7	0.05	0.00	0.06
PreK PPVT Score		86.4	86.9	87.3	0.01	0.01	0.01
Baseline PPVT Score		81.1	81.5	81.9	0.01	0.00	0.01
Baseline Social Skills		15.8	15.8	15.8	0.00	0.00	0.00
Baseline Behavior Problems		5.8	5.7	5.7	0.01	0.01	0.01
Percentage of Students Eligible for FRPL			61.0	59.8		0.02	
Percentage Minority Enrollment			59.4	57.1		0.04	

FRPL = free or reduced-price lunch; GED = General Equivalency Diploma; HS = high school; NH = non-Hispanic; PPVT = Peabody Picture Vocabulary Test.

Take, for example, the family risk index, which has the largest relative difference in the table. Almost 13 percent of the full sample of those eligible (or possibly eligible) for kindergarten have zero risk factors. Among those who had a teacher complete, 15 percent have zero risk factors. The relative difference is 19 percent. However, the difference itself (13 versus 15 percent) is still not worrisome, so that even the biggest differences are not all that large.

Neither of the two school-level characteristics that had statistically different unweighted distributions for respondents and nonrespondents[5] had a relative difference in the weighted distributions of 5 percent or higher, indicating that children with responding teachers came from a distribution of schools that resembled that of the full sample in terms of low-income and minority enrollment.

Conclusion. We were able to improve the kindergarten teacher response rate between our first and second years by obtaining the kindergarten information earlier—using information we obtained from the parent interview that was conducted in the spring prior to the kindergarten year—and by improving the school look-up table in our data collection instrument. As one would expect, our respondents and nonrespondents differed on a number of characteristics. However, these differences did not translate into meaningful differences in the profiles of the respondents and the full sample. We take this to mean that, although the response rate for kindergarten teachers was lower than that for Head Start teachers, we see little evidence of bias.

Note that the small differences shown occur before any weighting adjustments for nonresponse have been applied. By using nonresponse adjustment weights within Head Start center and age cohorts, which we have done, we attempt to decrease any potential bias, no matter how small. Comparable tables showing distributions using nonresponse-adjusted weights would make the minimal differences we have shown even smaller.

Limitations. We acknowledge that these differences are only for observed characteristics of children, their families, Head Start programs, and schools. We have no kindergarten teacher characteristics to compare. However, the kindergarten teacher survey data are rarely used alone. They are often used with data from the direct child assessment or parent interview[6] to describe the total kindergarten experience of children who attended Head Start before entering kindergarten.

Recommendations. It is important to start identifying schools and teachers early, and parents are a reliable source of information about where their children will attend kindergarten. We saw big improvement in kindergarten teacher response rates between spring 2008 and spring 2009. Although we lost some teachers because of incorrect information, this was less of an issue in the second year due to improvements in the parent interview look-up table.

However, the 68 percent response rate in the second year is still not good enough. Because our ability to collection data from teachers is contingent on parent participation, we need to find ways to increase the parent response rate, which was lower in the kindergarten year, and to find more ways to help parents better identify their child's

[5] Percentage of minority enrollment and percentage of students eligible for free or reduced-price lunches.

[6] Obtained independently from the kindergarten teacher data collection.

school and teacher. And we need better incentives for the kindergarten teachers, once identified, to respond. We are exploring the possibility of conducting an incentive experiment in FACES 2009 that would include a comparison of a pre- versus post-payment of the respondent incentive.