Head Start Family and Child Experiences Survey 2019 (FACES 2019) OMB Supporting Statement for Data Collection

OMB Information Collection Request
0970 - 0151

Supporting Statement
Part B

June 2021

Submitted by:
Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services

4th Floor, Mary E. Switzer Building
330 C Street, SW
Washington, DC 20201

Project Officers: Nina Philipsen and Meryl Barofsky

Content

S

Part B

B1. Objectives

Study Objectives

The Head Start Family and Child Experiences Survey (FACES) produces data on a set of key indicators in Head Start Regions I–X (FACES 2019) and Region XI (AIAN FACES 2019). The current request outlines activities to understand the needs of children and families a year and a half after the start of the COVID-19 (for coronavirus disease 2019) pandemic. For information about previous FACES information collection requests, see: https://www.reginfo.gov/public/do/PRAOMBHistory?ombControlNumber=0970-0151.

A spring 2022 information collection is also planned for both FACES and AIAN FACES. More information about the spring 2022 data collection will be provided, along with the specific data collection elements, in a subsequent request package.

In this request, we detail the sampling plans for the fall 2021 data collection activities, including selecting classrooms and children for the study and getting consent for children, conducting data collection, data analyses, and the reporting of study findings. As outlined in Supporting Statement Part A, ACF is requesting this additional fall 2021 wave of data collection to understand how children, families, and staff are faring in light of the COVID-19 pandemic.

Generalizability of Results

FACES 2019 and AIAN FACES 2019 are intended to produce nationally representative estimates of Head Start programs, their associated staff, and the families they serve. The results are generalizable to the Head Start program as a whole, with a few limitations. Head Start programs in U.S. territories are excluded, as are programs under the direction of ACF Region XII (Migrant and Seasonal Worker Head Start) and those under transitional management. Programs that are administrative only and do not directly provide services, those that have no center-based classrooms, and Early Head Start programs are also excluded from this study. These limitations will be clearly stated in published results.

Appropriateness of Study Design and Methods for Planned Uses

FACES 2019 and AIAN FACES 2019 were primarily designed to answer important questions to inform technical assistance and program planning, and questions of interest to the research community. The studies' logic models (Appendix U) guided the overall design of the studies. The AIAN FACES 2015 Workgroup also informed AIAN FACES 2019's study design. The studies' samples are designed so that the resulting weighted estimates are unbiased, sufficiently precise, and have adequate power to detect relevant differences at the national level.

The fall 2021 topical focus on the well-being of children, families, and staff is reflected in the proposed questionnaires, which are designed to capture information on current well-being, especially in light of the COVID-19 pandemic. The AIAN FACES 2019 questionnaires also reflect the importance of culture— underscored in the logic model (Appendix U) —in understanding Native culture, language use, and community supports. We will archive restricted-use data from FACES 2019 and AIAN FACES 2019 for

secondary data analysis by researchers interested in exploring non-experimental associations between children's Head Start experiences and child and family well-being, and the logic models will be included in documentation to support responsible secondary data use.

As noted in Supporting Statement Part A, this information is not intended to be used as the principal basis for public policy decisions, and it is not expected to meet the threshold of influential or highly influential scientific information.

B2.　Methods and Design

Target Population

The target population for FACES 2019 and AIAN FACES 2019, including the fall 2021 data collection, is all Region I through XI Head Start programs in the United States (in all 50 states plus the District of Columbia), their classrooms, and the children and families they serve. The study team plans to freshen the FACES 2019 program sample for fall 2021 by adding a few programs that came into being since the original sample was selected; and for continuing programs, the study team plans to reselect centers when one or both selected centers have closed—otherwise, the study team will keep the originally sampled centers in the sample. For AIAN FACES 2019, the study team does not plan to freshen the program sample for fall 2021, but does plan to select new center samples for all programs. In both studies, the study team will select new samples of classrooms and children for fall 2021.

The sample designs for fall 2019 were similar to the ones used for FACES 2014 and AIAN FACES in 2015. FACES 2019 and AIAN FACES 2019 used a stratified multistage sample design with four stages of sample selection: (1) Head Start programs, with "programs" defined as grantees or delegate agencies providing direct services; (2) centers within programs; (3) classes within centers; and (4) for a random subsample of programs (in FACES) or for all programs (in AIAN FACES), children within classes. To minimize the burden on parents or guardians who have more than one child selected for the sample, the study team also randomly subsampled one selected child per parent or guardian, a step that was introduced in FACES 2009. The study team used the Head Start Program Information Report (PIR) as the sampling frame for selecting FACES 2019 and AIAN FACES 2019 programs. This file, which contains program-level data as reported by the programs themselves, is updated annually, and the study team used the latest available PIR at the time of sampling each of the two studies. For later sampling stages, the study team obtained lists of centers from the sampled programs, and lists of classrooms and rosters of children from the sampled centers. For fall 2021, when members of the study team freshen the program sample for FACES, they will likely have to get updated program lists directly from the Office of Head Start (OHS), as the PIR data collections have been disrupted by the COVID-19 pandemic.

Sampling and Site Selection

Sample design

FACES 2019. The freshening of the program sample for fall 2021 will use well-established methods to ensure that the refreshed sample can be treated as a valid probability sample. The fall 2021 sample for FACES 2019 will primarily use the programs selected for fall 2019 data collection, but it will be freshened by selecting a small number of newer programs (those that had no chance of selection for fall 2019). The study team will freshen the centers in continuing programs if either of the originally sampled centers has closed. For all other programs, the study team will select a new sample of centers.

Because of the disruption of typical Head Start instruction and other services in response to the COVID-19 pandemic, defining classrooms the way the study has in the past (a group of children taught together by the same teacher) may be difficult in fall 2021, given the various virtual and hybrid instructional scenarios that may still be in place. To get around this issue, the study team plans to sample teachers instead of classrooms for fall 2021 data collection, and then select from the children instructed by the teacher. Because of the uncertainty about parental consent rates under the COVID-19 pandemic situation in fall 2021, the study team will treat the nonsampled children instructed by the sampled teachers as a potential back-up sample to release into the study to achieve study targets. The study team will use a process like the one carried out for fall 2019 data collection to select new samples of teachers and children.

The sample design for the FACES 2019 study is based on the one used for FACES 2014, which in turn was based on the designs of the five previous studies. Like the design for FACES 2014, the sample design for FACES 2019 involved sampling for two study components: the Classroom + Child Outcomes Study and the Classroom Study. The Classroom + Child Outcomes Study involved sampling at all four stages (programs, centers, classrooms, and children) and the Classroom Study involved sampling at the first three stages only (excluding sampling of children within classes). The sample design for the fall 2021 data collection maps to that of the Classroom + Child Outcomes Study. The study team does not describe the Classroom Study further in this submission.

Proposed sample sizes were determined with the goal of achieving accurate estimates of characteristics at the child level given various assumptions about the validity and reliability of the selected measurement tools with the sample, the expected variance of key variables, the expected effect size of group differences, and the sample design and its impact on estimates. At the child level, the study team will collect information from surveys administered to parents and teachers. To determine appropriate sample sizes at each nested level, the study team explored thresholds at which sample sizes could support answering the study's primary research questions. The study's research questions determined both the characteristics the study team aims to describe about children and the subgroup differences they expected to detect. The study team selected sample sizes appropriate for point estimates with adequate precision to describe the key characteristics at the child level. Furthermore, the study team selected sample sizes appropriate for detection of differences between subgroups of interest (if differences exist) on key variables. The study team evaluated the expected precision of these estimates after accounting for sample design complexities. (See section on "Degree of Accuracy" below.) The study team will work to achieve comparable sample sizes in fall 2021.

Like the ones for FACES 2014 and the fall 2019 data collection in FACES 2019, the fall 2021 child level sample will represent children enrolled in Head Start for the first time and those who are attending a second year of Head Start. This will allow for a direct comparison of first- and second-year program participants.

To minimize the effects of unequal weighting on the variance of estimates, the study team will sample with probability proportional to size (PPS) in the first two stages. At the third stage, the study team selects an equal probability sample of teachers within each sampled center and an equal probability sample of children within each sampled classroom. The measure of size for PPS sampling in each of the first two stages is the number of classrooms. This sampling approach maximizes the precision of teacher-level estimates and allows for easier sampling of teachers and children within teachers. The study team targeted 60 programs for fall 2019 data collection for Regions I–X. Within these 60 programs, the study team
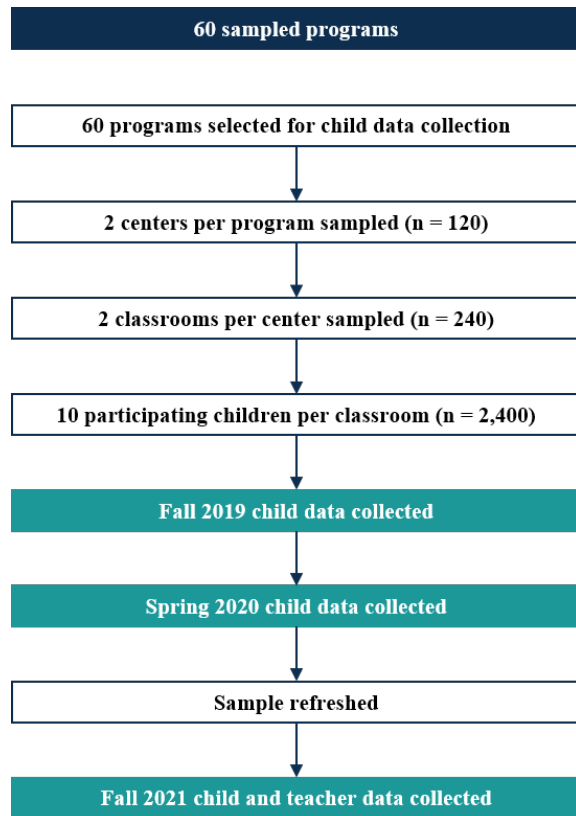
selected, if possible, two centers per program, two classes per center, and enough children to yield 10 consented children per class, for a total of about 2,400 children in fall 2019.

For follow-up data collection in fall 2021, the study team will select a refresher sample of programs and their centers so that the new sample will be representative of all programs and centers in Regions I–X at the time of follow-up data collection, and the study team will select a new sample of teachers and children in all centers. As part of the sample freshening, the study team plans to use as many as possible of the original 59 FACES 2019 programs that participated in fall 2019 in the fall 2021 data collection, but will supplement with a small sample of programs that came into being since the original program sample was selected. For the original programs that will participate again in 2021, the study team will also use their originally sampled centers in situations where their sampled centers are still providing Head Start services. Otherwise, the study team will sample new centers. In all centers, the study team will select new samples of teachers (two per center) and children (12 per classroom). Figure B.1 is a diagram of the sample selection procedures. At both the program and sampling stages, the study team uses a sequential sampling technique based on a procedure developed by Chromy.[1] The study team uses a systematic sampling technique for selecting classrooms (or teachers) and children.

For fall 2019, the study team initially selected double the target number of programs, and paired similar selected programs within strata. The study team then randomly selected one from each pair to be released as part of the main sample of programs. After the initially released programs were selected, the study team asked OHS to confirm that these programs were in good standing—that is, they had not lost their Head Start grants, nor were they in imminent danger of losing their grants. If confirmed, each program was contacted and recruited to participate in the study: the 60 programs sampled for the Classroom + Child Outcomes Study were recruited in spring 2019 (to begin participation in fall 2019). If the program was not in good standing, or refused to participate, the study team released into the sample the other member of the program's pair and went through the same process of confirmation and recruitment with that program. For fall 2021, the study team will go through the same OHS confirmation process for the continuing programs and for any new programs in the sample that are added as part of the freshening process. All released programs will be accounted for as part of the sample for purposes of calculating response rates and weighting adjustments. At subsequent stages of sampling, the study team will release all sampled cases, expecting full study participation from the selected centers and teachers. At the child level, the study team estimates that out of 12 selected children per teacher, the study team will end up with our target of 10 eligible children with parental consent. The study team expects to lose, on average, 2 children per teacher because they are no longer enrolled, because parental consent was not granted, or because of the subsampling of selected siblings. In fall 2021, the study team will consider the nonsampled children in each sampled teacher a potential backup sample to release if consent rates are lower than the study team has historically achieved.

---

[1] The procedure offers all the advantages of the systematic sampling approach, but eliminates the risk of bias associated with that approach. The procedure makes independent selections within each of the sampling intervals while controlling the selection opportunities for units crossing interval boundaries. See Chromy, J. R. "Sequential Sample Selection Methods." *Proceedings of the Survey Research Methods Section of the American Statistical Association,* Alexandria, VA: American Statistical Association, 1979, pp. 401–406.

Figure B.1. Flow of sample selection procedures for FACES 2019, fall 2021

60 sampled programs

60 programs selected for child data collection

2 centers per program sampled (n = 120)

2 classrooms per center sampled (n = 240)

10 participating children per classroom (n = 2,400)

Fall 2019 child data collected

Spring 2020 child data collected

Sample refreshed

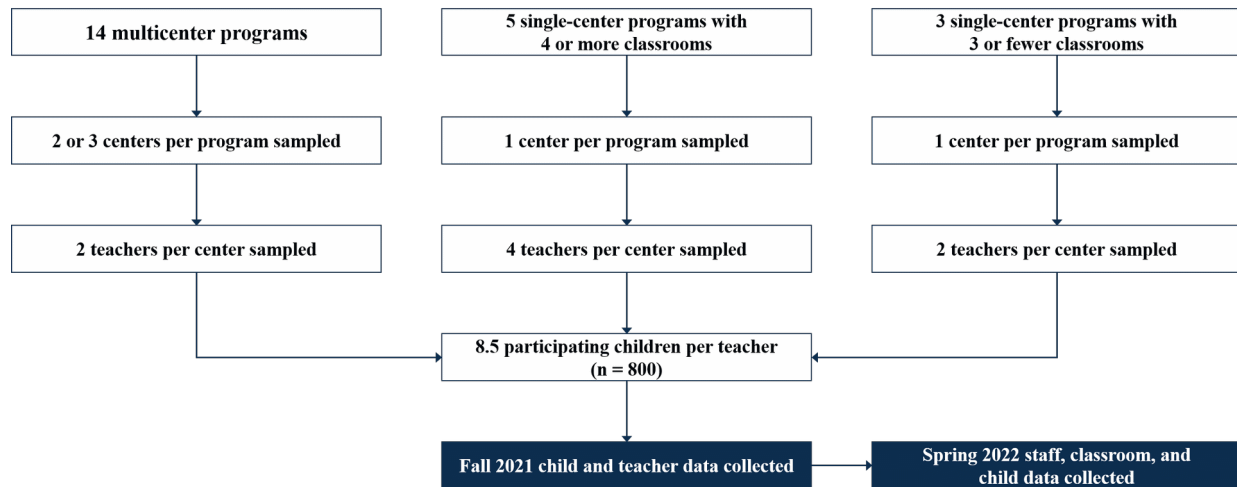| Fall 2021 child and teacher data collected |
|:---:|

In programs for which the study team will be selecting a new sample of centers, the team will use the Chromy procedure again to select them with PPS within each sampled program, using the number of classrooms as the measure of size. The study team will randomly select teachers and home visitors within centers with equal probability. The study team will group teachers whose classrooms have only a few children with other teachers in the same center for sampling purposes to ensure a large enough sample yield. Once teachers are selected, the study team will select an equal probability sample of 12 children per teacher, with the expectation that 10 children will be eligible and receive parental consent.

AIAN FACES 2019. In fall 2021, the study team plans to keep as many of the original 22 programs in the sample as possible. The study team does not plan to freshen the AIAN FACES 2019 program sample for fall 2021 the way they plan to do for FACES 2019. It is possible that not all programs will continue in the study. Therefore, to achieve our goal of 800 children in the study sample, the study team will select new samples of centers (up to three per program), teachers (two or four per center), and children (13 per teacher). Because of the uncertainty about parental consent rates in light of the COVID-19 pandemic, the study team will treat the nonsampled children instructed by the sampled teachers as a potential backup sample to release into the study to achieve study targets as needed. Within the 14 sampled programs with two or more centers, the intent of the design is to select up to three centers per program, two teachers per center, and enough children to yield an average of 8.5 consented children per teacher. Within the 8 sampled programs with one center, the study team will select up to four teachers and enough children to yield an average of 8.5 consented children per teacher. We anticipate a total of 42 centers and 90 teachers. At the child level, the study team estimates that if 13 children are selected per teacher, the result would be an average of 8.5 eligible children with parental consent, which is our target. The study team expects to lose, on average, 4.5 children per teacher, whether because they are no longer enrolled, because parental

consent was not granted, or because of the subsampling of selected siblings. Figure B.2 is a diagram of the sample selection procedures. The primary stratification of the program sample is based on program structure (number of centers per program, and number of classrooms per single-center program, as detailed in the next section).

Figure B.2. Flow of sample selection procedures for AIAN FACES 2019 fall 2021

14 multicenter programs	5 single-center programs with 4 or more classrooms	3 single-center programs with 3 or fewer classrooms
2 or 3 centers per program sampled	1 center per program sampled	1 center per program sampled
2 teachers per center sampled	4 teachers per center sampled	2 teachers per center sampled

8.5 participating children per teacher
(n = 800)

Fall 2021 child and teacher data collected → Spring 2022 staff, classroom, and child data collected

Statistical methodology for stratification and sample selection

The sampling methodology is described above. When sampling programs from Regions I–X, the study team formed explicit strata using census region, metro/nonmetro status, and percentage of racial/ethnic minority enrollment. The new programs eligible for the program sample freshening will form their own sampling stratum. When sampling programs from Region XI, the study team formed explicit strata based on geographic area for programs that had two or more centers; for those with only one center, the study team stratified by whether the program had four or more classrooms. Sample allocation was proportional to the estimated fraction of eligible classrooms represented by the programs in each stratum. For Regions I–X, the study team implicitly stratified (sorted) the sample frame by other characteristics, such as percentage of children who are dual language learners (DLLs) (categorized), whether the program is a public school district grantee, and the percentage of children with disabilities. No explicit stratification was used to select centers within programs, classes within centers, or children within classes, although some implicit stratification (such as the percentage of children who are dual language learners) was used to select centers. When freshening the FACES program sample for fall 2021, the study team will create a new stratum for programs that did not have a chance of selection for the fall 2019 program sample. For Region XI, the study team implicitly stratified (sorted) the sample frame by program state and by the proportion of enrolled children who were AIAN. No explicit stratification was used to select centers within programs, classes within centers, or children within classes, although implicit stratification by the percentage of children who are dual language learners was used to select centers.

Estimation procedure

For the fall 2021 data collection for both FACES 2019 and AIAN FACES 2019, the study team will create analysis weights to account for variations in the probabilities of selection and variations in the eligibility and cooperation rates among those selected. For each stage of sampling (program, center, teacher, and child) and within each explicit sampling stratum, the study team will calculate the probability

of selection. The inverse of the probability of selection within stratum at each stage is the sampling or base weight. The sampling weight takes into account the PPS sampling approach, the presence of any certainty selections, and the actual number of cases released. The study team will treat the eligibility status of each sampled unit as known at each stage. Then, at each stage, the study team will multiply the sampling weight by the inverse of the weighted response rate within weighting cells (defined by sampling stratum) to obtain the analysis weight, so that the respondents' analysis weights account for both the respondents and nonrespondents. This will be done for a variety of weights, each with different definitions of "respondent" for fall 2021 data, using different combinations of completed instruments.

Thus, the program-level weight for fall 2021 will adjust for the probability of selection of the program and study participation at the program level; the center-level weight will adjust for the probability of center selection and center-level participation; and the teacher-level weight will adjust for the probability of selection of the teacher and teacher-level study participation and survey response. For the FACES 2019 and AIAN FACES 2019 fall 2021 weights, the study team will then adjust for the probability of selection of the child instructed by a given teacher, whether parental consent is obtained, and whether various child-level instruments (Teacher Child Report and parent surveys) are obtained. The formulas below represent the various weighting steps for the cumulative weights through prior stages of selection, where P represents the probability of selection, and RR the response rate at that stage of selection.

$$W_{pgm} = \frac{1}{P_{pgm}} \cdot \frac{1}{RR_{pgm}}$$

$$W_{center} = W_{pgm} \cdot \frac{1}{P_{center}} \cdot \frac{1}{RR_{center}}$$

$$W_{class} = W_{center} \cdot \frac{1}{P_{class}} \cdot \frac{1}{RR_{class}}$$

$$W_{child} = W_{class} \cdot \frac{1}{P_{pgm-subsample}} \cdot \frac{1}{P_{child}} \cdot \frac{1}{RR_{child}}$$

Degree of accuracy needed to address the study's primary research questions

The complex sampling plan, which includes several stages, stratification, clustering, and unequal probabilities of selection, requires specialized procedures to calculate the variance of estimates. Standard statistical software assumes independent and identically distributed samples, which would indeed be the case with a simple random sample. A complex sample, however, generally has larger variances than would be calculated with standard software. Two approaches for estimating variances under complex sampling, Taylor Series and replication methods, can be estimated by using SUDAAN statistical software and special procedures in SAS, Stata, and other packages.

All of the analyses for the fall 2021 data collection for FACES 2019 will be at the child and teacher levels. AIAN FACES 2019 is designed for child-level estimates only. Given various assumptions about the validity and reliability of the selected measurement tools with the sample, the hypothesized variation expected on key variables, the expected effect size of group differences, and the sample design and its impact on estimates, the sample size should be large enough to detect meaningful differences. In Tables

B.1 and B.2 (for Regions I–X and Region XI, respectively), we show the minimum detectable differences with 80 percent power (and alpha = 0.05) and various sample and subgroup sizes, assuming different intraclass correlation coefficients for classroom- (where applicable) and child-level estimates at the various stages of clustering (see table footnote).

For point-in-time estimates, the study team is making the conservative assumption that there is no covariance between estimates for two subgroups, even though the observations may be in the same classes, centers, and/or programs. By conservative, we mean that smaller differences than those shown will likely be detectable.

Tables B.1 and B.2 show the minimum differences that would be detectable for point-in-time (cross-sectional) estimates at the teacher (where applicable) and child levels. The study team has incorporated the design effect attributable to clustering. We show minimum detectable differences between point-in-time child subgroups defined two different ways: (1) assuming the subgroup is defined by program-level characteristics, and (2) assuming the subgroup is defined by child-level characteristics (which reduces the clustering effect in each subgroup). Next, we give examples.

The columns farthest to the left ("Subgroups") show several sample subgroup proportions (for example, a comparison of male children to female children would be represented by "50, 50"). The child-level estimates represent all consented children in fall 2021 who will complete an instrument (for FACES, n = 2,040). For example, the n = 2,040 row within the "33, 67" section represents a subgroup comparison involving child-level respondents at the beginning of data collection for two subgroups, one representing one-third of that sample (for example, children in bilingual homes), the other representing the remaining two-thirds (for example, children from homes where English is the only language used).

The last few columns ("minimum detectable difference") show different types of variables from which an estimate might be made; the first two are estimates in the form of proportions, and the last shows the minimum detectable effect size—the MDD in standard deviation-sized units. The numbers for a given row and column show the minimum underlying differences between the two subgroups that would be detectable for a given type of variable with the given sample size and design assumptions. The MDD numbers given in Table B.1 assume that 85 percent of sampled classrooms will have completed teacher surveys. Similarly, both Tables B.1 and B.2 assume that 85 percent of sampled children who have parental consent will have a completed parent survey, and 85 percent will have a completed Teacher Child Report. The assumptions for AIAN FACES (Table B.2) are what we believe is *likely* – about 20 participating programs, and about 642 children with responses. We leave the burden estimate (Table A.8) at 22 programs and 800 children because that is how many we will reach out to and could participate—reflecting the maximum burden expected. We use the 20 programs and 642 children for the precision calculations so we have a more realistic estimate of what those MDDs will actually be.

Table B.1. FACES 2019 minimum detectable differences: Regions I–X (Fall 2021)

Teacher Subgroups				Minimum detectable difference		
Percentage in Group 1	Percentage in Group 2	Teachers in Group 1	Teachers in Group 2	Proportion of 0.1 or 0.9	Proportion of 0.5	Minimum detectable effect size[a]
50	50	102	102	.151	.252	.499
33	67	67	137	.161	.268	.531
15	85	31	173	.215	.358	.699
Child subgroups				Minimum detectable difference (Program-defined subgroups / Child-defined subgroups)		
Percentage in Group 1	Percentage in Group 2	Children in Group 1	Children in Group 2	Proportion of 0.1 or 0.9	Proportion of 0.5	Minimum detectable effect size
50	50	1,020	1,020	0.091/0.068	0.152/0.114	0.304/0.228
33	67	673	1,367	0.097/0.069	0.162/0.116	0.323/0.231
40	30	816	612	0.112/0.072	0.184/0.120	0.367/0.239

Note:　Conservative assumption of no covariance for subgroup comparisons. Assumes =.05 (two-sided), .80 power. For teacher-level estimates, assumes 60 programs, 120 centers, between-program ICC = .2, between-center ICC = .2, and an 85 percent teacher survey completion rate. For child-level estimates, assumes 60 programs, 120 centers, between-program ICC = .10, between-center ICC = .05, between-teacher ICC = .12, and an 85 percent instrument completion rate.

[a]The minimum detectable effect size is the minimum detectable difference in standard deviation-sized units.

Table B.2. AIAN FACES 2019 minimum detectable differences: Region XI (Fall 2021)

Child subgroups				Minimum detectable difference (Program-defined subgroups / Child-defined subgroups)		
Percentage in Group 1	Percentage in Group 2	Children in Group 1	Children in Group 2	Proportion of 0.1 or 0.9	Proportion of 0.5	Minimum detectable effect size[a]
50	50	321	321	0.158/0.118	0.263/0.197	0.522/0.392
33	67	212	431	0.168/0.120	0.280/0.200	0.555/0.398
40	30	257	193	0.192/0.124	0.319/0.207	0.630/0.413

Note:　Conservative assumption of no covariance for point-in-time subgroup comparisons. Covariance adjustment made for pre-post difference (Kish 1995, p. 462, Table 12.4.II, Difference with Partial Overlap). Assumes = .05 (two-sided), .80 power. Assumes 20 programs, 42 centers, 89 classrooms, between-program ICC = .10, between-center ICC = .05, between-classroom ICC = .12, and an 85 percent instrument completion rate.

[a]The minimum detectable effect size is the minimum detectable difference in standard deviation-sized units.

If we were to compare two equal-sized subgroups of the 204 classrooms with completed teacher surveys in Regions I–X (FACES) in fall 2021, our design would allow us to detect a minimum difference of .499 standard deviations with 80 percent power. At the child level, if we were to compare an outcome of around 50 percent with a sample size of 2,040 children in fall 2021, and two approximately equal-sized

child-defined subgroups (such as male and female), our design would allow us to detect a minimum difference of 11.4 percentage points with 80 percent power.

The main purpose of AIAN FACES 2019 is to provide descriptive statistics for this population of children. Comparisons between child subgroups are a secondary purpose, given the smaller sample size. If we were to compare an outcome of around 50 percent for two equal-sized subgroups (say, male and female) of the 642 children in Region XI (AIAN) with responses in fall 2021, our design would allow us to detect a minimum difference of 19.7 percentage points with 80 percent power.

B3.   Design of Data Collection Instruments

Development of Data Collection Instruments

The FACES 2019 and AIAN FACES 2019 data collection instruments are based on their respective logic models (presented in Appendix U), which were developed through expert consultation and coordination between ACF and contracted study team[2] to ensure the data's relevance to policy and the research field. AIAN FACES 2019 surveys were also developed in consultation with the AIAN FACES Workgroup.

The data collection protocol for fall 2021 includes surveys of parents and teachers as well as teacher reports of child development. The surveys used the fall 2019-spring 2020 instruments as the foundation to continue to capture information on child and family characteristics, family resources and needs, family and teacher well-being, and teachers' perspectives on children's development. The fall 2021 surveys have been updated to gather more information on family and teacher well-being in relation to the COVID-19 pandemic and parents' perspectives on children's development. Together this information will allow the data to be instrumental in addressing research questions about who is participating in Head Start and about the associations between factors such as social supports, family well-being, and child development.

Wherever possible, the surveys use established scales with known validity and reliability. When there were not enough existing instruments we could use to measure the constructs of interest, we reviewed surveys or new items (notably those studying experiences related to COVID-19) to consider for inclusion. This work fills a gap in the knowledge base about how the population attending Head Start has changed since the onset of the COVID-19 pandemic, and about the supports that population might need. Appendix Q provides the FACES and AIAN FACES instrument content matrices.

Ahead of the data collection, the study team will pretest the parent survey and teacher survey with no more than 9 respondents each to test for question clarity and flow in the content related to family and staff well-being. The study team will also pretest the Teacher Child Report with no more than 9 respondents for its suitability given COVID-related changes to teaching.

B4.   Collection of Data and Quality Control

All data will be collected remotely by the contractor. Modes for all instruments are detailed in Table B.3.

Table B.3. FACES 2019 and AIAN FACES 2019: fall 2021 data collection activities

Component	Administration characteristics	Fall 2021
Parent survey	Mode	CATI/web
	Time	35 minutes

---

[2] ACF has contracted with Mathematica to carry out this information collection.

Component	Administration characteristics	Fall 2021
	Token of Appreciation	$30
Teacher Child Report	Mode	Paper/web SAQ
	Time	10 minutes per child
	Token of Appreciation	$10 per child
Teacher survey	Mode	Paper/Web SAQ
	Time	10 minutes
	Token of Appreciation	n.a.

CATI = computer-assisted telephone interviewing; SAQ = self-administered questionnaire.

n.a. = not applicable

The fall 2021 wave of FACES 2019 and AIAN FACES 2019 will deploy monitoring and quality control protocols that were developed for and used effectively during previous waves of the study.

Recruitment protocol

Following OMB approval, the study team will send programs that were sampled and recruited for the fall 2019 round of data collection an invitation to participate in this special data collection. The invitation will include an official request from OHS, along with a memo from Mathematica and fact sheet about the study. It will fully inform program directors about the continuation of the study, the assistance the study team will need to recruit and sample families, and our planned data collection activities (Appendix R). Should any sampled program be ineligible or decline to participate, the study team will release its replacement program and repeat the program recruitment process (for FACES 2019 fall 2021; the study team will not have replacement programs in AIAN FACES 2019 fall 2021). For AIAN FACES 2019 fall 2021 collection, the study team anticipates that programs will require continued approval from a governing body. If necessary, study staff will use a template (Appendix T) to create a presentation for each program. Each AIAN FACES program also will acknowledge an Agreement of Collaboration and Participation (Appendix S).

Additionally, the study team will send correspondence to remind Head Start staff and parents about upcoming surveys (Appendix V for FACES 2019 Fall 2021 Special Respondent Materials; Appendix W for AIAN FACES 2019 Fall 2021 Special Respondent Materials). The web administration of Head Start teacher and parent surveys will allow the respondents to complete the surveys at their convenience. The study team will ensure that the language of the text in study forms and instruments is at a comfortable reading level for respondents. Paper-and-pencil survey options will be available for teachers who have no computer or Internet access, and parent surveys can be completed via computer or by telephone. Computer-assisted telephone interviewing (CATI) staff will be trained on refusal conversion techniques. AIAN FACES 2019 CATI staff will also be trained on cultural awareness for fall 2021 data collection.

Monitoring telephone interviews

For the parent telephone interview, professional Mathematica Survey Operation Center monitors will monitor the telephone interviewers and observe all aspects of the interviewers' administration—from dialing through completion. Each interviewer will have his or her first interview monitored and will receive feedback. For ongoing quality assurance, over the course of data collection the study team will monitor 10 percent of the telephone interviews. Monitors will also do heavier monitoring of interviewers who have had issues requiring correction during previous monitoring sessions. In these situations,

monitors will ensure that interviewers are conducting the interview as trained, provide feedback again as needed, and, if necessary, determine whether the interviewer should be removed from the study.

Monitoring the web instruments

For each web instrument, the study team will review completed surveys for missing responses, review partial surveys to determine whether they need follow-up with respondents, and conduct a preliminary data review after the first 10–20 completions to confirm that the web program is working as expected and to check for inconsistencies in the data. The web surveys will be programmed to include soft checks to alert respondents to potential inconsistencies while they are responding.

Monitoring the response rates

The study team will use reports generated from the sample management system and web instruments to actively monitor response rates for each instrument by program and center. The reports will give study team members up-to-date information on the progress of data collection, including response rates, allowing the team to quickly identify challenges and implement solutions to achieve the expected response rates.

B5.   Response Rates and Potential Nonresponse Bias

Response Rates

There is an established record of success in gaining program cooperation and obtaining high response rates with center staff, children, and families in research studies of Head Start, Early Head Start, and other preschool programs. To achieve high response rates for the fall 2021 data collections, the study team will continue to use the procedures that have worked well on prior FACES studies, such as offering multiple modes for survey completion, sending e-mail and hard-copy reminders, and providing tokens of appreciation. In fall 2019, FACES 2019 marginal unweighted response rates ranged from 75 to 93 percent, and AIAN FACES 2019 response rates ranged from 75 to 88 percent. (Spring 2020 response rates were somewhat lower for the instruments fielded in the early months of the COVID-19 pandemic.)

The recruitment approaches that have already been described, most of which have been used in prior FACES studies, will help ensure a high level of participation. For FACES 2019, 79 percent of eligible programs participated for fall 2019 data collection, and for AIAN FACES 2019, the study team recruited 54 percent of the eligible programs. At the child level, the study team obtained parental consent from 91 percent of children sampled for FACES 2019 and from 75 percent of children sampled for AIAN FACES 2019. Obtaining the expected high response rates reduces the potential for nonresponse bias, making any estimates from FACES 2019 or AIAN FACES 2019 data more generalizable to the Head Start population. The study team will calculate both unweighted and weighted and marginal and cumulative response rates at each stage of sampling and data collection. Following the American Association for Public Opinion Research (AAPOR 2016) industry standard for calculating response rates, the numerator of each response rate will include the number of eligible completed cases. We define a completed case as one in which all critical items for inclusion in the analysis are complete and within valid ranges. The denominator will include the number of eligible selected cases. Table B.4 summarizes the FACES 2019 and AIAN FACES 2019 response rates for fall 2019 and spring 2020. The study team expects the response rates in fall 2021 to be the same as those achieved in fall 2019, with the exception of the child assessment, which is not fielding this fall. The study team expects an 85 percent response rate for the special fall teacher survey.

Table B.4. Final response rates for fall 2019 and spring 2020 approved information requests

Data collection component	Expected response rate	Final response rate
FACES 2019		
FACES parent consent form (with consent given)	90%	91%
FACES fall child assessment	85%	93%
FACES fall parent survey	85%	75%
FACES fall Teacher Child Report	85%	92%
FACES spring program director survey	85%	76%
FACES spring center director survey	85%	98%
FACES spring classroom observation	100%	n.a.[a]
FACES spring teacher survey	85%	62%
FACES spring child assessment	85%	n.a.[a]
FACES spring parent survey	85%	68%
FACES spring Teacher Child Report	85%	70%
AIAN FACES 2019		
AIAN FACES parent consent form (with consent given)	90%	75%
AIAN FACES fall child assessment	85%	86%
AIAN FACES fall parent survey	85%	75%
AIAN FACES fall Teacher Child Report	85%	88%
AIAN FACES spring program director survey	85%	82%
AIAN FACES spring center director survey	85%	68%
AIAN FACES spring classroom observation	100%	35%[a]
AIAN FACES spring teacher survey	85%	69%
AIAN FACES spring child assessment	85%	32%[a]
AIAN FACES spring parent survey	85%	67%
AIAN FACES spring Teacher Child Report	85%	70%

[a]In spring 2020, the classroom observation and child assessment were not fielded in FACES 2019 due to the COVID-19 pandemic. In AIAN FACES 2019, these instruments were conducted in 7 of the AIAN FACES programs before fielding stopped in response to the COVID-19 pandemic.

n.a. = not applicable.

Nonresponse

Once data collection is complete, the study team will create a set of nonresponse-adjusted weights to use for creating survey estimates and to minimize the risk of nonresponse bias. The weights will build on sampling weights that account for differential selection probabilities as well as nonresponse at each stage of sampling, recruitment, and data collection. Each weight will define a "respondent" based on a particular combination of instruments being completed (for example, a parent survey and a Teacher Child Report). When marginal response rates[3] are low (below 80 percent) the study team plans to conduct a nonresponse bias analysis to compare distributions of program- and child-level characteristics, compare the characteristics of respondents to those of nonrespondents, and then compare the distributions for respondents when weighted using the nonresponse-adjusted weights to see if any observed differences

---

[3] "Marginal response rate" is used here to mean the response rate among those for whom an attempt was made to complete the instrument; and does not account for any study nonparticipation in prior stages of sampling.

appear to have been mitigated by the weights. Program-level characteristics are available from the PIR. Characteristics measured for children will be limited to what is collected on the lists or rosters used for sampling. Item nonresponse tends to be low (for example on both FACES 2019 and AIAN FACES 2019, data on key children's characteristics such as race/ethnicity, age, and sex are present for all sample members or missing in less than 1 percent of cases), with the exception of the household income variables on the parent survey, which typically have an item nonresponse rate of between 20 to 30 percent. The study teams plan to impute income, as has been done in the past on FACES.

B6.   Production of Estimates and Projections

All analyses will be run using the final analysis weights, so that the estimates can be generalized to the target population. Documentation for the restricted use analytic files will include instructions, descriptive tables, and coding examples to support the proper use of weights and variance estimation by secondary analysts.

B7.   Data Handling and Analysis

Data Handling

Once the electronic instruments are programmed, Mathematica uses a random data generator (RDG) to check questionnaire skip logic, validations, and question properties. The RDG produces a test data set of randomly generated survey responses. The process runs all programmed script code and follows all skip logic included in the questionnaire, simulating real interviews. This process allows any coding errors to be addressed before data collection.

During and after data collection, Mathematica staff responsible for each instrument will make edits to the data when necessary. The survey team will develop a document for data editing to identify when survey staff select a variable for editing—documenting the current value, the new value, and the reason for the edit. A programmer will read the specifications from these documents and update the data file. All data edits will be documented and saved in a designated file. The study team expects that most data edits will correct interviewer coding errors identified during frequency review (for example, filling missing data with "M" or clearing out "other specify" verbatim data when the response has been back-coded). This process will continue until all data are clean for each instrument.

Data Analysis

The analyses will aim to (1) describe children and families participating in Head Start in fall 2021 and over time, (2) relate family characteristics (such as resources and supports) to family well-being and children's development, and (3) describe Head Start teachers. Analyses will employ a variety of methods, including descriptive statistics (means, percentages), simple tests of differences over time and, for FACES 2019, across subgroups (t-tests, chi-square tests), multivariate analysis (regression analysis, hierarchical linear modeling [HLM]), and trend analysis (t-tests, chi-square tests, regression analysis). For all analyses, the study team will calculate standard errors that take into account multilevel sampling and clustering at each level (program, center, classroom, child), as well as the effects of unequal weighting. The study team will use analysis weights, taking into account the complex multilevel sample design and nonresponse at each stage.

Descriptive analyses

Descriptive analyses will provide information on characteristics at a single point in time, overall and by key subgroups. For example, for questions on the characteristics of Head Start teachers (for example, teacher mental health)[4] and the characteristics of Head Start children and families (for example, family characteristics or children's skills at the beginning of the Head Start year), the study team will calculate averages (means) and percentages. The study team will calculate averages and percentages using t-tests and chi-square tests to assess the statistical significance of differences between subgroups. Additionally, the study team will examine open-ended questions for themes that emerge from the responses.

Multivariate analyses

The study team will use multiple approaches for questions relating resources and needs to the mental health of parents and teachers. These questions can be addressed by estimating regression analyses in which resources (such as Head Start services or social supports) and needs (such as assistance with housing) predict mental health outcomes (such as depression, anxiety, or stress). The study team will include responses to open-ended questions to provide qualitative context to these quantitative analyses.

Trend analyses

To examine changes across years—from fall 2019 to fall 2021—the study team will use t-tests and chi-square tests for simple comparisons (one year versus another). The study team will use trend analysis to examine whether children's development or family or teacher well-being change across time. For example, to compare parent mental health from fall 2019 (or earlier rounds) to fall 2021, the study team will employ a regression framework to examine the relationships between parent mental health and the year in which the well-being was measured, controlling for child and family characteristics.

Data Use

FACES 2019 and AIAN FACES 2019 will provide data on a set of key indicators for Head Start programs in their respective regions, which is of particular interest coming out of the COVID-19 pandemic. The study team will release a FACES 2019 data user's manual and an AIAN FACES 2019 data user's manual to inform and assist researchers who might be interested in using the data for future analyses. The manuals will include (1) background information about the study, including its logic model; (2) information about the FACES 2019 or AIAN FACES 2019 sample design, with the number of study participants, response rates, and weighting procedures; (3) an overview of the data collection procedures, data collection instruments, and measures; (4) data preparation and the structure of the FACES 2019 or AIAN FACES 2019 data files, including data entry, frequency review, data edits, and creation of data files; and (5) descriptions of scores and composite variables. Limitations will be clearly stated in materials resulting from this information collection.

Plans for reporting and dissemination

Mathematica will produce several publications based on analysis of data from each study:

- Key indicators will highlight descriptive findings on children, families, and teachers. The intention is to quickly produce findings federal agencies can use.
- Descriptive data tables with findings from all surveys and a description of the study design and methodology will be accessible to a broad audience.

---

[4] Because AIAN FACES 2019 supports analyses at the child level, it can answer questions on characteristics of Region XI Head Start children's teachers.

- Specific topical briefs of interest to the government, introduced in the data table reports but with greater depth will be focused and accessible to a broad audience, using graphics and figures to communicate findings.

- Restricted-use data files and documentation will be available for secondary analysis.

B8.   Contact Persons

The following individuals are leading the study team:

Nina Philipsen, Ph.D. Senior Social Science Research Analyst Office of Planning, Research, and Evaluation Nina.Hetzner@acf.hhs.gov	Meryl Barofsky, Ph.D. Senior Social Science Research Analyst Office of Planning, Research, and Evaluation Meryl.Barofsky@acf.hhs.gov
Alysia Blandon, Ph.D. Senior Social Science Research Analyst Office of Planning, Research, and Evaluation Alysia.Blandon@acf.hhs.gov	Laura Hoard, Ph.D. Senior Social Science Research Analyst Office of Planning, Research, and Evaluation Laura.Hoard@ACF.hhs.gov
Lizabeth Malone, Ph.D. Project Director Mathematica LMalone@Mathematica-Mpr.com	Louisa Tarullo, Ph.D. Co-Principal Investigator Mathematica LTarullo@Mathematica-Mpr.com
Nikki Aikens, Ph.D. Co-Principal Investigator Mathematica NAikens@Mathematica-Mpr.com	Sara Bernstein, Ph.D. Deputy Project Director Mathematica SBernstein@mathematica-mpr.com
Ashley Kopack Klein, M.A. Deputy Project Director Mathematica AKopackKlein@mathematica-mpr.com	Andrew Weiss Survey Director Mathematica AWeiss@mathematica-mpr.com
Sara Skidmore Survey Director Mathematica SSkidmore@mathematica-Mpr.com	Barbara Carlson, M.A. Senior Statistician Mathematica BCarlson@mathematica-mpr.com
Margaret Burchinal, Ph.D. Research Professor of Psychology Graham Child Development Institute burchinal@unc.edu	Marty Zaslow, Ph.D. Society for Research in Child Development mzaslow@srcd.org

To complement the study team's knowledge and experience, we also consulted with a working group of outside experts, as described in Section A.8 of Supporting Statement Part A.

Attachments

Attachment 23. Fall 2021 special telephone script and recruitment information collection for program directors, Regions I–X

Attachment 24. Fall 2021 special telephone script and recruitment information collection for program directors, Region XI

Attachment 25. Fall 2021 special telephone script and recruitment information collection for on-site coordinators, Regions I–X

Attachment 26. Fall 2021 special telephone script and recruitment information collection for on-site coordinators, Region XI

Attachment 27. FACES 2019 fall 2021 special teacher sampling form from Head Start staff

Attachment 28. FACES 2019 fall 2021 special child roster form from Head Start staff

Attachment 29. FACES 2019 fall 2021 special parent consent form for fall 2021 and spring 2022 data collection

Attachment 30. FACES 2019 fall 2021 special Head Start parent survey

Attachment 31. FACES 2019 fall 2021 special Head Start teacher child report

Attachment 32. FACES 2019 fall 2021 special Head Start teacher survey

Attachment 33. AIAN FACES 2019 fall 2021 special teacher sampling form from Head Start staff

Attachment 34. AIAN FACES 2019 fall 2021 special child roster form from Head Start staff

Attachment 35. AIAN FACES 2019 fall 2021 special parent consent form for fall 2021 and spring 2022 data collection

Attachment 36. AIAN FACES 2019 fall 2021 special Head Start parent survey

Attachment 37. AIAN FACES 2019 fall 2021 special Head Start teacher child report

Attachment 38. AIAN FACES 2019 fall 2021 special Head Start teacher survey

Appendices

Appendix P: Previously approved and completed data collection activities

Appendix Q: FACES and AIAN FACES 2019 fall 2021 instrument content matrices

Appendix R: FACES and AIAN FACES 2019 fall 2021 special program information packages

Appendix S: AIAN FACES 2019 special agreement of collaboration and participation

Appendix T: AIAN FACES 2019 special tribal presentation template

Appendix U: Logic models

Appendix V: FACES 2019 fall 2021 special respondent materials

Appendix W: AIAN FACES 2019 fall 2021 special respondent materials

Appendix X: FACES and AIAN FACES 2019 and 2020 nonresponse bias analyses