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

NOVEMBER 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

Contents

B1. Objectives....1
Study Objectives....1
Generalizability of Results....1
Appropriateness of Study Design and Methods for Planned Uses....1
B2. Methods and Design....2
Target Population....2
Sampling and Site Selection....3
B3. Design of Data Collection Instruments....10
Development of Data Collection Instruments....10
B4. Collection of Data and Quality Control....11
Recruitment Protocol....12
Monitoring Telephone Interviews....12
Monitoring the Web Instruments....12
Monitoring the Field Data Collection....12
Monitoring the Response Rates....13
B5. Response Rates and Potential Nonresponse Bias....13
Response Rates....13
Nonresponse....14
B6. Production of Estimates and Projections....15
B7. Data Handling and Analysis....15
Data Handling....15
Data Analysis....15
Data Use....16
B8. Contact Persons....17
Attachments....18
Appendices....18

Tables

B.1 FACES 2019 minimum detectable differences: Regions I–X (Spring 2022).....9

B.2 AIAN FACES 2019 minimum detectable differences: Region XI (Spring 2022).....9

B.3 FACES 2019 and AIAN FACES 2019: spring 2022 data collection activities.....11

B.4 Final response rates for fall 2019 and spring 2020 approved information requests.....14

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 planned additional data collection activities for FACES program- and classroom-level data collection in spring 2022, as well as a spring 2022 wave of FACES and AIAN FACES data collection to follow up on the fall 2021 work. This spring wave will address how families and staff are faring during the ongoing COVID-19 pandemic.[1] For information about previous FACES information collection requests, see https://www.reginfo.gov/public/do/PRAOMBHistory?ombControlNumber=0970-0151.

In this request, we detail the sampling plans for the spring 2022 data collection activities, including selecting classrooms, conducting data collection, analyzing data, and reporting study findings.

Generalizability of Results

FACES 2019 and AIAN FACES 2019 are designed 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 to answer 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 spring 2022 data collection reflects the latest round of the FACES Classroom Study (described further in the sample design section B2, below), which provides information on the drivers of program quality, with a larger sample for staff surveys and classroom observations. It will also include new items to describe how programs are using new quality improvement (QI) funds first awarded in 2020. The child-level data collection builds on the fall 2021 wave with its topical focus on the current well-being of children, families, and staff; and on change in response to the COVID-19 pandemic, as reflected in the proposed questionnaires. The AIAN FACES 2019 questionnaires also reflect the importance of Native culture and language—underscored in the logic model (Appendix U). 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 use of secondary data.

[1] Coronavirus disease 2019

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 spring 2022 data collection, is Regions 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 will have already freshened a random one-third of the FACES 2019 program sample ahead of fall 2021 data collection, and plans to freshen the remainder of the FACES 2019 program sample for spring 2022 by adding a few programs that came into being since the original sample was selected. For continuing programs not included in the fall 2021 data collection, the study team plans to reselect centers when one or both of the originally selected centers have closed—otherwise, the study team will keep the originally sampled centers in the sample. The study team will select new samples of classrooms for all centers in programs that were not part of the fall 2021 data collection. For AIAN FACES 2019, the study team does not plan to freshen the program sample for fall 2021, but for fall 2021 data collection, they will have already selected new center, classroom, and child samples for all programs. These samples will carry through to spring 2022.

Because the freshened samples for fall 2021 and spring 2022 are primarily composed of the fall 2019 and spring 2020 program samples, respectively, we provide here a bit of background on their design. 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;[2] 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. When members of the study team freshen the FACES program sample for spring 2022 data collection, they will get updated program lists directly from the Office of Head Start (OHS), as the PIR data collection for 2019–2020 were disrupted by the COVID-19 pandemic, and the PIR data for 2020–2021 will not yet have been released when the sample is being freshened.

[2] In fall 2021 and spring 2022, because programs may alter classroom configurations in response to the pandemic, we will select teachers instead of classrooms, or groups of children associated with teachers.

Sampling and Site Selection

Sample design

FACES 2019. The study team will use well-established methods to freshen the sample, ensuring the refreshed sample can be treated as a valid probability sample. The spring 2022 sample for FACES 2019 will primarily use the programs selected for spring 2020 data collection, but it will be freshened through selection of a small number of newer programs (those that had no chance of selection for fall 2019 or spring 2020). The study team will only freshen the centers in continuing programs if either of the originally sampled centers has closed; otherwise, we will keep the originally selected centers in the sample. For all newly sampled programs, the study team will select a new sample of centers.

Because typical Head Start instruction and other services were disrupted by 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. A variety of virtual and hybrid instructional scenarios have been instituted, and they may still be in place in spring 2022. To get around this issue the study team plans to draw a sample of teachers instead of classrooms for spring 2022 data collection using a process like the one carried out for fall 2021 data collection. For selected teachers who are teaching more than one center-based classroom in spring 2022, we will randomly select one for classroom observation; however, we will ask the teachers selected for the sample to complete the spring teacher survey about both classrooms they teach, and to complete teacher child reports for all children across the two classrooms who are in the sample.

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 spring 2022 data collection maps to the Classroom Study. The study team describes both the Classroom + Child Outcomes Study and the Classroom Study in this submission, as both make use of data to be collected in spring 2022.

Proposed sample sizes were determined with the goal of achieving accurate estimates of characteristics at the child and classroom levels, 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 expect 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 to allow 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. (Refer to section on "Degree of Accuracy" below.) The study team will work to achieve comparable sample sizes in spring 2022.

As was the case for the spring 2020 data collection in FACES 2019, the spring 2022 child-level sample will represent all children enrolled in Head Start in the fall of the program year who are still participating in Head Start in the spring. Likewise, the samples at the teacher, classroom, center, and program levels will represent those populations in spring 2022.

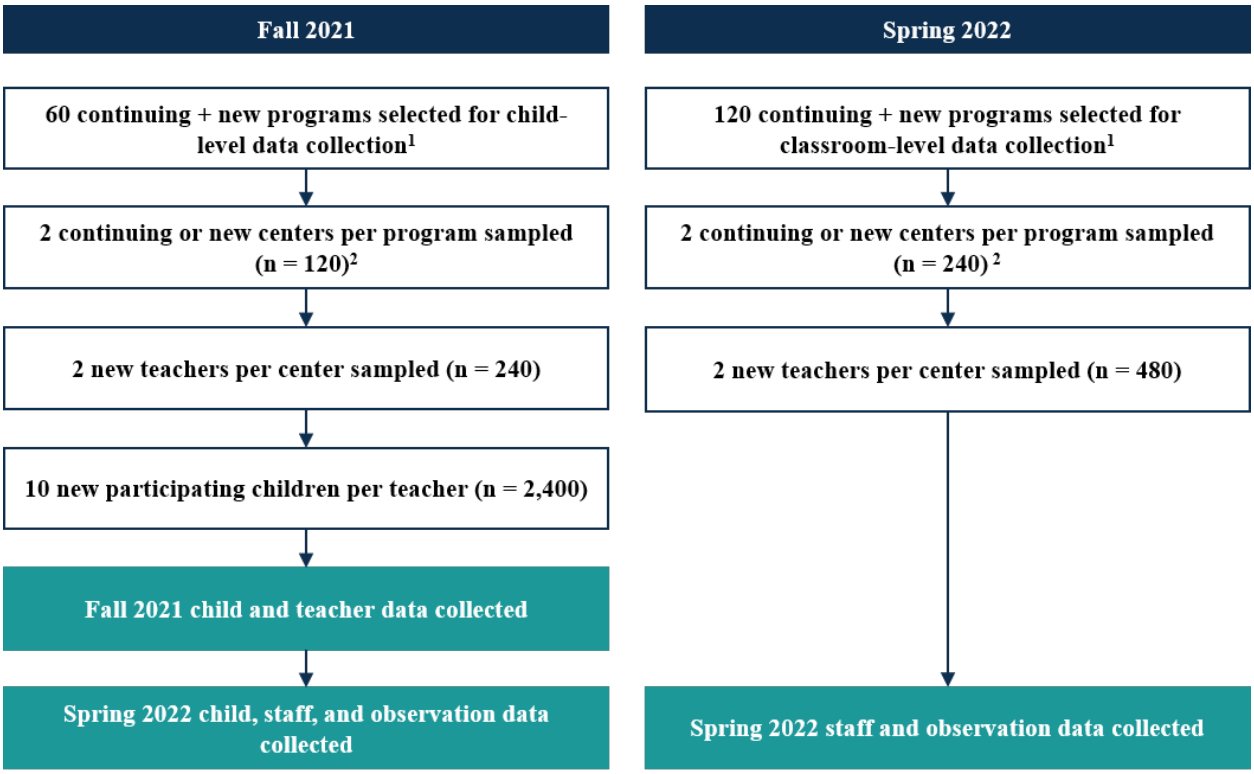
To minimize the effects of unequal weighting on the variance of estimates, the study team will sample with probability proportional to size (PPS) in the first two stages. At the third stage, the study team selects an equal probability sample of teachers within each sampled center and an equal probability 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 per program and per center, respectively. This sampling approach maximizes the precision of teacher-level estimates and allows for easier sampling of teachers and the children with those teachers. The study team targeted 60 programs for fall 2019 and spring 2020 data collection for Regions I–X at the child level. For spring 2020 data collection at the classroom level, we focused on 180 programs, including the 60 just described. Within these 180 programs, the study team selected, if possible, two centers per program, two classes per center, and (for the 60 fall programs) 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 spring 2022, the study team will select a refresher sample of programs and their centers to yield a new sample representative of all programs and centers in Regions I–X at the time of follow-up data collection, and the study team will also select a new sample of teachers in all centers and of children in a random subsample of centers. As part of the sample freshening for spring 2022, the study team plans to use in the spring 2022 data collection as many as possible of the original 165 FACES 2019 programs that participated in spring 2020, 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 2022, the study team will also use their originally sampled centers if those centers are still providing Head Start services. Otherwise, the study team will draw a new sample of centers for those programs. In all centers, the study team will select new samples of teachers (two per center). 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.[3] The study team uses a systematic sampling technique for selecting teachers.

For the continuing programs, and for any new programs in the spring 2022 sample that are added as part of the freshening process, the study team will go through the same OHS confirmation process used for fall 2019 (as well as for spring 2020 and fall 2021). 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 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 start participating 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. 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.

[3] 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, spring 2022



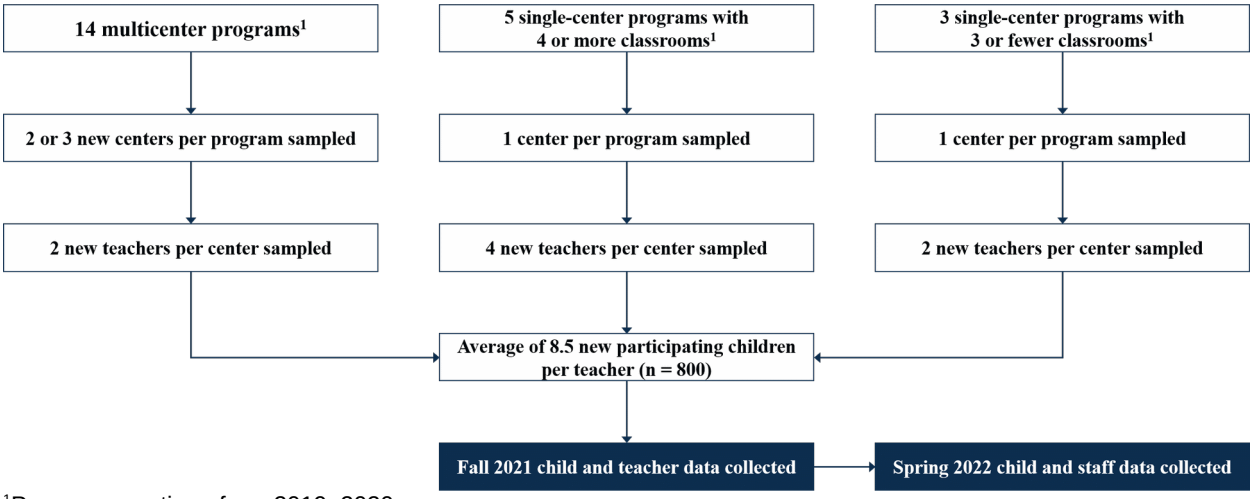
[1]Programs continue from 2019–2020 with freshening to reflect Head Start programs in 2021.

[2]For continuing programs whose sampled centers have not closed we will keep the same centers, otherwise we will select new centers.

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 within centers with equal probability.

AIAN FACES 2019. No new sampling will take place for AIAN FACES for spring 2022 data collection. We will collect data from the programs, centers, classrooms, and children included in the fall 2021 sample.

Figure B.2. Flow of sample selection procedures for AIAN FACES 2019, spring 2022



[1]Programs continue from 2019–2020.

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. Sample allocation was proportional to the estimated fraction of eligible classrooms represented by the programs in each stratum. The new programs eligible for the program sample freshening—those that did not have a chance of selection for the fall 2019 program sample—will form their own sampling 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 (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 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. For Region XI, the study team implicitly stratified (sorted) the sample frame by the state the program was in 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 spring 2022 data collection for both FACES and AIAN FACES, 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 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 spring 2022 data (or fall 2021–spring 2022 longitudinal estimates), using different combinations of completed instruments.

Thus, the program-level weight for spring 2022 will adjust for the probability of selection of the program, study participation, and (for FACES only)[4] response to the program director survey at the program level; the center-level weight will adjust for the probability of center selection, center-level participation, and (for FACES only) response to the center director survey; and the teacher-level weight will adjust for the probability of selection of the teacher and teacher-level study participation and (for FACES only) survey response. The FACES classroom observation weight will account for any within-teacher subsampling of classrooms and whether the observation was completed. For the FACES and AIAN FACES spring 2022 weights, for children still in the Head Start program, 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 (the 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 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.

[4] Analysis weights for AIAN FACES are produced at the child level only; however, these building-block weights at the program, center, and teacher levels account for sampling and study participation for AIAN FACES.

The analyses for the spring 2022 data collection for FACES 2019 will be at the program, center, teacher, classroom, and child 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 (refer to table footnotes).

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.

In Tables B.1 and B.2, 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 who will complete an instrument (for FACES, n = 2,040) in spring 2022. For example, the "33, 67" row within the child section represents a subgroup comparison involving child-level respondents in spring 2022 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 in Table B.1 assume that 85 percent of sampled classrooms will have completed teacher surveys, and all will have classroom observations; however, there will be fewer classroom observations than teachers, so the estimates in Table B.1 apply to both instruments. Both Tables B.1 and B.2 assume that 90 percent of the children in the sample who had parental consent in fall 2021 will still be in the sampled Head Start center in spring 2022. Of these, 85 percent 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 578 children still in those centers (also assumed to be 90 percent of the fall sample) with responses in the spring. We leave the burden estimate (Table A.10) 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 578 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 (Spring 2022)

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	306	306	.087	.145	.288
33	67	202	410	.092	.154	.307
15	85	92	520	.122	.203	.404
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	918	918	0.092/0.069	0.153/0.115	0.306/0.230
33	67	606	1,230	0.098/0.070	0.163/0.117	0.325/0.234
40	30	734	551	0.111/0.073	0.186/0.121	0.369/0.242

Note: Conservative assumption of no covariance for subgroup comparisons. Assumes =.05 (two-sided), .80 power. For teacher-level estimates, assumes 180 programs, 360 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 (Spring 2022)

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	289	289	0.159/0.120	0.265/0.200	0.526/0.397
33	67	191	387	0.170/0.122	0.283/0.203	0.559/0.404
40	30	231	174	0.193/0.127	0.322/0.211	0.635/0.420

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 612 teachers with completed teacher surveys in Regions I–X (FACES) in spring 2022, our design would allow us to detect a minimum difference of .288 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 1,836 children in spring 2022, and two approximately equal-sized child-defined subgroups (such as male and female), our design would allow us to detect a minimum difference of 11.5 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 578 children in Region XI (AIAN) with responses in spring 2022, our design would allow us to detect a minimum difference of 20.0 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 the contracted study team[5] (refer to section B8 for study team contacts) 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 spring 2022 includes surveys of parents, teachers, center directors, and program directors, as well as teacher reports of child development and classroom observations (FACES only). The parent and teacher surveys use the fall 2019–spring 2020 instruments as the foundation to continue to capture information on child, family, and teacher characteristics; family resources and needs; family and teacher well-being; teachers' experiences and perspectives on children's development; and classroom activities. The center and program director surveys use the spring 2020 instruments as the foundation to continue to capture information about director characteristics; staff recruitment, education, and training; center and program management and activities; and center and program systems and resources. As in fall 2021, the spring 2022 parent and teacher surveys will collect more information on family and teacher well-being in relation to the COVID-19 pandemic and parents' perspectives on children's development. The spring 2022 teacher and director surveys have also been updated to collect more information about staff well-being in relation to the COVID-19 pandemic, and about the types of supports centers and programs provide to support staff well-being and provide trauma-informed care. The spring 2022 director surveys have been updated to gather information about centers' and programs' recruitment activities during the COVID-19 pandemic, and—in FACES—programs' use of new quality improvement funds first awarded in 2020. Together data from these instruments can address research questions about who is participating in Head Start; about the associations between factors such as social supports, family well-being, and child development; and about approaches programs and centers are using to support family and staff well-being during the COVID-19 pandemic.

Wherever possible, the surveys use established scales with known validity and reliability. When there were not enough items on existing FACES or AIAN FACES data collection instruments that we could

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

use to measure the constructs of interest, we reviewed other surveys to identify relevant items or wrote new items (notably those studying experiences related to the COVID-19 pandemic and staff well-being) 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 pre-tested the staff surveys with fewer than ten respondents each to test for question clarity and flow in the content related to new items developed for the current wave.

B4. Collection of Data and Quality Control

Modes for all instruments are detailed in Table B.3.

Table B.3. FACES 2019 and AIAN FACES 2019: spring 2022 data collection activities

Component	Administration characteristics	
Parent survey	Mode	CATI/web
	Time	35 minutes
	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	40 minutes (FACES); 35 minutes (AIAN FACES)
	Token of appreciation	n.a.
Center director survey	Mode	Web SAQ/Paper upon request
	Time	35 minutes (FACES); 30 minutes (AIAN FACES)
	Token of appreciation	n.a.
Program director survey	Mode	Web SAQ/Paper upon request
	Time	40 minutes(FACES); 30 minutes (AIAN FACES)
	Token of appreciation	n.a.
Classroom observations (FACES only)	Mode	Paper with web data entry
	Time	4 hours*
	Token of appreciation	Classroom gift

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

n.a. = not applicable.

*Classroom observations do not impose burden on respondents.

The spring 2022 wave of FACES and AIAN FACES will deploy monitoring and quality control protocols that were developed for and used effectively during previous waves of the study.

Recruitment Protocol

The recruitment materials and plan for 120 spring-only FACES programs in spring 2022 were approved by OMB and recruitment is planned to begin on November 2021. To recruit new respondents, the study team will send correspondence to remind Head Start staff and parents about upcoming surveys (Appendix Y for FACES 2019 spring 2022 respondent materials; Appendix Z for AIAN FACES 2019 spring 2022 special respondent materials). The web administration of Head Start staff and parent surveys will allow the respondents to complete the surveys at their convenience. The study team will ensure 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 and directors 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 CATI staff will also be trained on cultural awareness and cultural humility for data collection.

Monitoring Telephone Interviews

For the parent telephone interview, professional Mathematica Survey Operations Center (SOC) monitors will monitor the telephone interviewers and observe all aspects of the interviewers' administration—from dialing through completion. Each interviewer will have their 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 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 Field Data Collection

The study team's approach to monitoring field data collection is multifaceted. First, FACES liaisons will hold regular calls with on-site coordinators—program or center staff who help with logistical details at the program—to discuss each site visit and any challenges that arise. Second, trained SOC staff will review all materials returned from the field for completeness and follow up with field and program staff, if needed. Third, gold standard observers will accompany each field staff conducting classroom observations at least once during the field period to conduct a quality assurance (QA) observation. The study team will compare the field staff's classroom observation scores with the gold standard observers' scores to determine the field staff's inter-rater reliability. The study team will provide feedback on errors, offer an opportunity for additional coaching with the study's gold standard observer who conducted the QA observation, give feedback from the observation trainers, and give the staff the opportunity to engage in a second observation in which the inter-rater reliability scores are checked again. Field staff who are unable to achieve the desired level of inter-rater reliability (80 percent of codes within one point) on the classroom observation will not be allowed to continue conducting observations. In these instances, the

study team will use the gold standard observer's scores for the observation and bring in another team member to conduct the field staff observer's remaining observations.

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 spring 2022 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 marginal unweighted response rates ranged from 75 to 93 percent, and AIAN FACES 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, 74 percent of eligible programs participated in spring 2020 data collection.

Obtaining the expected high response rates reduces the potential for nonresponse bias, making any estimates from the 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 spring 2022 to be better than those achieved in spring 2020, which were adversely impacted by the pandemic.

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%	59%
FACES spring classroom observation	100%	n.a.[a]
FACES spring teacher survey	85%	62%
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 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 teacher survey	85%	69%
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 seven 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[6] 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 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

[6] "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.

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.

B6. Production of Estimates and Projections

All analyses will be run using the final analysis weights, so 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 edit 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 spring 2022 and over time, (2) describe Head Start staff, (3) relate parent mental health to available resources, and (4) relate teacher mental health to classroom quality and child behavior (focusing on FACES only given that study's larger classroom sample size). Analyses will employ a variety of methods, including descriptive statistics (means, percentages), simple tests (*t*-tests, chi-square tests) of differences over time (and, for FACES 2019, across subgroups), 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), and the study team will use analysis weights that take 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, by program size [FACES], or for AIAN children only [AIAN FACES]). As an example, for questions on staff wellness and well-being (such as flexible work hours)[7] and the needs and supports for staff and families (such as resources for physical care), the study team will calculate averages (means) and percentages. The study team can also 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 emerging themes.

Multivariate analyses

The study team will use multiple approaches for questions relating classroom quality and child well-being to the mental health of teachers. These questions can be addressed by estimating regression analyses in which teacher mental health (such as depression or anxiety) predicts classroom quality (FACES only) and child behavior (such as classroom practices and child social skills). The study team will also use these methods for questions relating resources to parent mental health, where regression analyses predict parent mental health (such as depression or anxiety) based on resources available to the parent (such as Head Start services or social supports). The study team will include responses to open-ended questions to provide qualitative context to these quantitative analyses.

Data Use

FACES 2019 and AIAN FACES 2019 will continue to 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 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 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, staff, classrooms, and programs. 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.
- Specific topical briefs of interest to the government will be produced, with information introduced in the data table reports explored in greater depth. These 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.

[7] Because AIAN FACES 2019 supports analyses at the child level, this would be characteristics of Region XI Head Start children's teachers.

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. Co-Project Director Mathematica LMalone@Mathematica-Mpr.com	Andrew Weiss Co-Project Director Mathematica AWeiss@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
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 outside experts, as described in Section A.8 of Supporting Statement Part A.

Attachments

Attachment 11. FACES 2019 Head Start teacher survey

Attachment 12. FACES 2019 Head Start program director survey

Attachment 13. FACES 2019 Head Start center director survey

Attachment 20. AIAN FACES 2019 Head Start teacher survey

Attachment 21. AIAN FACES 2019 Head Start program director survey

Attachment 22. AIAN FACES 2019 Head Start center director survey

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

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

Attachment 30. FACES 2019 special Head Start parent survey

Attachment 31. FACES 2019 special Head Start teacher–child report

Attachment 36. AIAN FACES 2019 special Head Start parent survey

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

Attachment 39. FACES 2019 spring 2022 special teacher sampling form from Head Start staff

Appendices

Appendix I: FACES 2019 spring 2022 classroom observation components

Appendix P: Previously approved and completed data collection activities

Appendix Q: FACES and AIAN FACES 2019 spring 2022 instrument content matrices

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

Appendix U: Logic models

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

Appendix Y: FACES 2019 spring 2022 respondent materials

Appendix Z: AIAN FACES 2019 spring 2022 special respondent materials