2021–2022 Study of Family and Staff Well-Being in Head Start FACES Programs (2021–2022 Study)

Data User's Manual

September 2024

Maya Reid, Ashley Kopack Klein, Elizabeth Doran, Barbara Lepidus Carlson, Judy Cannon, Nikki Aikens, Sara Skidmore, Katie Gonzalez, Xinwei Li, Lizabeth Malone, Sara Bernstein

Submitted to:

Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services
330 C Street, SW, 4th Floor
Washington, DC 20201
Project Officers: Nina Philipsen and Laura Hoard
Contract Number: 47QRAA18D00BQ/
75D30121F12883

Submitted by:

Mathematica
1100 1st Street, NE, 12th Floor
Washington, DC 20002-4221
Telephone: (202) 484-9220
Facsimile: (202) 863-1763
Project Director: Sara Bernstein
Reference Number: 51307.01.T08.002.000

This page has been left blank for double-sided copying.

Acknowledgments

The authors would like to express their appreciation to Project Officers Alysia Blandon, Nina Philipsen and Laura Hoard; to Krystal Bichay-Awadalla, Jacquelyn Gross, Neda Senehi, and Kylee Probert; and to other federal staff at OPRE and the Office of Head Start. We thank the Mathematica team, including Andrew Weiss, Alex Hollister, Kathleen Feeney, Diletta Mittone, Morgan Woods, Rhiannon Jones, Tatiana Santiago, Timothy Luyster, John McCarthy, David Naglee, Sam Lem, Scott Peters, Joe McClurkin, Lauren O'Keefe, Ryan McInerney, Cole Garvey, Tiffany Waits, Macy Miller; Amelia Forman; Kevin Manbodh; Xinwei Li; Meghan Clough; Max Miller; Raquel Whitt; Kelsey Bagwell; Davis Straske; Natalie Reid; Tutrang Nguyen; Charlotte Cabili; Mary Kalb; Eric Grau; Cheri Vogel; Hanzhi Zhou; Cathy Lu; Myah Scott; Aden Bhagwat; Will Ratner; Jeffrey Harrington; Addison Larson; Effie Metropoulos; Molly Cameron; Jim Cameron; Felita Buckner; as well as Season Bedell; Richard Godwin; Karen Markoswki; Lucy Tindall; Sean Harrington; and Daniella Turner at the Survey Operations Center; and all of the Mathematica field and telephone staff who collected the data. We are also grateful for the contributions of our partners at Juárez and Associates. Most of all, we offer our gratitude to the staff, families, and children of the 176 Head Start programs across the country who once again opened their doors (virtually and/or physically) during this difficult period and shared their time with us.

This page has been left blank for double-sided copying.

Contents

Appendices

Exhibits

Getting started

The 2021–2022 Study of Family and Staff Well-Being in Head Start FACES Programs (2021–2022 study), builds on the Head Start Family and Child Experiences Survey (FACES), which has been a source of national information about Head Start programs and participants since 1997. Like FACES studies, the 2021–2022 study includes data from a large multistage sample of Head Start programs, centers, teachers, and families from across the U.S.; however, participation and response rates were lower than expected, resulting in data that may not be nationally representative of Head Start children and their families, teachers, centers, and programs. Therefore, data from most instruments collected as part of the 2021–2022 study should not be compared to FACES data (see Chapter VI for more information on cross-study analyses). This user's manual provides information to support researchers and data users who want to learn more about the 2021–2022 study and conduct secondary analysis of the descriptive data. The manual begins with guidelines for working with the 2021–2022 study data, an overview of the manual, and a list of acronyms for reference.

Guidelines for 2021–2022 study data users

We recommend that users follow these key steps when working with the 2021–2022 study data:

1. Before analyzing any data, review this user's manual for the following:
   - The **research questions** the study was designed to examine, and the limitations on use relative to previous FACES studies (see in particular chapters I and VI)
   - The **sampling approach** for the study
   - The **instruments used to collect data and the content** gathered in data collection
   - The **codebooks** for the data files
   - The available **composite variables**
2. Review the instruments to understand how questions were posed to study participants
3. Review survey skip patterns to understand what questions were and were not asked of participants based on their responses to earlier questions

In this manual, data users will find the following information:

- Chapter I presents an introduction of the 2021–2022 study—the motivation for the study, its design, the population of children and programs it represents, its analytic purposes, how it is different from FACES studies, and other unique features. It includes guidelines to follow and limitations to keep in mind when analyzing data from the 2021–2022 study. It describes the special considerations for using the data, including selecting weights, calculating variances, and conducting specific types of analyses. Data users should refer to this important introductory material when working with data from the 2021–2022 study.
- Chapter II discusses the study's sampling procedures and study design.
- Chapter III provides an overview of content included in the data collection instruments, including teachers' ratings of children and the Head Start parent, teacher, and director surveys.
- Chapter IV describes the sampling and recruitment of study participants, data collection, and quality control activities, as well as response rates for each study instrument.
- Chapter V discusses data preparation, including data entry, frequency review, data edits, and coding.

- Chapter VI covers the structure of the data files, content, and use, including information on analysis weights, variance estimation, and nonresponse bias analysis. Data users should refer to this chapter when working with data from the 2021–2022 study to understand potential limitations of the data for their specific analytic weights.
- Chapter VII presents scores and key composite variables (including reliability data as appropriate) for child, family, household, teacher, classroom, and program characteristics included in the data files.

In addition, the manual includes the following appendices:

- **Appendix A.** Elements of the FACES Design and Key Measures Used (and Child Outcomes Captured): FACES 1997–FACES 2019 and the 2021–2022 study
- **Appendix B.** Copyright Permissions
- **Appendix C.** Instrument Content Matrices
- **Appendix D.** Instruments
- **Appendix E.** Spring 2022 Center-/Program-Level File Codebook
- **Appendix F.** Fall 2021–Spring 2022 Classroom-/Teacher-Level File Codebook
- **Appendix G.** Fall 2021–Spring 2022 Child-Level File Codebook
- **Appendix H.** Descriptions of Composite Variables
- **Appendix I.** Nonresponse Bias Analyses

List of Abbreviations

AIAN	American Indian and Alaska Native
ACF	Administration for Children and Families
BPI	Behavior Problems Index
CATI	Computer-assisted telephone interviewing
CES-D	Center for Epidemiological Studies Depression Scale
CLASS	Classroom Assessment Scoring System
COR	HighScope Child Observation Record
DLL	Dual language learner
ECE	Early care and education
ECERS-R	Early Childhood Environment Rating Scale, Revised
ECLS-K	Early Childhood Longitudinal Study-Kindergarten Class of 1998–99
FACES	Head Start Family and Child Experiences Survey
FCC	Family child care
GAD–7	General Anxiety Scale–7
GED	General Education Diploma/General Education Development Test
GOLD	Teaching Strategies Gold
HPOG	Health Profession Opportunity Grants

ICPSR	Interuniversity Consortium for Political and Social Research
ICC	Intraclass correlation coefficient
IEP	Individualized education program
IFSP	Individual family service plan
MSHS	Migrant and Seasonal Head Start
NCPFCE	National Center on Parent, Family, and Community Engagement
NHES	National Household Education Survey
NEILS	National Early Intervention Longitudinal Study
NICHD	National Institute of Child Health and Human Development
NLSY	National Longitudinal Survey of Youth
PFCE	Parent, Family, and Community Engagement
PII	Personally identifying information
PIR	Program Information Report
PPS	Probability proportionate to size
Pre-K CLASS	Classroom Assessment Scoring System-Pre-K
PSU	Primary sampling unit
RAPID-EC	Rapid Assessment of Pandemic Impact on Development-Early Childhood
SIPP	Survey of Income and Program Participation
SMS	Sample management system
SOC	Survey Operations Center
SSI	Supplemental Security Income
SSRS	Social Skills Rating System
TANF	Temporary Assistance for Needy Families
TCR	Teacher Child Report
TCU SOF	Texas Christian University Survey of Organizational Functioning
T/TA	Training and technical assistance
USDA	U.S. Department of Agriculture
WIC	Special Supplemental Nutrition Program for Women, Infants, and Children

This page has been left blank for double-sided copying.

I. Introduction

Since 1997, the Head Start Family and Child Experiences Survey (FACES) has been a source of information on the population in Head Start; staff qualifications, credentials, and opinions; Head Start classroom practices and characteristics; and the outcomes of children and families. The most recent nationally representative study was conducted in the 2019–2020 program year. The motivation and goals of the Study of Family and Staff Well-Being in Head Start Family and Child Experiences Survey Programs (the 2021–2022 study) came from a need that arose as the COVID-19 pandemic continued into another year of affecting Head Start families' and staff's lives.

The 2021–2022 study design draws from FACES 2019 implementation and plans. In fall 2019, FACES 2019 completed a data collection wave aligned with previous FACES studies to include direct child assessments, parent surveys, and Teacher Child Reports (TCRs). FACES 2019 then planned to conduct a spring 2020 wave that would include the same child-level components, as well as classroom observations and surveys of teachers, center directors, and program directors. However, when COVID-19 was declared a pandemic by the World Health Organization and a national emergency by the U.S. Centers for Disease Control and Prevention, FACES 2019 canceled in-person direct child assessments and classroom observations for spring 2020. Spring 2020 data collection continued remotely for TCRs and surveys of parents, teachers, center directors, and program directors. The FACES 2019 design then called for a spring 2022 data collection that would include only classroom observations and staff surveys.[1]

Because of the continued impact of the COVID-19 pandemic on Head Start families and staff, FACES 2019 did not continue in its planned manner. Instead, the Office of Planning, Research, and Evaluation in the Administration for Children and Families (ACF), U.S. Department of Health and Human Services, funded Mathematica and its partner—Juárez and Associates—to conduct a new data collection effort, the 2021–2022 study, for the fall and spring of the 2021–2022 Head Start program year. The goal was to provide data on the characteristics and needs of families and staff after nearly two years of the COVID-19 pandemic. The 2021–2022 study includes data from a large multistage sample of Head Start programs, centers, teachers, and families from across the U.S.; however, participation and response rates were lower than expected and there is evidence that the weights did not fully mitigate the risk for nonparticipation and nonresponse bias. See sections A.1 and D.6 in this chapter and Chapter VI for more information on the limitations of these data.

A. Design of the 2021–2022 study

The design of the 2021–2022 study, including the sampling plan, instruments, procedures, and data analysis plans, drew heavily from the design of FACES 2019. However, the 2021–2022 study featured changes to more fully capture data on emerging policy issues and staff and family well-being in the context of the COVID-19 pandemic. The study also enhanced its focus on several areas, including families' experiences with Head Start, the intersection of Head Start with other early care and education (ECE) systems, families' economic conditions, and classroom instruction.

The sample design for the 2021–2022 study included two studies: a Program, Staff, and Family Study and a Program and Staff Study.[2] The two studies answer different sets of questions but also overlap to a

[1] For more information on FACES 2019, see the FACES 2019 User's Manual (Kopack Klein et al. 2021).
[2] The 2021–2022 study follows the redesigned approach used in FACES 2014 (Kopack Klein et al. 2018) and FACES 2019.

limited degree. All participating programs were involved in the Program and Staff Study (176 programs); a subsample of 60 of those programs were involved in additional data collection as part of the Program, Staff, and Family Study (Exhibit I.1). Both studies' samples drew from the sample of participating programs in FACES 2019 (see Chapter II for more information).

Planning for the 2021–2022 study during the evolving pandemic required modifying the sampling, recruitment, and data collection procedures relative to those used for FACES 2019. Most notably, we planned to eliminate all in-person activities, including in-person sampling, child assessments, and classroom observations.[3] In spring 2022, we added in-person parental consent distribution and collection in a subset of 15 programs participating in the Program, Staff, and Family Study. All other data were collected remotely in fall 2021 and spring 2022.

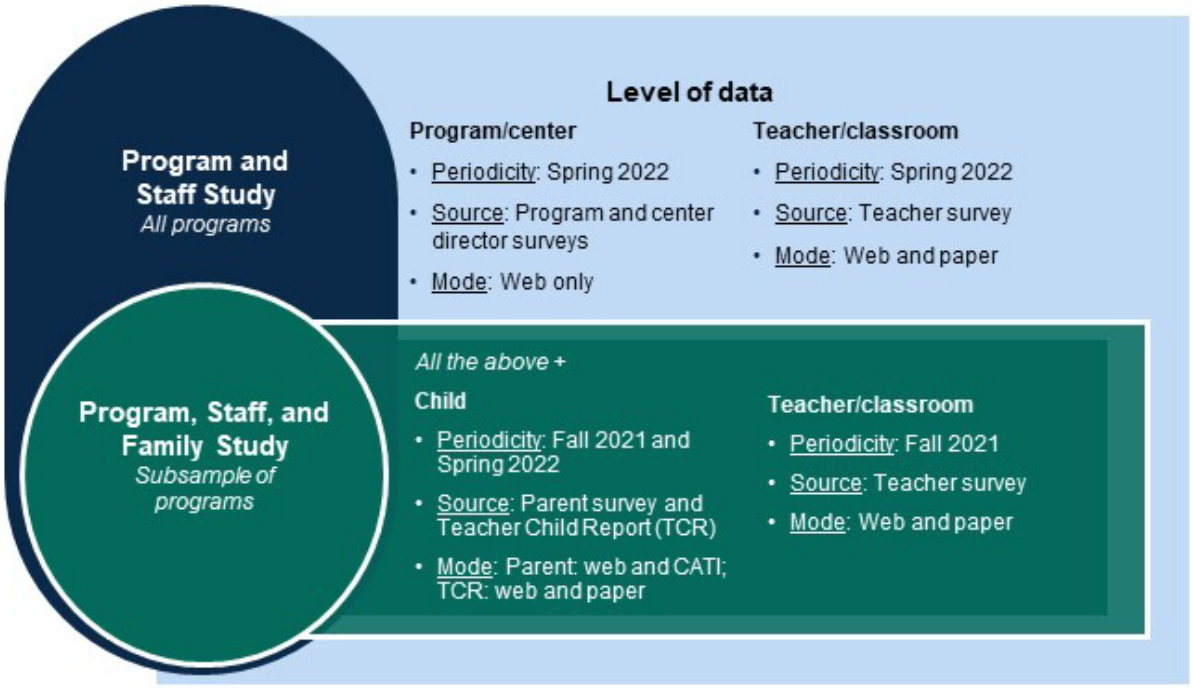
As outlined in Exhibit I.1, the Program, Staff, and Family Study provided information about the children and families who participated in Head Start and their Head Start teachers, whereas the Program and Staff Study provided information on Head Start programs, classrooms, and staff (directors and teachers). The Program, Staff, and Family Study was conducted in fall and spring of the 2021–2022 Head Start program year. In fall 2021 and spring 2022, parents and caregivers of children enrolled in the 60 programs participating in the Program, Staff, and Family Study completed surveys, and teachers completed a survey and rated the children's social and emotional skills, classroom behavior, and approaches to learning. In addition, teachers reported on any concerns about the children and how those concerns were addressed. The Program and Staff Study was conducted in spring 2022 and included teacher, center director, and program director surveys in all 176 participating programs (60 programs that were selected and participated in fall 2021 plus 116 additional programs).[4] A detailed description of the two studies' samples, including expected and actual sample sizes, appears in Chapter II.[5]

[3] We originally planned to conduct classroom observations in spring 2022; however, we decided to cancel observations out of safety concerns for Head Start program staff, children, and families, and our staff due to the rise of the Omicron variant in January 2022.

[4] The 60 programs selected for the Program, Staff, and Family Study were also part of the Program and Staff Study. These programs and the new programs selected for spring 2022 comprised the total sample for the Program and Staff Study.

[5] Sample sizes reported in this chapter are actual sample sizes.

Exhibit I.1. 2021–2022 study and its components



CATI = computer-assisted telephone interview.

1. 2021–2022 study sample and who it represents

The 2021–2022 study includes a sample of Head Start programs in Regions 1 through 10 and the children and families they serve. The sample does not include American Indian and Alaska Native (AIAN) programs (Region XI);[6] Migrant and Seasonal Head Start (MSHS) programs (Region XII);[7] programs in Puerto Rico and other U.S. territories; programs not directly providing services to 3-, 4-, and 5-year-olds; and programs under transitional management or soon to be defunded. Within the sampled programs, we selected centers. Within those centers, we selected teachers who had a least one child supported by Head Start. Then, for the 60 programs participating in the Program, Staff, and Family Study, rather than sampling children within a classroom, we selected all children receiving instruction from each selected teacher, regardless of child funding type.[8] Users should keep these exclusions in mind when reporting

[6] A separate study for Region XI AIAN programs—the 2021–2022 Study of Family and Staff Experiences in AIAN Head Start FACES Programs (2021–2022 AIAN study)—was conducted in parallel during the 2021–2022 program year. Although conducted at the same time, the 2021–2022 study and 2021–2022 AIAN study are not meant to be combined for analytical purposes. A separate 2021–2022 AIAN study data file and user's manual is available (Reid et al forthcoming). For more information on this study, see https://www.acf.hhs.gov/opre/project/2021-2022-study-family-and-staff-experiences-american-indian-and-alaska-native-head.

[7] In 2017, the Office of Planning, Research, and Evaluation funded the Migrant and Seasonal Head Start Study, which focused on Region XII. See https://www.acf.hhs.gov/opre/research/project/migrant-and-seasonal-head-start-study for details.

[8] Because typical Head Start instruction and other services were disrupted by the COVID-19 pandemic, we anticipated that it might be difficult to define classrooms the way we had done so in the past (a group of children

their findings and describing the populations of programs and children represented by the 2021–2022 study sample.

The 2021–2022 study aimed to describe the national population of Head Start programs, centers, teachers, classrooms, and children during the 2021–2022 program year. However, we were unable to fully meet this goal because of challenges related to the COVID-19 pandemic. Although we *selected* a nationally representative sample of Head Start programs, fewer of them participated than expected, despite our extension of the planned parental consent collection and data collection windows (discussed further in Chapter IV). We selected probability samples of centers, teachers, and children within the participating programs. Weights are available for analysis to account for the probability that children and their teachers, centers, and programs were selected for the study; lessen the risk of bias due to study nonparticipation and survey nonresponse; and provide results that represent, to the extent possible, all programs, centers, teachers, classrooms, and children in Head Start. The responding sample may not *fully* represent the population due to higher-than-expected nonresponse that may not have been adequately addressed with weighting adjustments (see section D6 below and Chapter VI for more information on nonresponse bias analysis).

Despite these limitations, the 2021–2022 study sample design supports many analyses for programs and teachers, as well as children. The Program and Staff Study's larger program sample (relative to that of the Program, Staff, and Family Study) allows for more powerful analyses of staff well-being. This design yields more precise classroom (and teacher) estimates and increases the ability to detect differences in staff well-being across different program contexts. At the child level, the Program, Staff, and Family Study's sample follows children through fall 2021 and spring 2022, allowing users to examine children's developmental change over the course of one program year. In Chapter II, we discuss in more detail the 2021–2022 study sample and its sampling design.

2. Analytic purposes of data from the 2021–2022 study

Data from the 2021–2022 study may be used for a variety of analytic purposes. For example, child-level data collected as part of the Program, Staff, and Family Study provide rich descriptive information on children and families who participated in Head Start in the 2021–2022 program year, including characteristics of the child, the household and its members (including the physical and mental health of the respondent), and aspects of the child's home life, as well as children's current learning skills, social-emotional development, approaches to learning, and developmental conditions and attendance. Data from this study may also be used to inform answers to questions about key subgroups of the Head Start population, such as children who are dual language learners (DLLs); children with identified disabilities; or those with several family risks (for example, children whose parents have less than a high school education or those who live in single-parent families).

Data from the Program and Staff Study in spring 2022 provide rich information for describing Head Start programs and the context in which Head Start children and families received services during the 2021–

taught together by the same teacher). Centers had instituted a variety of virtual and hybrid instructional scenarios early in the COVID-19 pandemic, and we anticipated that those might still be in place in 2021–2022. Thus, unlike in FACES 2019, where in the third stage of sampling we sampled at the classroom level, for the 2021–2022 study we sampled at the teacher level. As discussed in this manual, analyses can still be conducted at both the teacher and classroom levels. For more information on how sampling teachers instead of classrooms affected the sample of children invited to participate in the 2021–2022 study, see Chapter II: Sample and Sample Design.

2022 program year—including staffing and recruitment, staff education and training, curriculum and assessment, program management, use of program data and information, staff background and professional experiences, the impact of COVID-19 on programs and centers, teachers' instructional practices, teachers' beliefs, and staff mental health and well-being. See sections A.1 and D.6 in this chapter and chapter VI for information on the limitations of these data.

3. Key content areas of focus

a. The COVID-19 pandemic

The 2021–2022 study aimed to better understand the impact of the COVID-19 pandemic on programs and centers, as well as on Head Start teachers and the families served by Head Start programs. The parent surveys included questions regarding children's behavioral changes since March 2020, children's participation in virtual or remote learning activities, and families' experiences with the pandemic. Questions also asked about employment and household income changes due to the pandemic, as well as open-ended items aimed at understanding families' challenges and supports during the pandemic. The teacher surveys gathered information about personal and professional challenges teachers faced due to the pandemic and the types of supports they found useful. The program director survey gathered information about emergency preparedness and the largest lasting changes to the respondent's program as a result of the pandemic. Similarly, the center director survey gathered information about the largest lasting changes to the respondent's center as a result of the pandemic.

b. Staff and family well-being

The 2021–2022 study instruments demonstrate a substantial focus on Head Start staff and family well-being. The parent survey included questions regarding parents' physical and mental health, including parental stress and anxiety. Additionally, the 2021–2022 study gathered information on teachers' stress and anxiety, and supports for staff wellness and overall well-being in the teacher surveys. The program director survey included questions regarding supports for staff well-being and program directors' mental health, including depressive symptoms, anxiety, and job stress. Similarly, the center director survey asked questions on center directors' mental health, including depressive symptoms, anxiety, and job stress.

c. Social and community connections

The 2021–2022 study also focused on social and community connections for Head Start families, including questions on the parent surveys aimed at understanding whether specific community services would be helpful or have recently been helpful to families. We also added items to the parent surveys focused on families' knowledge of and comfort in accessing community supports and resources.

d. Emerging policy issues

In spring 2022, in collaboration with the Office of Head Start and the Office of Management and Budget, the 2021–2022 study enhanced its focus on emerging policy issues related to understand the changing landscape of early care and education, staff wellness, staff recruitment, and staff retention. We added items to the spring 2022 parent survey to understand parents' preferences for the format (virtual, in-person, or hybrid) for Head Start activities, why parents chose Head Start, and parents' plans for early care and education in the next school year and the reasons informing those plans. We added items to the spring 2022 teacher survey to gather information about staff recruitment and retention, staff stress and burnout, staff wellness, and teachers' overall wellbeing. We added items to the spring 2022 center

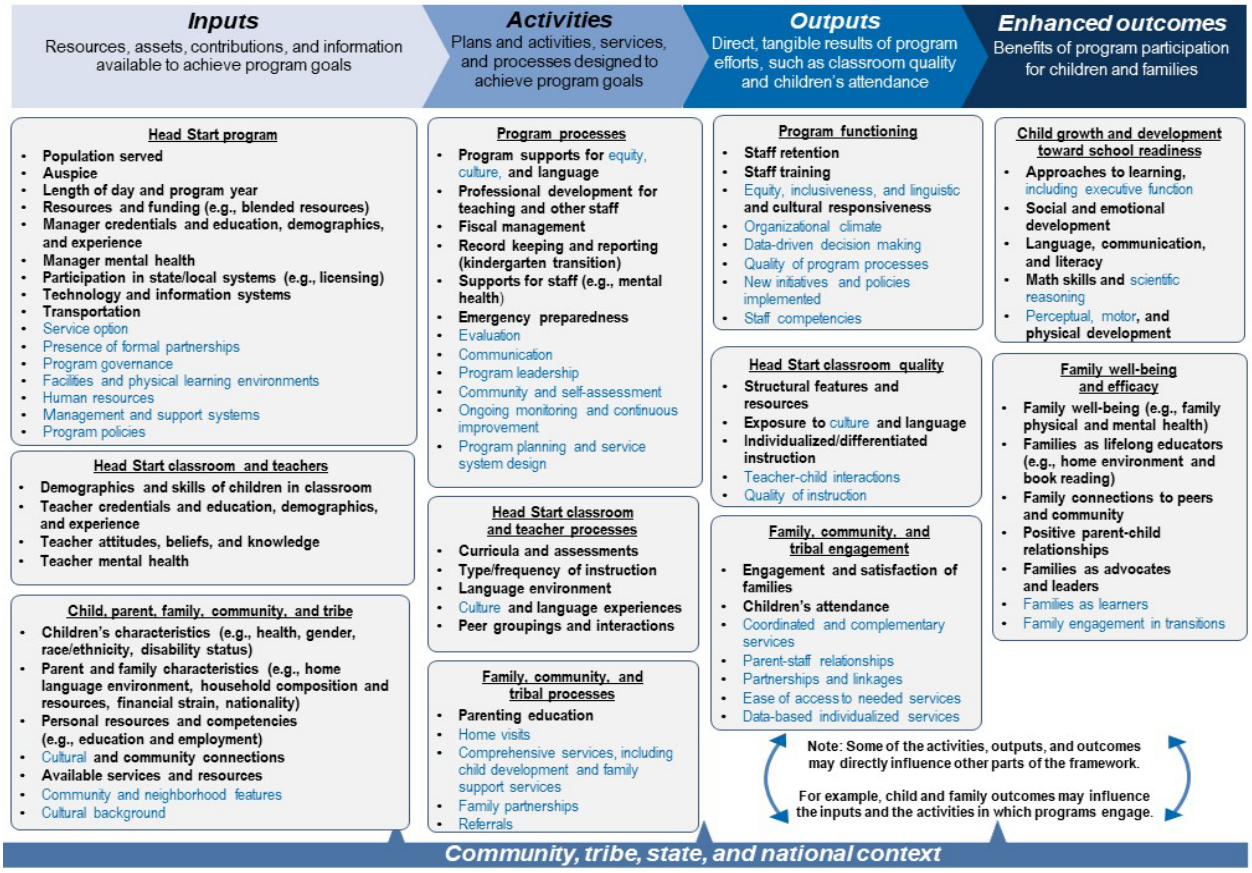
director survey to capture additional information about staff recruitment and retention and staffing challenges. Lastly, we added items to the spring 2022 program director survey to ask whether programs made an effort to recruit specific groups of families; about supports for staff wellness and overall wellbeing offered by the program; about the types of staff compensation offered by the program; about topics included in programs' emergency management/disaster preparedness and response plans; and about activities or expenses related to staff compensations and benefits programs implemented in the last 12 months.

B. Logic model

The key focus content of the 2021–2022 study fits within a broader logic model, depicted in Exhibit I.2, which documents key components of Head Start and how they support the program's intended outcomes. It encompasses Head Start as a whole and applies to all of the contexts and regions in which the program operates. The logic model shows the pathways from inputs to the ultimate goal of achieving enhanced outcomes for children and families. The underlying assumptions are as follows:

- Program inputs (for example, resources and funding, staff characteristics) affect the activities provided (for example, staff supports, curricula, and assessments). Those activities in turn produce key outputs (for example, quality of instruction and children's attendance), which ultimately lead to child and family outcomes.
- Some of the activities, outputs, and outcomes may directly influence other layers of the model. For example, child and family outcomes may influence the activities in which programs engage.
- All inputs, activities, outputs, and outcomes are influenced by a broader context. For example, federal, state, or local policies influence the inputs available to Head Start programs and families.

Exhibit I.2. Logic model for Head Start



Note: The logic model depicts Head Start more generally, beyond what the FACES studies can measure. The Black type indicates constructs measured in the 2021–2022 study; the blue type indicates constructs not measured.

C. Research questions

The 2021–2022 study is designed so users can answer a wide range of research questions crucial for aiding program directors and policymakers. The 2021–2022 study aims to describe (1) the children and families who participate in Head Start—families' resources and needs, parents' mental well-being, and children's cognitive and social-emotional skills; (2) the changes or trends in those characteristics over time, including since the onset of the COVID-19 pandemic; (3) the factors or characteristics that might explain differences in children's skills and development and family experiences; (4) the characteristics and mental well-being of Head Start teachers; (5) staff recruitment and retention, including staff compensation and benefits; and (6) the impact of the COVID-19 pandemic on program operations and emergency preparedness.

As Exhibit I.3 details, the 2021–2022 study is designed to address many questions related to these topics. This table presents research questions the study was designed to answer. Due to higher than expected nonresponse, there are limitations in how the data may be used to answer some of these research questions. For more information, see section D.6 below, Chapter VI, and Appendix I. These questions were formulated in collaboration with ACF and are designed to gather information on many topics of

interest to a variety of audiences, including the Office of Head Start; Head Start program leaders, administrators, and advocates; early childhood education practitioners and training and technical assistance (T/TA) providers; and interdisciplinary users. Users can access ACF products based on the 2021–2022 study and addressing questions about Head Start programs, classrooms, children, and families at https://www.acf.hhs.gov/opre/project/o-2021-2022-study-family-and-staff-well-being-head-start-family-and-child-experiences.

The data from the programs in the Program, Staff, and Family Study can address questions about the children and parents who participate in the program, including about children's development across one year in the Head Start program for both newly entering children and those returning for a second year. The study also supports research questions related to subgroups of interest, such as families with low income and specific racial/ethnic groups, as well as policy issues that emerge during the study. In addition, the research questions investigate the characteristics of Head Start programs, centers, and teachers, and the classrooms they teach. Users can use the same data to answer questions about the relationships between program and classroom characteristics and child and family well-being. The data from the larger sample of programs in the Program and Staff Study are most useful for answering questions about Head Start programs, classrooms, teachers, and program and center directors.

Exhibit I.3. Research questions the 2021–2022 study was designed to answer, by study

Research question	Program, Staff, and Family Study	Program and Staff Study
Describing Head Start programs and classrooms		
What are the characteristics of Head Start programs, including services offered, structural characteristics, and program policies and practices? Have these characteristics changed over time?		X
What are the characteristics of Head Start classrooms? Have classroom characteristics changed over time?		X
What are the characteristics and qualifications of Head Start teachers and management staff? Have these characteristics changed over time?		X
Did programs use the new quality improvement funds on activities or investments in the first year it was awarded? If so, what were those expenses?		X
What are the characteristics and mental well-being of Head Start teachers in spring 2022?		X
What are the types of compensation and benefits Head Start programs provide to staff? What positions receive each type of compensation and benefits? Have programs increased staff compensation and benefits in the past year?		X
To what extent have centers experienced challenges due to staff turnover and staff shortages in the past year?		X
How have Head Start programs prepared for future emergencies?		X
Describing family characteristics and children's outcomes		
What are the demographic characteristics of children and families served by Head Start in fall 2021? Have these characteristics changed over time?	X	

Research question	Program, Staff, and Family Study	Program and Staff Study
What are the demographic characteristics of children and families served by Head Start in spring 2022? Have these characteristics changed over time?	X	
What are the resources, supports, and needs of families served by Head Start in fall 2021 and spring 2022—notably related to economic situations and mental and physical health? Have these resources, supports, and needs changed over time?	X	
What are the cognitive and social-emotional skills of children in spring 2022?	X	
What are families' early care and education needs, and how do families make early care and education decisions?	X	
How frequently and in what types of activities are families involved in children's Head Start programs, and how satisfied are families with their experiences with Head Start?	X	
What is the relationship between program and social supports and families' economic situation and mental well-being?	X	

Note: This table presents research questions the study was designed to answer. Due to higher than expected nonresponse, there are limitations in how the data may be used to answer some of these research questions. For more information, see section D.6 below, Chapter VI, and Appendix I.

D. Guidelines for data users of the 2021–2022 study

The 2021–2022 study features several instruments (see Exhibit I.1 above and Chapter III), a range of data files and analysis weights (see Chapter VI), and composite variables (see Chapter VII) to support secondary analysis by users. Given the complexity of the 2021–2022 study design and data set, users need to follow several steps when working with the data. Below we outline the most important steps for learning more about the data in the remainder of this user's manual.

1. Review the study design and develop a plan based on users' specific research questions

As highlighted earlier in this chapter, the 2021–2022 study was designed to answer a variety of questions about Head Start programs, centers, classrooms, teachers, children, and families. It is imperative to match the sample to the research question (Exhibit I.4). This step will help users identify the most relevant instruments and data files needed.

Based on the sample identified in this matching step, the data may be in one of three data files available for the 2021–2022 study analyses: (1) a spring 2022 center-/program-level file, (2) a fall 2021–spring 2022 classroom-/teacher-level file, and (3) a fall 2021–spring 2022 child-level file (discussed further in Step 3). The sample identified will also guide decisions for weights (discussed further in Step 4) because the data files include a number of cross-sectional analysis weights to support analyses at all levels. For child- and teacher-level estimates, there are program year (fall 2021–spring 2022) weights for longitudinal analyses.

Exhibit I.4. Matching research questions to the 2021–2022 samples

If you are interested in…	Sample
Characteristics of participating Head Start programs and centers	Use data from the 176 programs and 340 centers found in the spring 2022 center- and program-level file
Characteristics of participating Head Start classrooms and teachers	Use data from the 236 teachers in the fall 2021 classroom- and teacher-level file, and the 631 teachers in the spring 2022 classroom- and teacher-level file
Head Start children and families who were receiving services through the entire program year	Use data from the 1,837 children and parents who participated in the study in fall 2021 and were still participating in spring 2022
Relationships between program- or classroom-level characteristics on child and family well-being during the 2021–2022 Head Start program year	Use data from the programs or teachers with linked fall 2021–spring 2022 child-level data—that is, the 60 programs and 236 teachers in the Program, Staff, and Family Study sample

Note: These data should not be assumed to be representative of the population of Head Start children, families, centers, and programs.

2. Identify content of interest

Users should become familiar with the study's instruments (Appendix D) to know exactly how the 2021–2022 study posed individual questions and groups of questions. Users should also note when specific item-level information is not available for review (for example, for copyrighted measures and instruments) and check for respondent paths and skip patterns in the instruments. For example, some questions may not be asked based on responses to earlier questions, or respondents may follow different pathways through the instruments in the spring based on their responses in the fall. In some cases, failure to consider valid skip patterns could result in incorrect reporting of findings. For example, when producing an estimate of the overall percentage of children whose parents usually speak to them in Spanish at home, users could overestimate the percentage if they do not recognize that respondents are asked about the specific language used when speaking with the child only if they have previously indicated the use of any language other than English in the home. Chapter VI provides additional information on special codes for item-level missing data due to such skip patterns.

Additionally, as part of identifying content of interest, users should consider the types of composite variables (or variables that combine information from multiple survey items) available in the data files. These variables include information on (1) child and family characteristics; (2) family processes and parenting; (3) Head Start teacher and classroom characteristics; and (4) center and program characteristics (see Chapter VII for more information).

3. Determine which files have the necessary data

Three data files are available for the 2021–2022 study: (1) a spring 2022 center-/program-level file, (2) a fall 2021–spring 2022 classroom-/teacher-level file, and (3) a fall 2021–spring 2022 child-level file. In Chapter VI, we provide details on the data file structure and content of each file. Each data file also has a corresponding codebook to confirm the variables available. Chapter VI also includes information on merging data from different files, including descriptions of identifiers in the data that may assist in the merging process. Most often, users use one file to conduct analysis. However, depending on the research question of interest, they may draw on data at different levels or in different data files from the 2021–

2022 study. For example, they may be interested in using data on children's Head Start teachers with data on child outcomes to explore the relationships between teacher characteristics and gains in school readiness skills. To do so, they must merge spring 2022 teacher source data (individual item-level data) with data from the child-level data file.

4. Determine the most appropriate weight to use

The 2021–2022 study data contain a variety of analysis weights. Although lower-than-expected participation and response rates may have contributed to the data in the 2021–2022 study not being fully representative of the populations of Head Start children and their families, teachers, classrooms, centers, and programs, using weights with the data is important to ensure that findings represent these populations to the extent possible. The 2021–2022 study uses a complex sample design (described fully in Chapter II). Not all programs, centers, classrooms, and children had an equal probability of selection. (The 2021–2022 study does not rely on a simple random sample.) Furthermore, not all sampled children, families, programs, and teachers participated in the 2021–2022 study. In addition, across the program year, some children left Head Start and were ineligible for spring 2022 follow-up by design. The 2021–2022 study analysis weights account for variations in the probabilities of selection, as well as for eligibility and differences in cooperation rates among those selected. They also account for sample attrition over time.

We recommend that data users employ weights to make responses from sampled respondents more representative of those from the target population. Use of analysis weights reduces the potential bias in findings associated with differential selection and nonresponse. With about 25 different analysis weights, users must decide on the best weight for answering their research question. In Chapter VI, we provide a full list of the weights and details on their construction. We also review the factors to consider in choosing the best weight for a particular analysis.

5. Conduct appropriate analyses

Users should approach analysis of 2021–2022 study data thoughtfully for the following reasons: the complexity of the study design, the lower than anticipated participation rates and resultant potential biases, the available indirect assessment scores and composite variables, and the weights that can be used when analyzing these data. Below we describe the most important steps users should follow when analyzing data from the 2021–2022 study.

a. Running analyses and checking results

To run analyses, users should start with an original, write-protected data file and then separate the data coding and preparation files (based on the review of instruments and composite variables) from the files to be used to run the analyses. They should always keep a record of the syntax across each of these steps. Users should first run analyses with unweighted data, and check data recodes through cross-tabulations and frequencies. Unweighted analyses will help users identify potential data coding errors in their work, as well as small cell sizes. Users should examine the data quality of analysis variables by running descriptive statistics and frequencies.

For all analyses, users should consider the potential of missing data and their type (item missing or instrument [unit] missing) (Chapter VI discusses the 2021–2022 study approach to missing data codes). The analysis weights provided with the data set adjust for unit missing data (unit nonresponse), but not for item-level missing data. With few exceptions (for example, household income), we have not imputed

missing data. Users should assess the plausibility of their findings. If results are surprising or seem to go in unexpected directions, the analyses should be examined for potential errors.

Users should also review the data file codebooks (Appendices E, F, and G) to check variable values and frequencies, noting when values obtained in their analysis do not align with the data. Before beginning any analyses, users should check for the availability of composite variables (Chapter VII and Appendix H). Use of such variables—which are created by combining (1) responses to several questions; (2) data across the two waves of the study; or (3) data from different sources (for example, parent survey data and data from the sample accrual or consent process)—can save users considerable time. The table sets that accompany the 2021–2022 study reports use these variables. A comparison of user results to the results published in the tables is a good test of a data set's correct application.[9]

b. Calculating variances

Standard errors are used to test hypotheses and explore group differences when making an inference to a population, as well as to indicate the precision of point estimates. They are a measure of the variance in the estimates associated with the selected sample and reflect that the selected sample is one of many possible samples. Most standard procedures in commonly used statistical software packages assume that a simple random sample is the source of the data. Therefore, software packages tend to underestimate the standard errors for complex sample designs such as the 2021–2022 study, which increases the chances of finding what appear to be statistically significant results when none are present.

Analysis of the 2021–2022 study data requires procedures appropriate for complex, multistage, clustered designs with unequal probabilities of selection. The 2021–2022 study observations are not independent (an assumption of a simple random sample)—that is, for some measures, children in the same classroom (or center) are more likely to share characteristics with one another than with children in other classes, centers, or programs. In using 2021–2022 study data, users should therefore use a design-based variance estimation method, such as the Taylor Series methods. In Chapter VI, we provide details on variance estimation using the Taylor Series methods. An alternative approach, which may be used during the preliminary or exploratory stage of data analysis, approximates the variance associated with complex sample designs by applying average design effects for similar outcomes and subgroups. In Chapter VI, we describe the alternative method in greater detail and list design effects for different 2021–2022 study weights.

c. Teacher-level versus classroom-level analyses

Data from the 2021–2022 study may be used to answer questions about Head Start classrooms and teachers. The classroom-/teacher-level file contains data collected from Head Start teachers about themselves and their classrooms. Given that some Head Start teachers taught more than one class, they provided separate information for each class but information about their background and demographics only once, so two records appear in the data file—one for each class. The teacher background and demographic variables are duplicated to appear in each record. Analyses of teacher-specific data (that is, estimates about Head Start teachers) thus must use the classroom-/teacher-level file and a teacher-level weight. The teacher-level weights are only provided for the primary classroom for the affected teachers

[9] 2021–2022 study reports are posted on the 2021–2022 study page on the ACF website at https://www.acf.hhs.gov/opre/project/o-2021-2022-study-family-and-staff-well-being-head-start-family-and-child-experiences.

(the primary classroom is defined as the classroom the teacher teaches earliest in the week; for teachers with an AM and a PM class, this would be their AM class). For analyses of classroom data at the class level, users must use the classroom-/teacher-level file and a classroom weight. In Chapter VI, we provide details on the structure and contents of the different data files and list the classroom and teacher weights (and all child weights) available for a given wave of data collection.

d. Cross-study analyses (2021–2022 study vs. FACES 2019)

Data from the fall 2021 teacher survey support cross-study analyses with data from FACES 2019 or earlier rounds of FACES. Due to lower than expected response rates, associated concerns about nonresponse bias (see section 6 below), and the resulting conclusions we can draw about who the 2021-2022 study sample represents, we do not recommend conducting cross-study analyses with data from any of the other 2021-2022 study instruments. See Chapter VI for additional information on potential limitations of cross-study analyses.

6. Consider small sample sizes and indicators of nonresponse bias

Users should be especially cautious when analyzing special populations (for example, children with specific types of disabilities), because the 2021–2022 study did not oversample special populations. Therefore, users should consider the sample size, standard errors, and reliability of the findings before drawing conclusions about special populations.

Additionally, program participation, parent consent, and survey response rates for the 2021–2022 study were lower than expected, increasing the risk of nonresponse bias. Nonresponse bias can occur when the survey responses of nonrespondents would have been different enough from those of respondents to make the results unrepresentative of those that would have been obtained had there been no nonresponse—that is, to bias them. To assess the risk of nonresponse bias, we compared respondents (weighted to adjust for nonresponse) to the full sample (respondents and nonrespondents) on variables available for both respondents and nonrespondents and plausibly correlated with key survey outcomes. One limitation of nonresponse bias analysis is that there are likely to be characteristics that are driving nonresponse but are not known for nonrespondents (for example, level of depressive symptoms).

In this section, we summarize results of the nonresponse bias analyses at a very high level; additional detail can be found in Chapter VI and Appendix I. In particular, Appendix I includes guidance for users about which variables to control for in their analyses, based on results of nonresponse bias analyses for specific weights.

a. Fall 2021 nonresponse bias analyses

We conducted five nonresponse bias analyses of data collected in fall 2021: (1) program participation; (2) teacher survey response; (3) child consent; (4) parent survey response; and (5) parent survey **and** teacher child report completion in combination. In the nonresponse bias analysis for program participation, we compared the 60 participating fall 2021 programs with the 104 eligible and selected programs. We found some differences between weighted estimates of covariates – that is, information available for both respondents and nonrespondents that are plausibly related to key survey outcomes – for the participating programs and the full sample that were large enough to conclude there is nonignorable risk of nonresponse bias for items on the surveys. As outlined in Chapter VI and Appendix I, program-level participation is a building block for other weights, so users should control for the covariates where we found these larger differences in their analyses.

When analyzing child consent and response to the teacher and parent surveys, most differences between the weighted respondents' covariate estimates and those of the full sample were small, indicating that these weights mitigated observable differences between respondents and nonrespondents for those covariates that were tested. When we examined response to the parent survey and teacher child report in combination, we found differences between weighted respondents' estimates of covariates and those of the full sample that were large enough to conclude there is nonignorable risk of nonresponse bias for items on the survey.

b. Spring 2022 nonresponse bias analyses

We conducted nine nonresponse bias analyses of data collected in spring 2022: (1) program director survey response; (2) center director survey response; (3) spring teacher survey response; (4) spring teacher survey, center director survey, and program director survey response in combination; (5) child consent; (6) fall **or** spring parent survey response; (7) fall **and** spring parent survey response; (8) fall **or** spring parent survey **and** spring TCR response in combination; and (9) fall **or** spring parent survey **and** fall TCR **and** spring TCR response in combination.

When considering child consent and response to the center director survey, the spring teacher survey, the fall **or** spring parent survey, and the fall **and** spring parent survey, most differences between weighted estimates of covariates from respondents and those from the full sample were small, indicating that these weights mitigated observable differences between respondents and nonrespondents for the covariates that were tested.

When considering response to the program director survey; combined spring teacher survey, center director survey, and program director survey response; combined fall **or** spring parent survey **and** spring TCR response; and combined fall **or** spring parent survey **and** fall TCR **and** spring TCR response, the results indicated there may be remaining indicators risk of nonresponse bias for items on the surveys after weighting based on the covariates that were tested. Several of the covariate estimate differences remaining between the full sample and the respondents after weighting adjustments are large enough for us to conclude that there is nonignorable risk of nonresponse bias for items on the surveys.

II. Sample and sampling design

The 2021–2022 study follows the redesigned approach used in FACES 2014 (Kopack Klein et al. 2018) and FACES 2019. The sample design for the 2021–2022 study was intended to be nationally representative of Head Start programs and the children and families they serve, and was applied in two studies: a Program, Staff, and Family Study and a Program and Staff Study. The two studies answer different sets of questions, but overlap to a limited degree (Chapter I). All participating programs were involved in the Program and Staff Study (with a goal of 180 programs); a subsample of 60 of them were involved in additional data collection as part of the Program, Staff, and Family Study. Each follows a separate design that considered cost and feasibility constraints to maximize efficiency but still meet desired precision targets. Both studies' samples drew from the sample of participating programs in FACES 2019. The following sections will define each stage of sampling separately for the two studies within the larger 2021–2022 study.

The 2021–2022 study aimed to describe the national population of Head Start programs, centers, teachers, classrooms, and children during the 2021–2022 program year. However, we were unable to fully meet this goal because of challenges related to the COVID-19 pandemic. Although we *selected* a nationally representative sample of Head Start programs, centers, teachers, and children, fewer of them participated than expected. Statistics are weighted to (1) account for the probability that children and their teachers, centers, and programs were selected for the study; (2) lessen the risk of bias due to study nonparticipation and survey nonresponse; and (3) provide results that represent, to the extent possible, all programs, centers, teachers, classrooms, and children in Head Start. However, participation and response rates were lower than expected and there is evidence that the weights did not fully mitigate the risk for nonparticipation and nonresponse bias. See Chapter VI for more information on the limitations of various types of estimates from these data.

For the 2021–2022 study, we started with the participating program sample from FACES 2019, removed any programs that subsequently became ineligible (lost funding or stopped providing services), and freshened it in summer 2021 to give a small number of newer programs—those that began providing services after initial program selection—a chance of selection. This process allowed the selected sample to be representative of all Head Start programs in the 2021–2022 program year. Of the 165 programs that participated in FACES 2019, 119 participated again in the 2021–2022 study and 46 did not (13 ineligible and 33 refusals). To freshen the sample, we released 17 newer programs, of which 11 participated in the 2021–2022 study. We also released 83 backup programs[10] from FACES 2019, of which 46 participated, to fill out the sample. The final participating program sample for the 2021–2022 study totaled 176 (60 randomly subsampled for the Program, Staff, and Family Study and the remaining 116 participating only in the Program and Staff Study). These 176 programs were organized into four cohorts (Cohorts 1, 2, 4, and 5) as defined by their point of study entry and level of data collection, as shown in Exhibit II.1.

[10] This release included 14 programs that were sampled in FACES 2019 and did not refuse participation; however, they did not participate at that time due to pandemic disruptions. Of these, eight participated in spring 2022.

Exhibit II.1. Programs participating in the 2021–2022 study, by study entry point (cohort)

	2021–2022 study participation		
FACES 2019 participation	Did not participate in the 2021–2022 study	Program, Staff, and Family Study (child- and staff-level data collection)	Program and Staff Study only (staff-level data collection only)
Did not participate in FACES 2019		20 (Cohort 4)	37 (Cohort 5)
Child- and staff-level data collection	19	40 (Cohort 1)	0
Staff-level data collection only	27	0	79 (Cohort 2)
Total 2021–2022 study programs = 176	**n.a.**	**60**	**116**

Note: All programs participated in the Program and Staff Study, which collected data from staff in spring 2022. Programs in the Program, Staff, and Family Study, comprising a subsample, collected data at the staff- and child-levels in both fall 2021 and spring 2022. Cohort 3 (not reported here) refers to a separate study, the 2021–2022 Study of Family and Staff Experiences in Head Start AIAN FACES Programs.

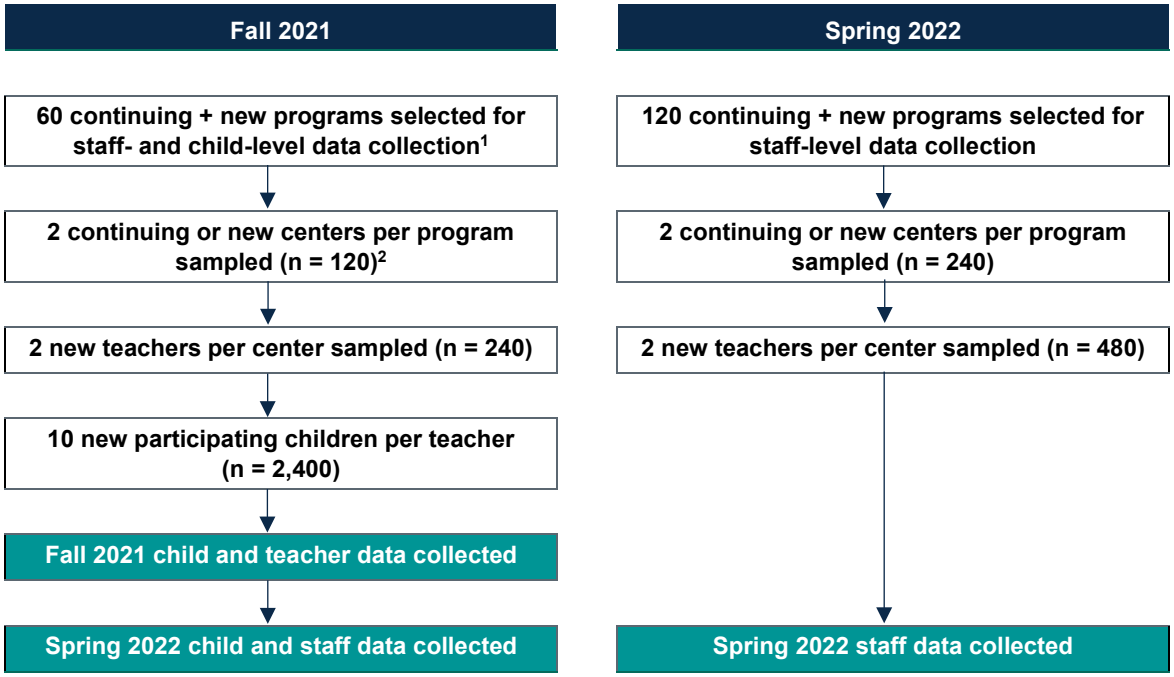
A. Multistage sampling approach

1. Program, Staff, and Family Study (Cohorts 1 and 4)

The 2021–2022 study used a multistage sample design, with the goal of having an efficient, sufficiently large representative national sample to be able to detect policy-relevant differences at the child level for the Program, Staff, and Family Study. The sampling target for the Program, Staff, and Family Study was 60 programs, and those 60 programs were a random subset of the 180 programs targeted in the larger Program and Staff Study. This section describes the sampling approach for these 60 programs only; we describe the sampling approach for the Program and Staff Study later in this chapter.

The four stages of the sampling design for the Program, Staff, and Family Study were as follows: (1) Head Start programs, with programs defined as grantees or delegate agencies providing direct services; (2) centers within programs; (3) teachers within centers; and (4) children taught by each teacher. When we sampled more than one child from the same family, we subsampled one child to minimize burden on participating families. For the first two sampling stages, we kept the sampled programs and centers from the FACES 2019 programs that participated in child- and staff-level data collection if they were still providing services in 2021–2022 and willing to continue participating in the study. Note that due to hybrid teaching situations during the pandemic, we opted to sample teachers rather than classrooms in the third sampling stage for the 2021–2022 study. (Sampling teachers still allowed us to produce classroom-level estimates because sampling classrooms in the past allowed us to produce teacher-level estimates.) Another pandemic-related change in sampling was in the fourth sampling stage: we considered the nonsampled children within each sampled teacher as backups to be released if necessary. Ultimately, we released into the sample all children within each sampled teacher.

Exhibit II.2. Flow of 2021–2022 study sample selection procedures and expected sample sizes



[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.

To minimize the effects of unequal weighting on the variance of estimates, the 2021–2022 study used a design that involved sampling with probability proportional to size[11] (PPS) in the first two stages (program and center). Following that stage, we sampled equal numbers of teachers within a center and equal numbers of children within a teacher, selected with equal probability at the final two stages; the goal was to give each teacher and child, respectively, an approximately equal chance of selection into the sample.[12] The Program, Staff, and Family Study was designed to select 60 programs, two centers per program, and two teachers per center (Exhibit II.2). Because it sometimes was necessary to pair a small teacher caseload with a larger teacher caseload before sampling to achieve the planned number of children in the 2021–2022 study, our initial sample in the Program, Staff, and Family Study included 236 teachers.[13] Within each selected teacher (or teacher pair), we initially sampled 12 children to result in 10 with parental consent per teacher, but ultimately sampled all children within sampled teachers, for a total of 1,363 consented children across the 60 programs in fall 2021.

[11] The measure of size was the number of classrooms in the program or center.

[12] The 2021–2022 Study is designed to maximize precision at the classroom/teacher level rather than the child level; however, in view of limited variation in classroom size, the precision of child-level estimates is not adversely affected to any great extent.

[13] When a teacher had fewer than 12 children, we paired them with another teacher to form a larger sampling unit—a teacher pair. If selected, both teachers were included in the study, and the child sample was proportionally allocated between the two teachers in the pair. We sampled 27 teacher pairs in fall 2021. We also had 21 sampled centers with only one classroom selected.

In Exhibit II.3, we compare the expected sample sizes (based on the study's design and sample selection procedures illustrated in Exhibit II.2 above) with the actual sample sizes we were able to achieve for both the Program, Staff, and Family Study (with data collection in 60 programs in both fall 2021 and spring 2022) and the Program and Staff Study (with data collection in all 176 programs in spring 2022 only). In fall 2021 and spring 2022, the COVID-19 pandemic was still affecting the U.S. Consequently, participation and response rates in the spring 2022 data collection were much lower than expected for the parent survey and Teacher Child Reports (TCRs), and classroom observations were not conducted.

Exhibit II.3. Expected and actual sample sizes

Total	Expected	Actual
Eligible and participating programs 2021–2022	180	176
Fall 2021 study entry[a]	60	60
Spring 2022 study entry	120	116
Centers selected and participating (up to two per program)	360	340
Fall 2021 study entry	120	113
Spring 2022 study entry	240	227
Teachers selected and participating (up to two per center)	720	631
Fall 2021 study entry	240	235
Spring 2022 study entry	480	396
Fall 2021 study entry only		
Children with parental consent in fall 2021	2,400	1,363
Children with a parent survey in fall 2021	2,040	785
Children with a Teacher Child Report in fall 2021	2,040	887
Children still in sampled Head Start program in spring 2022	2,160	2,732[b]
Eligible children with parental consent in spring 2022	2,160	1,837
Children with a parent survey in spring 2022	1,836	928
Children with a Teacher Child Report in spring 2022	1,836	1,250

a. "Fall 2021 study entry" includes those in the Program, Staff, and Family Study; "Spring 2022 study entry" includes the additional programs in the Program and Staff Study only. Together, they make up the Program and Staff Study sample of programs.

b. Fifty-one additional children moved to a nonsampled Head Start program and are considered part of the study's target population, though no data collection was attempted for them. Of these, only 19 children had parental consents in fall 2021.

To control the characteristics of the sample, the 2021–2022 study used explicit and implicit stratification at the first two stages of selection as appropriate. Explicit stratification involves forming strata (mutually exclusive groups within the sampling frame based on specified characteristics) from which separate samples are selected, whereas implicit stratification involves sorting the frame by specified characteristics within strata before sampling and then using a sequential sampling technique. By selecting the sample with implicit stratification within explicit strata, we ensured that the sample resembled the frame in its stratification variables. Stratification may also be used to oversample certain subgroups; however, we did not oversample any subgroups in the 2021–2022 study.

At each of the first two stages of sampling (program and center), the 2021–2022 study used a sequential sampling technique based on a procedure developed by Chromy (1979).[14] This procedure offers all of the advantages of the systematic sampling approach (selecting every nth case after a random start) but eliminates the risk of bias associated with that approach because of the random selection of the case within zones. The Chromy procedure allows for PPS sampling and explicit and implicit stratification, as described above. We carried out teacher and child selection remotely as simple random samples of teachers within a center and of children within the sampled teachers' rosters.

a. Sampling Head Start programs

The sampling frame for programs initially selected in 2019 was based on the final Head Start Program Information Report (PIR) database for program year 2017–2018 (the most current PIR available at the time of sampling). The sampling unit for the first stage was at the PIR reporting level—the grantee or delegate agency (what we call the program).[15] The PIR included about 3,400 Head Start and Early Head Start programs, and the sampling frame was a subset of that file. The sampling frame included all Head Start programs in the 50 states and the District of Columbia that met the study's eligibility criteria. We considered the following programs ineligible for the study and therefore excluded them before sampling:

- Programs in Puerto Rico and other U.S. territories
- Programs under American Indian and Alaska Native Head Start (Region XI) and Migrant and Seasonal Head Start (Region XII)
- Programs that did not provide direct services to children in the target age group (Early Head Start programs and grantees that provided only administrative services to their delegate agencies)
- Defunded programs under transitional management or entire programs temporarily out of operation

The project did not exclude from the FACES sampling process any programs that might have been involved in other ongoing Head Start studies. Before program sampling began, however, we coordinated with the Office of Head Start to obtain updated information about grantees, including those no longer receiving grants, new grantees not yet in the PIR, and those with changes in grant numbers, and then updated the PIR information accordingly to use as the sampling frame for Head Start programs. After various exclusions and the addition of new grantees, our sample frame comprised 1,515 Head Start programs.

In March 2019, as part of FACES 2019, we selected programs with PPS, using the number of classrooms in a program (from the 2017–2018 PIR) as the size measure. For particularly large programs (that had a measure of size guaranteeing selection), we selected them with certainty. We found six such certainty programs based on the FACES 2019 first sample release (for the target sample size of 180, without backups), though none was certain to be randomly subsampled for the 60 programs with child- and staff-level data collection. Explicit sampling strata included the following program characteristics: census region, urbanicity (metropolitan statistical area or not), and percentage of racial/ethnic minority

[14]The procedure makes independent selections within each of the sampling intervals while controlling the selection opportunities for units crossing interval boundaries.

[15] In 12 cases, we combined two or three programs (with different grant numbers) into a single program for sampling purposes if they had the same program name, program director, and program address, and overlapping centers. Two such program pairs ended up in the final sample of participating programs; we treated each as a single program for data collection and analysis.

enrollment.[16] After collapsing some strata that had just a few programs, we had 10 explicit sampling strata. Within the collapsed stratum variable, the implicit stratification was based on (1) whether the program had a population in which at least 25 percent of the children were dual language learners (DLLs), (2) its program status as a public school district grantee, (3) its ACF region (Regions 1 through 10), (4) the percentage of children in the program who were DLLs (categorized),[17] and (5) the percentage of children with disabilities. Explicit stratification allowed us to decide how to allocate the sample across strata to improve the sample's efficiency. Allocation of the sample proportional to the fraction of classrooms represented by the programs in each stratum can reduce variation in the selection probabilities of classrooms and children.[18]

We needed to allow for the possibility that a selected program might later turn out to be ineligible or refuse to participate so, as part of FACES 2019, we initially selected twice the number of programs needed within each stratum.[19] We then formed pairs of selected programs, as sorted by size (number of classrooms) within the same explicit stratum.[20] With equal probability,[21] we next selected one program within each pair for the initial sample release, leaving the other program in the pair available as a potential replacement for the released program. If we had to use the other program in the pair as a replacement, we treated both programs as released into the sample for purposes of calculating weights and response rates. We randomly selected 60 pairs (out of 210) to be included in the staff- and child-level data collection for FACES 2019 (plus 10 extra pairs). We used 16 such replacement programs, ending up with 59 programs participating in staff- and child-level data collection in FACES 2019.[22] Such a method offers the advantage of an uncomplicated way of replacing a nonparticipating program with a similar program while maintaining the ability to quantify the probability of selection and achieving the expected sample size.[23] Among the 59 programs that participated in FACES 2019, 40 participated again in the Program, Staff, and Family Study in fall 2021. To account for those that did not continue participation, we released 31 backup programs never released in fall 2019; of these, 15 participated in fall 2021.

[16] The PIR asked separately about Hispanic ethnicity and race for children enrolled in Head Start programs. To avoid double counting an unknown number of Hispanic children who are Black, we created a new minority stratification variable: (1) 40 percent or more of the children were Hispanic, or 40 percent or more of the children were Black; or (2) less than 40 percent of the children were Hispanic, and less than 40 percent of the children were Black. This stratification variable divided the programs in the frame roughly in half (56 versus 44 percent).

[17] This variable represents the percentage of children who were DLLs (based on child counts in the PIR). The *first* variable for implicit stratification is a dichotomous variable based on that percentage, indicating whether 25 percent or more of the children were DLLs. Later in the sort sequence, we sorted by a 10-category version of the same variable (based on deciles). Given that the percentage of children who are DLLs is an important control variable, we wanted to sort by that variable first, but at the same time wanted to limit the number of categories in this *first* sort variable to two to allow for meaningful sorting on the variables that followed, including the 10-category version of this characteristic.

[18] By reducing variation in the selection probabilities, we reduced the variation in analysis weights, which are the inverse of the selection probabilities, and improved the precision for survey estimates.

[19] We also selected some extra pairs to prepare for cases in which both members of a pair turned out to be nonparticipants. Two extra pairs were released for the sample of programs participating in staff- and child-level data collection.

[20] The Chromy procedure is associated with a serpentine ordering of the cases in the frame by the implicit stratification factors so that consecutive cases in the sorted frame are similar on several implicit stratification factors.

[21] The exception was the six certainty programs, which were always released first.

[22] These programs replaced 16 that refused and one that was ineligible.

[23] This method was also used in FACES 2014, FACES 2009, and FACES 2006.

We based the sample freshening for the 2021–2022 study on an updated list of programs provided by the Office of Head Start in summer 2021, which we compared to the list used for selecting the original sample. After the various exclusions described above, there were 152 new programs on the frame for the 2021–2022 study. We randomly selected 50 of them to be used as the subframe for the Program, Staff, and Family Study sample, with the remaining 102 to be used as the subframe for the Program and Staff Study only sample (described later). We sampled eight programs (two initial releases and six backups) for the Program, Staff, and Family Study sample, treating these new programs as a single explicit stratum when sampling, but we implicitly stratified the sample selections by urbanicity (metropolitan statistical area or not), census region, and number of classrooms. We ultimately released all eight programs into the sample, with five participating and three refusing, bringing us to 60 programs participating in fall 2021.

b. Drawing the sample of centers

Within each selected and participating program, we randomly selected two centers. In summer and early fall 2021, we asked all 60 of the recruited programs to participate in the Program, Staff, and Family Study to provide (1) a list of their centers and the number of classrooms they expected to have in the fall, and (2) other information needed for sampling (such as the number of classrooms and children per center and the percentage of DLL children in each center). Any centers that did not provide direct services to children, and those considered partnerships, were ineligible and we excluded from the sampling frame.[24] If a program had only one or two centers, we included all of them in the study sample.

Among the 40 FACES 2019 programs that agreed to participate in the 2021–2022 study, we kept their originally sampled centers if they were still providing Head Start services in fall 2021, but selected a new center sample of two per program if either originally sampled center had become ineligible in the previous two years. For any programs participating for the first time in the 2021–2022 study, we selected two centers per program. There were 69 centers in 36 programs that participated in fall 2019 and fall 2021, with the remaining 45 centers in 24 programs selected as part of new fall 2021 center samples.[25]

We selected very large centers (that had a measure of size guaranteeing selection) with certainty and other centers with PPS (based on expected numbers of classrooms). At this stage, we used no explicit stratification, although the frame was stratified implicitly (sorted) by the percentage of children who were DLLs. We did not build in a sample replacement strategy at this stage for the few centers that might not have participated or had become ineligible because in earlier FACES studies, participation among selected centers was fairly high. Our actual sample in the fall 2021 Program, Staff, and Family Study comprised 60 programs and 114 centers.

c. Drawing the sample of classrooms/teachers

In fall 2021, we asked each center to provide an electronic list of its lead teachers plus the names of any home visitors, the teaching mode (in person only, virtual/remote only, hybrid only, home visit, or other), and the number of enrolled Head Start children associated with each teacher or home visitor. To ensure a

[24] The Head Start PIR form defines a child care partner as "An individual child care center, umbrella organization operating multiple child care centers, child care resource and referral (CCR&R) network, family child care network, or other entity with which the Head Start program has formal contractual agreements to provide child care services to enrolled children that meet the Head Start Program Performance Standards."

[25] Note that some of these 45 *may* have participated in fall 2019, but because the other originally sampled center in that program became ineligible, we had to resample their centers. It is possible that such a center was randomly selected both times.

complete picture of the Head Start program for the Program, Staff, and Family Study, we included in the 2021–2022 study the Head Start services that a home visitor provided (in an individual home or a family day care setting). For sampling and analytic purposes, we treated each home visitor as a separate classroom and attached them to the center that the children visited for socialization and their families visited for other services. The 2021–2022 study did not over- or under-sample home visitors compared with center-based classrooms, and did not explicitly stratify by center- versus home-based classrooms. Given that their numbers were so few compared with the number of children in center-based classrooms (about 2 percent of total enrollment), the children served by home visitors were not the subject of separate analysis.

As described in Chapter IV, the study liaison entered the information for each teacher into their laptop, which was loaded with a sampling program. We paired teachers with fewer than 12 children with another teacher in the same center for sampling purposes to ensure a large enough sample (about 10 children).[26] We combined the smallest teacher caseload that needed pairing with the largest teacher caseload in the center, combined the next smallest with the next largest, and so forth, until we completed all needed pairing. For participating centers with only one or two teachers, we included all teachers. For all other participating centers, we selected a systematic sample of two teachers, implicitly stratifying by whether the teacher was teaching in person or not (using a program on the study liaison's laptop) to help ensure that all teaching modes would be represented in the sample.

Because all or nearly all teachers selected within participating centers were expected to participate, we did not establish a sample replacement strategy at this stage. The Program, Staff, and Family Study's targeted yield was 240 participating teachers in the sample. Our sample in the Program, Staff, and Family Study selected 236 eligible teachers (209 teacher sampling units, some of which were teacher pairs). All but one teacher remained as eligible participants between fall 2021 and spring 2022, so there were 235 of these fall 2021 teachers participating in the Program, Staff, and Family Study in spring 2022.

d. Drawing samples of children and parents

Immediately after sampling teachers, the study liaisons asked the selected and participating centers to provide the information needed for the last stage of sampling, which involved the selection of children and their parent or guardian (primary caregiver). We selected enough children in the initial sample to compensate for projected loss because of children not participating in the study for a variety of reasons (for example, child absent for an extended period, parent refusal, or child no longer in the program) and for loss because of sibling (within-family) subsampling. We obtained the requested teacher rosters from each selected teacher. The rosters included one record for each child in the teacher's caseload. In addition to the teacher/classroom indicator, each record specified the child's name, date of birth, enrollment date, Early Head Start participation, funding source, and instruction type (in person only, virtual/remote only, hybrid only, home visit, or other). In addition, for selected children only, the roster form collected information on sex, home language, parent/guardian name, and a sibling indicator to be used in case more than one child from the same family was selected.

For each selected teacher (or teacher pair), we selected 12 children, with each child having an equal probability of being selected. (For teacher pairs, we allocated the 12 selections between the two teachers

[26] The term "teacher" refers to both individual teachers and teacher pairs formed for sampling purposes. We also use teacher to refer to home visitors associated with centers. A teacher pair was a single sampling unit, but individual teachers in the pair were separate data collection and analytic units.

proportionally, according to the total number of children in each teacher's caseload.) After identifying siblings within the center's selected child sample, the laptop sampling program then randomly selected one child per family. (We define "family" and "sibling" as more than one child with the same primary caregiver; a parent or guardian.) The selection process reduced the number of children to 96.5 percent of the children in the first sample, or about 27.54 children per center out of 28.53 initially selected. We considered nonsampled children within sampled teachers as a backup sample. In December 2021, we released all of these children into the sample after seeing that the number of consented children was lower than expected. If we found any additional siblings among the backups, we retained the originally sampled child in the sample and "deselected" the backup sibling.

In fall 2021, the 2021–2022 study excluded children no longer in the sampled center at the end of the data collection period (as of January 31, 2022). In spring 2022, it excluded children no longer in the sampled center at the beginning of the data collection period (as of April 29, 2022). These children were still considered part of the target population at baseline (when the sample was selected). Sample selection occurred on a rolling basis between September 2021 and January 2022. The first parent surveys were completed in October 2021, and the first teacher surveys and TCRs in November 2021. Any children who joined the center after the sample was selected were not eligible to be in the sample; in other words, the baseline study population was defined at the time of sampling. Children selected for and later excluded from the sample (because of ineligibility, lack of parental consent, or other reasons) were included in the sample for purposes of constructing weights and calculating response rates.

On-site coordinators and other Head Start staff at each participating program worked to gain parental consent for selected children throughout fall 2021. We achieved parental consent rates of 44 percent in fall 2021 (1,363 out of 3,105 sampled and eligible[27] children). Due to delays in obtaining consent in fall 2021, we did not attain the targeted number of participating children per center; thus, unlike previous FACES studies, we continued collecting parental consent through the spring 2022 data collection field period. By spring 2022, only 2,783 sampled children were eligible (they were still receiving Head Start services), and 1,837 of them had parental consent (66 percent). Additional sampling at the child level did not take place for the spring 2022 data collection. Child-level data collection took place in spring 2022 for those children sampled in fall 2021 who were still receiving Head Start services in the sampled center and given parental consent to participate in the study.

At the child level, the 2021–2022 *selected* study sample represents all children enrolled in Head Start at baseline (fall 2021). The 2021–2022 study follows the children through fall and spring of one program year. The estimates from spring 2022 represent those children enrolled in Head Start in fall 2021 and still receiving Head Start services in spring 2022. This allows for estimates of children's gains in school readiness skills during their year in Head Start, based on actual measures obtained at the beginning and end of that year.

2. Program and Staff Study (Cohorts 1, 2, 4, and 5)

The Program and Staff Study was designed to target 180 participating Head Start programs, including the 60 programs targeted in the Program, Staff, and Family Study. We used a multistage sample design with three stages: (1) Head Start programs, (2) centers within programs, and (3) teachers within centers. We

[27] There were 3,135 children considered eligible for weighting and response rate purposes. The number 3,105 excludes those children who left the sampled Head Start center during the fall 2021 data collection.

did so to achieve high enough sample counts to detect policy-relevant differences at the classroom/teacher level (and even at the program level) by using an efficient, representative national sample.

To minimize the effects of unequal weighting on the variance of estimates, the design for the Program and Staff study involved sampling with PPS in the first two stages (program and center). Following that stage was sampling equal numbers of teachers with equal probability within centers at the final stage, with the goal of giving each teacher an approximately equal chance of selection into the sample. We selected and gained participation from 116 programs in addition to the 60 programs selected for the Program, Staff, and Family Study, for a total of 176 programs. Then, unless we were keeping the originally sampled centers for programs that participated in spring 2020, we selected two centers in each program. In all programs, we selected two teachers per center (Exhibit II.2). We added these newly selected centers and teachers to those already selected for the fall 2021 data collection. The sampling steps and procedures for these programs are the same as those described for the Program, Staff, and Family Study, with a few minor exceptions noted below.

a. Sampling Head Start programs

We selected all programs for the spring 2020 round of the FACES 2019 staff-level data collection (and their backups) at the same time, in March 2019. We randomly selected one-third of the sampled programs and their backups to participate in staff- and child-level data collection as part of FACES 2019 and selected the remaining two-thirds to participate in staff-level data collection (Exhibit II.2). Among the 106 programs that participated in the spring 2020 wave of FACES 2019 (in addition to the 59 that also participated in the fall 2019 wave of FACES 2019), 79 agreed to participate in the Program and Staff Study in spring 2022. For spring 2022, we approached 14 programs that had not participated in the spring 2020 round of the FACES 2019 study but were not refusals at that time to assess their willingness to participate in spring 2022 data collection. Of these, 1 had become ineligible and 5 refused, leaving 8 participating in spring 2022. In spring 2022, we released 38 backup programs from the spring 2020 sample, of which 4 had become ineligible and 11 refused, leaving 23 backup programs.

As described above in section A1a, we freshened the FACES 2019 program sample in 2021 for both the Program, Staff, and Family Study and the Program and Staff Study. We randomly sampled new programs that had no chance of selection in 2019, first randomly selecting 50 of these 152 programs to form the sampling frame for the Program, Staff, and Family Study; the remaining 102 formed the frame for the new programs for spring study entry in the Program and Staff Study. We sampled 24 new programs (six initial releases and 18 backups) and eventually released three backups. We treated these new programs as a single explicit stratum when sampling but implicitly stratified the sample selections by urbanicity (metropolitan statistical area or not), census region, and number of classrooms. To prepare for the spring 2022 data collection, we began recruiting the programs sampled for the Program and Staff Study only and released backup programs as necessary.

We ultimately recruited 116 additional programs for the Program and Staff Study (in addition to the 60 programs recruited for the Program, Staff, and Family Study): 79 continuing their participation from spring 2020, eight that did not participate in spring 2020 but had not refused, 23 backup programs from spring 2020 that were not released at that time, and six new programs.

b. Drawing the sample of centers

In late 2021 and early 2022 we selected two centers per program for those among the additional 116 programs (described above) that required new centers sampled, resulting in 224 additional centers

continuing from spring 2020 or newly selected, for a total of 345 centers sampled across the 176 programs (147 from spring 2020, 116 from fall 2021, and 82 from spring 2022), one of which was ineligible (closed).[28] In spring 2022, a total of 340 centers from all 176 programs were eligible for and participating in the Program and Staff Study.

c. Drawing the samples of classrooms and teachers

In spring 2022, we sampled teachers, as described above, for the Program, Staff, and Family Study. However, unlike the fall teacher selections, we excluded home visitors from the teacher sampling frames in these spring-only programs because we could not conduct an originally planned classroom observation for them. Further, for these programs, we had no need to pair small teacher caseloads with those of other teachers because no sampling at the child level took place within these sampled teachers. These additional teacher selections for the Program and Staff Study included 396 eligible teachers in spring 2022. When combined with the fall 2021 sample of 235 eligible teachers, there were 631 teachers in the Program and Staff Study.

B. Attrition and participation

We based all expected completion rates—comprising expected retention and cooperation rates—for the 2021–2022 study on the experience of FACES 2019. When designing the sample for the 2021–2022 study, we expected to have data on 180 programs, 360 centers, 720 classrooms, and about 2,160 children in spring 2022 (Exhibit II.4). The 2021–2022 study did not follow children who left the Head Start program between fall and spring. Mathematica found that children leaving Head Start accounted for a 10 percent sample loss between fall 2021 and spring 2022. Exhibit II.4 shows our assumptions for sample sizes in the 2021-2022 Study.

Exhibit II.4. The 2021–2022 study child sample size assumptions

	Fall 2021	Spring 2022
Children initially selected	2,880	n.a.
Proportion remaining after sibling subsampling	0.93	n.a.
Proportion initially eligible and consented	0.90	n.a.
Proportion of children not leaving Head Start since fall	n.a.	0.90
Eligible and consented children and parents	2,400	2,160
Response rate to the parent survey	0.85	0.85
Response rate to the Teacher Child Report (TCR)	0.85	0.85
Completed parent surveys	2,040	1,836
Completed TCRs	2,040	1,836

n.a. = not applicable

After accounting for initial consent and attrition among those who left Head Start between fall and spring and were no longer part of the study population, we expected participation rates for both children and their caregivers to decrease slightly between the two data collection periods. As seen in Exhibit II.4, of

[28] Note that there were three centers that participated in the study in spring 2022 but for which it was too late in the program year to sample their teachers, so we considered those centers to be nonparticipants when constructing the teacher sampling weights.

the initial sample of 2,400 children with parental consent, we expected 2,160 children and their parents and caregivers (90 percent) to be retained through spring of 2022.

Although this exhibit shows the sample sizes for the two instruments at the child level (parent surveys and teacher child ratings), we do not include observations at the classroom level (teacher surveys), the center level (center director survey), or the program level (program director survey). We expected nearly 100 percent cooperation among center and program directors (resulting in about 360 completed center director surveys and 180 completed program director surveys). We also expected that a high percentage of teachers in the classrooms would have completed the teacher survey. As mentioned above and discussed in detail in Chapter IV, the pandemic made in-person data collection classroom observations infeasible in spring 2022 and dampened response rates in the instruments we collected remotely.

We initially sampled a total of 3,252 children. As seen in Exhibit II.5, among those children, 3,139 remained after sibling subsampling, 1,363 (44 percent of 3,105 eligible) were eligible and received parental consent in fall 2021, and 785 (58 percent of consented) of their primary caregivers participated in the initial data collection in fall 2021. Teachers provided reports on 887 of the consented children (65 percent) in the fall. By spring 2022, only 2,732 of the 3,139 sampled children were still eligible for the study, but the number of consents among the eligible children increased to 1,837.

Exhibit II.5. The 2021–2022 study child actual sample size

ACTUAL	Fall 2021	Spring 2022
Children initially selected	3,252	n.a.
Proportion remaining after sibling subsampling	0.97	n.a.
Proportion initially eligible and consented	0.44	n.a.
Proportion of children not leaving Head Start since fall	n.a.	0.87
Eligible and consented children and parents	1,363	1,837

n.a. = not applicable

C. Power

Given the sample design noted above and its impact on the variance of estimates, including assumed rates of consent, response, and attrition, as well as design effects, the expected sample size would have been large enough to detect meaningful differences for various types of analyses at the classroom and child levels. Our actual sample sizes were smaller than expected due to challenges experienced during recruitment and data collection (Ch IV). Exhibits II.6 and II.7 show the target and actual sample sizes available for each level of analysis. The numbers in Exhibit II.8 are based on expected sample sizes and Exhibit II.9 is based on the actual numbers of teachers with completed surveys in the spring (n = 358) and consented children with a parent survey in either the fall or spring (n = 1,158). Both assume 80 percent power, Type I error rate of 0.05, various sample and subgroup sizes, and different assumptions about the impact of clustering on variance.

1. Classroom level

As seen in Exhibit II.6, when making classroom-level estimates based on 358 completed teacher surveys in the spring 2022 Program and Staff Study, the study is able to detect (1) differences of 22 percentage points between two equal-sized subgroups of classrooms for a percentage-based outcome near 50 percent

and (2) differences of 0.438 standard deviations for a measure such as the teacher job satisfaction scale and some of the teacher beliefs about teaching subscales (for example, the Didactic subscale), which have standard deviations of about 0.7. Drawing on findings from earlier FACES studies, we assumed an intraclass correlation coefficient (ICC, or clustering effect) of 0.20 between programs and an ICC of 0.20 between centers for classroom-level estimates.

2. Child level

For estimates at the child level (Exhibit II.6), from the parent survey in fall 2021 or spring 2022, we determined that the number of children was large enough to be able to detect differences of about 18.1 percentage points in an outcome of around 50 percent between two child-level subgroups of equal size. We assume an ICC of 0.10 between programs, 0.05 between centers, and 0.12 between classrooms for child-level estimates (based on findings from earlier FACES studies). For example, if we examined the percentage of children rated in excellent health (in either the fall or spring) with a sample size of 1,158 children, and compared two approximately equal-sized subgroups defined by child-level characteristics (such as male and female), the sample sizes in the 2021–2022 study would permit the detection of a minimum difference of 18.1 percentage points with 80 percent power.

Exhibit II.6. The 2021–2022 study fall 2021 target versus actual sample sizes

		Target	Actual
Child-level data	**Number of eligible and consented children**	**2,400**	**1,363**
	Completed parent surveys	2,040	785
	Completed TCRs	2,040	887
Teacher-level data	**Number of selected teachers**	**240**	**239**
	Completed teacher surveys	204	193

Exhibit II.7. The 2021–2022 study spring 2022 target versus actual sample sizes and instrument completion and response rates

		Target	Actual
Child-level data	**Number of eligible and consented children**	**2,160**	**1,837**
	Completed parent surveys	1,836	928
	Completed TCRs	1,836	1,250
Teacher-level data	**Number of selected and eligible teachers**	**720**	**631**
	Completed teacher surveys	612	358
Program-level data	**Number of recruited programs**	**180**	**176**
	Completed program director surveys	153	132
Center-level data	**Number of selected and eligible centers**	**360**	**344**
	Completed center director surveys	306	237

Exhibit II.8. Minimum detectable differences in the 2021-2022 study as designed

Point-in-time subgroup comparisons for classrooms and children							
	Classroom subgroups				Minimum detectable difference		
Time point	Percentage in Group 1	Percentage in Group 2	Classes in Group 1	Classes in Group 2	Percentage of 10 or 90	Percentage of 50	Minimum detectable effect size
Spring 2022	50	50	360	360	8.4	14.0	28.0
	33	67	238	482	9.0	14.9	29.8
	15	85	108	612	11.9	19.8	39.2
	Child subgroups				Minimum detectable difference (Program-defined subgroups / Child-defined subgroups)		
Time point	Percentage in Group 1	Percentage in Group 2	Children in Group 1	Children in Group 2	Percentage of 10 or 90	Percentage of 50	Minimum detectable effect size
Fall 2021	50	50	1,200	1,200	9.1/6.7	15.1/11.2	30.1/22.4
	33	67	792	1,608	9.6/6.8	16.1/11.3	32.0/22.7
	40	30	960	720	11.0/7.0	18.3/11.7	36.4/23.3
Spring 2022	50	50	1,080	1,080	9.1/6.8	15.2/11.3	30.3/22.6

Note: The exhibit reflects a conservative assumption of no covariance for point-in-time subgroup comparisons Assumes α = .05 (two-sided), .80 power. For classroom-level estimates, assumes 180 programs, 360 centers, between-program ICC=.2, between-center ICC = .2. For child-level estimates, assumes 60 programs, 120 centers, between-program ICC = .10, between-center ICC = .05, between-classroom ICC = .12.

The minimum detectable effect size is the minimum detectable difference in SD-sized units.

Exhibit II.9. Actual minimum detectable differences in the 2021-2022 study

Point-in-time subgroup comparisons for teachers and children							
Time point and instrument	Teacher subgroups				Minimum detectable difference		
	Percentage in Group 1	Percentage in Group 2	Classes in Group 1	Classes in Group 2	Percentage of 10 or 90	Percentage of 50	Minimum detectable effect size
Spring 2022 teacher survey	50	50	179	179	13.2	22.0	0.438
	33	67	118	240	14.1	23.5	0.465
	15	85	90	269	15.3	25.5	0.505

	Child Subgroups				Minimum detectable difference (Program-defined subgroups / Child-defined subgroups)		
Time point	Percentage in Group 1	Percentage in Group 2	Children in Group 1	Children in Group 2	Percentage of 10 or 90	Percentage of 50	Minimum detectable effect size
Fall 2021 or Spring 2022 Parent Survey	50	50	579	579	14.1/10.8	23.5/18.1	0.466/0.360
	33	67	382	776	15.0/11.1	25.0/18.5	0.495/0.368
	40	30	464	348	17.1/11.6	28.4/19.4	0.563/0.386

Note:Conservative assumption of no covariance for point-in-time subgroup comparisons. Assumes α =.05 (two-sided), .80 power.

Teacher numbers assume 176 programs, 340 centers; between-program ICC=.2, between-center ICC = .2. Child numbers assume 60 programs, 114 centers, 236 teachers; between-program ICC = .10, between-center ICC = .05, between-teacher ICC = .12.

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

III. Data collection instruments

In this chapter, we explain the content of instruments used in the 2021–2022 study to describe Head Start programs, classrooms, children, and families, as well as historical reliability for these measures from FACES studies where applicable. For information on the construction of variables and scales, as well as reliability (Chronbach's alphas) for the 2021-2022 study data collection, please see Chapter VII. For information on whether each item was included in FACES 2019 or FACES 2014, please see Appendix C. Most items in each instrument were included in previous FACES studies; in each section of this chapter, we have highlighted items included for in the 2021-2022 Study that had not been previously used in FACES studies. If a scale was created, it is mentioned here and described in more detail in Chapter VII.

- **The parent survey** collected information in several areas, including the characteristics of the child, the household and its members (including the physical and mental health of the respondent), and aspects of the child's home life.
- The **Teacher Child Report (TCR)** collected teachers' ratings of the children's current learning skills, social-emotional development, and approaches to learning, as well as information on developmental conditions and attendance.
- **The teacher survey** asked about the teacher's training and educational background, professional experience, instructional practices, feelings about teaching and the Head Start program, mental health and well-being, and the impact of the COVID-19 pandemic. It also asked about learning activities and curriculum, planning and assessment, the classroom environment, and professional development.
- **The program director survey and center director survey** addressed staffing and recruitment, staff education and training, curriculum and assessment, program management, use of program data and information, directors' backgrounds, the impact of COVID-19 on programs and centers, and directors' mental health and well-being.

In Exhibit III.1, we present information on the data collection instruments by wave.

Exhibit III.1. 2021–2022 study data collection instruments, by wave

	Fall 2021	Spring 2022
Parent survey	X	X
Teacher Child Report (TCR)	X	X
Teacher survey	X	X
Program director survey		X
Center director survey		X

A. Approach to identifying and developing instrument items

Instrument items selected for the 2021–2022 study sought to provide data on the characteristics and needs of families and staff after nearly two years of the COVID-19 pandemic. Many of items used in the 2021–2022 study were also included in previous FACES studies. Appendix C details the contents of the data collection instruments and notes whether the same topics were covered in FACES 2014 or FACES 2019 (the FACES studies conducted prior to the 2021–2022 study). Many new topics involved identifying existing scales, drawing on survey questions from other large early childhood studies (such as the U.S.

Department of Education's Early Childhood Longitudinal Studies or the Rapid Assessment of Pandemic Impact on Development-Early Childhood), or developing new survey items. Where applicable in this chapter, we provide information on scales' performance (that is, scale reliability) based on developer information or from prior FACES studies. In Chapter VII, we provide information on scale reliability in the 2021–2022 study. Appendix B contains the copyright permission references for any copyrighted instruments used in the TCR. Appendix D contains the 2021–2022 study instruments (with copyrighted items removed).

In April and May 2021, study staff at Mathematica conducted pretests of the parent surveys, teacher surveys, and Teacher Child Reports (TCRs) developed for the fall 2021 wave of the 2021-2022 Study, with the goals of ensuring the items in these instruments were clear and easy to understand, particularly items regarding the COVID-19 pandemic and items that had not been used in previous FACES studies. In May and June 2021, study staff at Mathematica conducted pretests of the teacher, center director, and program director surveys developed for the spring 2022 wave of the 2021-2022 Study, with similar goals as the earlier pretests. For the center director survey, an additional goal was to understand whether respondents could answer questions about how their centers use new quality improvement funds first awarded in 2020.

B. Teacher Child Report (TCR)

Head Start teacher ratings of children are important sources of information about children's learning and behavior, especially given that direct child assessments were not conducted as part of the 2021–2022 study. These teacher ratings, along with select items from the fall 2021 and spring 2022 parent survey, constitute the indirect child assessments for the 2021–2022 study.

1. Administration of TCR

Using a TCR, teachers in fall 2021 and spring 2022 of the Head Start year noted how children attended their class (in person, virtual or remote, hybrid) and rated each child on a set of items that assessed the child's current learning skills, cooperative classroom behavior (or social skills), problem behaviors (or classroom conduct), and approaches to learning. Teachers also provided reports of children's developmental conditions. The 2021–2022 study encouraged Head Start teachers to complete the TCR on the web for each sampled child. Head Start teachers without access to the web or with a preference for a paper TCR used a paper option. In Chapter IV, we provide details on the procedures for collecting the TCR data, the percentage of teachers completing it by web versus paper, and the response rates by wave. In Chapter VII, we provide details on summary scores derived from these items and current reliability. Appendix C details whether each measure was used in FACES 2014 or FACES 2019.

2. Children's attendance

Teachers answered questions about each child's attendance in their class. As in FACES 2019, they reported the number of days and hours per week the class met and how many days the child had missed class. For the first time, teachers reported whether children attended in person, virtually or remotely, or hybrid, and how many days per week and hours per day the teacher saw the child in person or virtually.

Additionally, the teacher reported the type of class—for example, a full-day class—and which days of the week the class met.[29]

3. Children's current learning skills

Teachers used a series of 15 items to rate each child's prereading, early mathematics, early writing, fine motor, and language skills by describing the child's ability in these areas. Nine items used in previous FACES studies and adapted from the National Household Education Survey (NHES) (U.S. Department of Education, National Center for Education Statistics 2000) assessed children's accomplishments in a variety of tasks, such as recognizing letters of the alphabet, counting, and writing. We created a summary score for teacher-reported child literacy behaviors using five of these items; see Chapter VII for details.

Teachers also reported on items that were not used in previous FACES studies, including one item from the Early Childhood Longitudinal Study, Kindergarten (ECLS-K) Class of 2022–23, 2020 field test preschool parent survey (U.S. Department of Education, National Center for Education Statistics 2020), which assessed children's ability to recognize basic shapes. Five items developed by Mathematica assessed children's ability to describe differences between shapes, sort and order objects by attributes, subitize items, and understand addition; these were included for the first time as part of the 2021-2022 study. Scales differed by item, with most items rated on a scale of 1 to 4 or by a yes or no response.

4. Children's cooperative classroom behavior (social skills)

Twelve items in the TCR addressed how often a child engaged in cooperative classroom behaviors and social skills, such as following the teacher's directions, helping to put things away, complimenting classmates, and following rules when playing games. The items were drawn from the Personal Maturity Scale (Entwisle et al. 1987) and the Social Skills Rating System (SSRS) social skills scale (Gresham and Elliott 1990; Elliott et al. 1988) and have been used in previous FACES studies. The teacher indicated the extent to which a given statement was characteristic of the child, rated from 1 (never) to 3 (very often). The teacher also had the option of selecting that they did not have an opportunity to observe the child. We created a summary score for teacher-reported social skills, with high numbers indicating more frequent cooperative behavior. In FACES 2006, 2009, 2014, and 2019, the internal consistency reliability (Cronbach's alpha) coefficients ranged from 0.88 to 0.92.

Like in previous FACES studies, the 2021–2022 study uses items from the Personal Maturity Scale and SSRS to create its own subscale, and the full source scales have strong reliability. The full Personal Maturity Scale, which measures a child's interest or participation, cooperation or compliance, and attention span or restlessness, consists of 13 items forming three subscales, with Cronbach's alpha reliabilities ranging from 0.74 to 0.85. The SSRS social skills scale has a Cronbach's alpha coefficient of 0.94 for preschool teacher report forms.

5. Children's problem behaviors (classroom conduct)

Fourteen items in the 2021–2022 study's problem behaviors scale, which measure negative child behaviors associated with learning problems and later grade retention, come from an adaptation of the Personal Maturity Scale (Entwisle et al. 1987; see above for description), the SSRS problem behaviors

[29] This information was requested in spring 2022 to allow Mathematica to link children to the specific classrooms taught by teachers with multiple class sections; these questions were not necessary in earlier FACES studies where classrooms were the sampling unit, rather than teachers.

scale (Gresham and Elliott 1990; Elliott et al. 1988; see above for description), and the Behavior Problems Index (BPI) (Peterson and Zill 1986). The BPI includes behaviors such as "undercontrol" (for example, aggression, hyperactivity, and destructiveness) and "overcontrol" (for example, social withdrawal, depression, and somatic problems). These items have been used in previous FACSE Studies. As in those studies, the 2021–2022 study created its own problem behaviors scale, and the full source scales have strong reliability. The internal consistency of the BPI total score ranged from 0.88 to 0.89 in the National Health Interview Survey and the National Longitudinal Study of Youth (Berry et al. 2004).

Teachers responded to questions also used in previous FACES studies about the frequency of aggressive behavior (such as "hits/fights with others"); hyperactive behavior (such as "is very restless"); and anxious or depressed and withdrawn behavior (such as "is unhappy") by using a scale from 1 (not true) to 3 (very true or often true). The teacher also had the option of selecting that they did not have an opportunity to observe the child. We created summary scores for disruptive/aggressive behavior, hyperactive behavior, withdrawn behavior, and total problem behaviors, with higher scores indicating that the child exhibits problem behaviors more often. The full SSRS problem behaviors scale has a Cronbach's alpha coefficient of 0.84 for preschool teacher report forms. In FACES 2006, 2009, 2014, and 2019, the internal consistency reliability (Cronbach's alpha) coefficients ranged from 0.77 to 0.88.

6. Children's approaches to learning

Teachers rated each child in the 2021–2022 study on a scale of 1 (never) to 4 (very often) on the six items that compose the Approaches to Learning Scale from the ECLS–K (U.S. Department of Education, National Center for Education Statistics 2002). The teacher also had the option of selecting that they did not have an opportunity to observe the child. The items assess a child's motivation, attention, organization, persistence, and independence in learning. The scale has been used with diverse populations. It is relatively short, has established reliability (Cronbach's alpha = 0.89), and has demonstrated relationships with academic achievement in elementary school (Bodovski and Young 2011). We calculated a summary score for teacher-reported approaches to learning, where higher scores indicate the child exhibits positive approaches to learning behaviors more often. The internal consistency reliability (Cronbach's alpha) coefficients in FACES 2009, 2014, and 2019 ranged from 0.91 to 0.93.

7. Developmental conditions or concerns

To provide context for children's Head Start experiences and their status and growth in skills predictive of school readiness, teachers responded to questions adapted from the National Early Intervention Longitudinal Study (NEILS) Family Enrollment Interview 2003, which had been used in previous FACES studies. Teachers noted whether a child had a diagnosed disability or was undergoing evaluation for disabilities or developmental concerns and whether the child had an Individualized Education Plan (IEP) or Individual Family Service Plan (IFSP). They described specific concerns about the child's health or development and indicated any efforts to address such concerns. These items do not constitute a scale. We created composite indicators for whether the child had an IEP, the child's disability status, whether the child had multiple disabilities, and whether the child's disability was sensory, physical, speech, cognitive, or behavioral. See Chapter VII for more information on these composite variables.

C. Parent survey

Head Start identifies families as the primary influence in children's lives (Bradley et al. 2001; Amato and Fowler 2004; Weiss et al. 2006; Moiduddin et al. 2010; Aikens et al. 2017a, 2017b). As laid out in Head

Start's Parent, Family, and Community Engagement (PFCE) framework, preparing children for school requires a partnership between programs and families that addresses the needs and well-being of both the child and the family. Families' characteristics and experiences directly influence child outcomes, and are a factor in the school readiness skills associated with children's classroom performance (Bryk and Schneider 2003; Burchinal et al. 2002; Fantuzzo and McWayne 2002; Lopez et al. 1999; Yoshikawa et al. 2013). The 2021–2022 study parent survey collected key information on the characteristics of Head Start children and their families, including characteristics and experiences shaped by the COVID-19 pandemic.

1. Administration of parent survey

Parents and caregivers had the option of completing the parent survey online or by telephone in either English or Spanish. Parent surveys took place in fall 2021 and spring 2022. The survey took about 35 minutes to complete in the fall. In the spring, the survey took about 31 minutes for families that had completed a fall 2021 survey, and about 47 minutes for families completing a survey for the first time. In Chapter IV, we provide details on the survey administration, use of web-based versus telephone surveys, and response rates.

2. Relationship of the parent survey respondent to the child

The section of the 2021–2022 study parent survey about the composition and structure of the household included a question about the survey respondent's relationship to the Head Start child (whether the respondent was the child's biological or adoptive parent, legal guardian, or relative), which was also included in previous FACES studies. Responses to this item provided background on the person identified as the individual primarily responsible for the care of the child and who answered the questions about child and family characteristics. Additional questions on family household structure (see Section C.4) focused on the adults available to provide resources, caregiving, and attention to children.

3. Children's demographic information

The enrollment of newly entering children has increased in recent years, from 66 percent in fall 2014 to 88 percent in fall 2019 (Aikens et al. 2017b; Kopack Klein et al. 2021b). Hispanic and Latino/a children represent a significant share of the Head Start population; at least one-third of children in fall 2014 and fall 2019 were Hispanic or Latino/a (Tarullo et al. 2017; Kopack Klein 2021b). Changes in the composition of the Head Start population have implications for other characteristics of the population and for program services.

The 2021–2022 study's parent survey collected information on children's demographics, such as race/ethnicity and gender. The 2021–2022 study also used the parent survey in conjunction with the Sample Management System (SMS) for some demographic information. For example, for analytic purposes, when age and gender data were not available in the SMS from the child consent form, we extracted this demographic information from the parent survey. Similarly, when information on Head Start exposure (that is, newly entering children versus returning children) or participation in Early Head Start was not available in our roster form, we extracted it from the parent survey.

4. Families' household structure

The 2021–2022 study collected information on families' household structure because research has indicated that the number of parents in a household can impact a wide range of cognitive and social-emotional outcomes (Bronte-Tinkew et al. 2004; Dawson 1991; McLanahan and Sandefur 1994;

Morrison and Cherlin 1995; Ribar 2015; Peterson and Zill 1986) during childhood and into adulthood (Amato 2005; Aquilino 1996). Household structure can affect families' access to economic resources (Amato 2005; Bane and Ellwood 1983; Garfinkel and McLanahan 1986; Ribar 2015), parenting practices (Amato 2005; Ribar 2015), and overall experiences of stress (Amato 2005).

The total number of household members can also reveal possible overcrowding, which could adversely affect a child's well-being or health. Measuring the number of children in a household also provides some indication of the extent to which parents must divide their caregiving and attention among children. For these reasons, the 2021–2022 study collected a complete roster of all household members, including their age and relationship to the child in the sample.[30] The survey also included items about the marital status of the child's parents and, for parents not legally married, the nature of their relationship.[31] These items were all included in previous FACES studies. In addition, respondents described how frequently the child saw a nonresident parent who did not live in the household. We created composite variables for household size, family structure, parent marital status, single-parent households, whether both parents were born in the United States, and whether the parent lives in shared housing; see Chapter VII for more details.

5. Joint book reading frequency

To understand the factors influencing children's development, it is critical to measure aspects of the home learning environment. Research has shown that children with stimulating home learning environments demonstrate higher cognitive skills, better social-emotional development (with more positive social skills and fewer problem behaviors), and more positive approaches to learning (Baker and Rimm-Kaufman 2014; Bradley et al. 2001; Fantuzzo et al. 2004; Foster et al. 2005; Korucu et al. 2020; McWayne et al. 2004). Research has also suggested that cognitive stimulation early in life may have implications for brain development and cognitive potential (Shonkoff and Phillips 2000). By asking about the frequency with which family members read books to their child, the 2021–2022 study instrument obtained information about the home literacy environment. Previous FACES studies also asked about the frequency of book reading in the home. We created composite variables indicating which families read to the sample child three or more times in the past week and told stories to the sample child three or more times in the past week; see Chapter VII for more details.

6. Home language environment

Children's language development is affected by the quality and quantity of their experiences with language, both inside and outside of the home (Dickinson and Tabors 2001; Hart and Risley 1995). For DLLs—that is, children who speak a language other than English at home—support for both English skills and home language skills is associated with optimal development across domains (Bialystok 2005; Fuligni et al. 2014). As such, it is important to understand children's home language environments. To assess these environments, the 2021–2022 study items asked whether a language other than English was spoken in the home, what it was, and what language was usually used to speak to the child at home. These items have been included in previous FACES studies.

[30] The parent survey captured this information both for individuals who usually live in the household and for those living there temporarily.

[31] As noted in Chapter VII, the composite variable for parent marital status includes options for registered domestic partnership or civil union, as well as for living together in a committed relationship.

7. Children's behavior and warmth in relationship between parents/caregivers and their child

a. Children's behavioral changes since March 2020

In March 2020, COVID-19 was declared a pandemic by the World Health Organization and a public health emergency by the U.S. Centers for Disease Control and Prevention; in the immediate aftermath, households experienced increased unemployment and income instability, and early care and education settings (including Head Start) closed their physical buildings and changed their operations. This disruption to daily life and economic stability could have a negative effect on children's development (RAPID-Survey Project 2023; Gassman-Pines and Gennetian 2020; Gassman-Pines et. al 2022; Liu and Fisher 2022; Kratch et. al 2021; Russel et. al 2020) and therefore was important to capture. Parents and caregivers answered four new items indicating whether their child had experienced any of four areas of behavioral changes since March 2020 (including the development of new fears, increase in tantrums, physical ailments, and disrupted sleep). Parents were asked to indicate the presence or absence of these changes.

b. Children's approaches to learning

With children experiencing virtual instruction and more time at home, the 2021–2022 study asked parents to rate their child's approaches to learning for the first time, as a parallel to teacher ratings. Parents and caregivers rated their child on a scale of 1 (never) to 4 (very often) on six items similar in content to those of teachers, using the teacher version of the Approaches to Learning Scale from the ECLS–K (U.S. Department of Education, National Center for Education Statistics 2002). The study team adapted the language of the items for use with parents and caregivers in the home setting to assess a child's motivation, attention, organization, persistence, and independence in learning. The original teacher scale has been used with diverse populations with established reliability (Cronbach's alpha = 0.89) and has demonstrated relationships with academic achievement in elementary school (Bodovski and Young 2011).

c. Warmth of relationship between parents/caregivers and their child

The relationship between parents/caregivers and their child can provide an essential foundation for the child's development. Parents and caregivers completed six items from the preschool parent survey field test for the ECLS-K 2023 (for the instrument, please see the OMB package OMB# 1850-0750 v. 21; Attachment B1). The items assess the warmth and closeness in parents' and caregivers' relationship with their child on a scale of 1 (not true at all) to 4 (completely true). We created a summary score for parenting warmth using these items; please see Chapter VII for more details. Information on scale performance is not available for these items from prior studies, as the 2021-2022 study predated the ECLS-K:2024, for which these items were field tested.

8. Household routines

Families' interactions with their children at home prepare them for socialization in school. One measure of family socialization takes the form of routines that parents establish for their children. Family routines have been linked to better self-regulation and fewer problem behaviors, as well as stronger health and cognitive skills for children (Rijlaarsdam et al. 2015; Ferretti and Bub 2014; Fiese and Schwartz 2008). The quantity and quality of children's sleep have been linked to children's cognitive functioning and physical health (Atkins-Burnett and Aikens 2011; Gaylor and Burnham 2010; Patel and Hu 2008; Ravid et al. 2009; Schlieber and Han 2018; Sekine et al. 2002; Taras and Potts-Datema 2005). Sleep quantity

and quality have also been linked to children's behavior and ability to pay attention (Atkins-Burnett and Aikens 2011; Hiscock et al. 2007; Hofferth and Sandberg 2001; LaVigne et al. 1999; Schlieber and Han 2018). For the 2021–2022 study, the parent survey asked about routines for certain household activities, such as meals, children's bedtime, and sleep quantity and quality. These items have been used in previous FACES studies.

9. School-age children's participation in virtual or remote learning activities

To better understand the learning experiences of children during the COVID-19 pandemic, the parent survey asked about the child's Head Start virtual or remote learning activities for the first time as part of the 2021-2022 study. Items were modified from the Rapid Assessment of Pandemic Impact on Development-Early Childhood (RAPID-EC) survey of families with young children (RAPID Survey Project 2023). Items asked whether the child had participated in virtual or remote learning activities, how engaged the child was in those activities, and how useful the activities were. In households with other school-age children, the parent survey also asked how those children were attending school (that is, in person, virtually/remotely, hybrid, or via homeschooling), and—if attending virtually/remotely or hybrid—who assisted with their online learning, whether that assistance occurred while the person was also working, and how many hours they helped with online learning. We created composite variables to indicate whether the study child had any virtual or remote instruction and whether a parent or caregiver assists with online learning during working hours; for more details, see Chapter VII.

10. Parents'/caregivers' demographic information

Several items assessed the demographic background of parents and caregivers. For example, we obtained information about the age, country of origin, length of residence in the U.S., education, employment, and employment changes due to the COVID-19 pandemic of resident mothers and fathers, that is, mothers and fathers who live in the same household as the study child.

11. Parents'/caregivers' material resources

The parent survey included items about the material resources available to and challenges faced by families, all of which potentially affect the quality of a child's home environment and subsequent development. We obtained information about household income; total family income; household income changes during the COVID-19 pandemic; receipt of income supports (such as participation in the Special Supplemental Nutrition Program for Women, Infants, and Children, better known as WIC); and receipt of government stimulus payments during the COVID-19 pandemic. Previous FACES studies collected information about household income and receipt of income supports, while items on total family income, household income during the COVID-19 pandemic, and receipt of government stimulus payments during the COVID-19 pandemic were asked for the first time as part of the 2021-2022 study.

The survey also captured information on families' housing status, whether they lived with another family for a financial reason, how many times they had recently moved (and the main reason for moving most recently), and the quality of their housing. These items were new to the parent survey as part of the 2021-2022 study.

In addition, given that a family's access to nutritious food can affect young children's health and cognitive development (Zaslow et al. 2009; Ryu and Bartfel 2012), we also obtained information on the family's food security and perspective on their own economic situation. Using the six-item short form of the U.S. Department of Agriculture's (USDA) Guide to Measuring Household Food Security, Revised

2000 (Bickel et al. 2000), which has been shown to correctly identify the level of food security for 97.7% of households (Blumberg et al 1999) and been used in previous FACES studies, parents indicated to what extent statements regarding food security described them (such as "I/we could not afford to eat balanced meals"). We created composite variables from each of these six items and a summary food security status scale; for more details, see Chapter VII.

We measured financial strain by using items from the Economic Strain Questionnaire (Conger et al. 1993) which have been used in previous FACES studies. Financial strain is derived from four items that measure the extent to which parents feel they have enough money to afford the type of home, clothing, food, and medical care they need. We created summary scores for material needs; see Chapter VII for details.

We measured material hardship—the lack of access to basic material needs—by asking parents to answer yes or no to seven items on housing insecurity (inability to pay rent or mortgage); lack of basic utilities (inability to pay for utilities, such as water or gas); and unmet medical needs (inability to pay for necessary medical care) from the Health Professional Opportunity Grants study, which were also used in FACES 2019. We created summary scores for housing insecurity, lack of basic utilities, and unmet medical needs; see Chapter VII for details.

12. Child care arrangements

Enrollment in Head Start can be one component of a family's nonparental child care participation. Many children spend time in nonparental care, and increasing numbers enter such care at a very young age. In fact, both the quality and consistency of child care over time are important factors that influence children's development. Research has found that the quality and consistency of care are related to children's cognitive and social-emotional development (Burchinal et al. 2014; Hayes et al. 1990; Love et al. 1996; Horm et al. 2018; National Institute of Child Health and Human Development [NICHD] Early Child Care Research Network 1998, 2000; Shonkoff and Phillips 2000; Whitebook et al. 1989; Zaslow 1991; Owen et al. 2008; Votruba-Drzal et al. 2004).

In spring 2022, the 2021–2022 study parent survey asked new items about families' reasons for enrolling in Head Start, their child care plans for the next year, and reasons for planning to send the child some place new (if applicable). The survey also asked about child care arrangements outside of their regular arrangements.

13. Family health and families' experiences with COVID-19

Children's health status can directly influence their well-being, development, and readiness for school (Janus and Duku 2007; Currie 2005; Gullo 2018; Lynch 2011). To assess the context of children's health care, the 2021–2022 study parent survey asked about the child's current health status and where the child usually goes for routine medical care. These questions were also asked in previous FACES studies, and a composite variable was created to indicate whether the child usually goes to the emergency room for routine medical care. Additionally, the parent survey asked new questions about the parent's or caregiver's current health status. Positive parent and caregiver health outcomes have been associated with child's health status, both generally (Murphey et al. 2018; Mensah & Kiernan 2010) and related to the COVID-19 pandemic (Patrick et al. 2020), and therefore were important to capture.

To better understand families' experiences with COVID-19, the survey also asked new questions about whether anyone in the household or close friends or family had contracted COVID-19 and, if so, which

individuals had done so, whether they had been hospitalized for it, and whether they had died from COVID-19. Composite variables were created from these items; see Chapter VII for details.

15. Community and social supports

A supportive community and social network can mitigate stressful life events and the stresses of daily living and parenting. Greater social support for families has been linked to more responsive parenting, lower rates of parental depressive symptoms, and greater involvement in the child's school (Burchinal et al. 1996; Crockenberg 1981; Jackson 1999; Sheldon 2002; Sharabi and Marom-Galon 2018). Because a family's social support network is highly important for favorable child outcomes (Frankel et al. 2014; Sarche et al. 2009), the spring 2022 survey asked a series of questions about the ability to access sources of social support when the family faces emotional, financial, and parenting problems or an emergency. Earlier national studies, including the ECLS–B, NHES, National Longitudinal Survey of Youth (NLSY), and FACES studies made extensive use of the same questions. The survey also include new items about whether specific community services (such as help with housing, finding or training for a job, unemployment assistance) would currently be helpful for the household or whether they would have been useful to anyone in the household in the last 12 months. For the first time, the survey also included items from the Healthy Families Parenting Inventory's Mobilizing Resources Subscale that focus on parents' and caregivers' knowledge about and comfort with accessing community supports and resources (Krysik and LeCroy 2012), and has a Cronbach's alpha of 0.86 when used in home visitation programs (LeCroy et al. 2007).

16. Parent and caregiver mental health and parenting stress

Parents' and caregivers' mental health is of concern to Head Start because of its relevance to their well-being and interactions with their children (Downey and Coyne 1990; National Center on Parent, Family, and Community Engagement [NCPFCE] 2013; National Research Council and Institute of Medicine 2009; Shonkoff and Phillips 2000). Depressed parents or caregivers may have a withdrawn or intrusive parenting style, either of which can lead to children's social and emotional problems, such as internalizing and externalizing behavior problems and difficulties in reading social cues (Downey and Coyne 1990; Field 2000; Shonkoff and Phillips 2000). Similarly, parental psychological distress (e.g., anxiety) may have insensitive and unsupportive parenting practices, which can be linked to child and adolescent maladjustment, as well as poorer psychological outcomes (Glasheen et al. 2010; Masarik and Conger 2017).

The 2021–2022 study parent survey included the short form of the Center for Epidemiologic Studies–Depression Scale (CES–D), which has also been used in previous FACES studies. The 12-item version of the CES–D (Radloff 1977; Ross et al. 1983) measures levels of depressive symptoms. Items are rated on a 4-point frequency scale (1 = rarely or never, 4 = most or all of the time). The internal consistency reliability estimates (Cronbach's alpha) of the parent's level of depressive symptoms (constructed from responses to the short form of the CES–D) ranged from 0.86 to 0.91 in FACES 2006 through 2014. We created a summary score for parents' depressive symptoms and a categorical indicator of parents' depressive scores; see Chapter VII for details.

For the first time as part of the 2021-2022 study, the survey also included items from the General Anxiety Disorder-7 (GAD-7; Spitzer et al. 2006) that measure anxiety levels, and an item asking about changes in parenting stress or anxiety since March 2020. The items (Spitzer et al. 2006) provide information on respondents' level of anxiety. Items were rated on a 4-point frequency scale (1 = not at all over the last

two weeks, 4 = nearly every day over the last two weeks). It has established reliability (Cronbach's alpha = 0.92) and has demonstrated relationships with other mental health outcomes (Löwe et al. 2008; Spitzer et al. 2006). We created a summary score for parents' anxiety level, a categorical indicator of parents' anxiety level, and an indicator of whether parent stress and anxiety increased compared to before the COVID-19 pandemic; see Chapter VII for details.

Parents and caregivers also rated the amount of stress and confidence in their parenting for the first time. Using six items from the 2021 ECLS-K parent survey field test, they rated items on a scale of 1 (rarely or never) to 5 (always or most of the time).

17. Head Start involvement and satisfaction

Head Start, which considers family involvement to be a central tenet, recognizes the importance of family engagement in children's learning and development (NCPFCE 2013; McWayne et al. 2004) and how strong parent-staff relationships can support family engagement (Blue-Banning et al. 2004; Spielberg 2011). Research has shown that family involvement in programs and schools (volunteering, participation on committees, accompanying children on field trips, and so forth) is related to improved outcomes for children during both early childhood and the elementary school years (Downer and Mendez 2005; Glick and Hohmann-Marrott 2007; Marcon 1999; McWayne et al. 2004; Toldson and Lemmons 2013; Vinopal 2018).

The spring 2022 survey asked how often parents and caregivers participated in a variety of activities in the Head Start center (volunteering in the classroom, attending workshops, accompanying children on field trips, attending parent-teacher conferences, and so forth). These questions had been included in previous FACES studies. For a subset of activities, the spring 2022 survey added items asking parents and caregivers to report whether they would prefer to participate in the activity in person, virtually/remotely, or both.

In the spring 2022 survey, parents responded to three questions about programs' cultural competence from the Strengths-Based Practices Inventory (Green et al. 2004), including whether parents believed that Head Start program staff respected their family's cultural and/or religious beliefs, encouraged them to learn about their culture and history, and had materials for their child that positively reflected their cultural background. Finally, Head Start parents/caregivers reported on their level of satisfaction with program efforts to promote children's development and support parents in meeting family needs; whether Head Start provided transportation for their child; and satisfaction with the proximity of the center to their home, operating hours of the center, transportation provided by Head Start (if applicable), and the respect and encouragement that program staff showed for the family's cultural and religious background. These items had been included in previous FACES studies. No summary scores were created for these items.

18. Challenges and supports during the COVID-19 pandemic

Finally, the survey included two new open-ended items to understand families' challenges and supports during the COVID-19 pandemic. These items asked respondents to describe in their own words how the losses and challenges of COVID-19 and current events related to racial injustice in the U.S. impacted their community, and what they found most helpful to cope with these challenges.

D. Teacher surveys

The 2021–2022 study teacher surveys collected information on the characteristics of classrooms and teachers related to the quality of care and education provided by the Head Start program. In fall 2021, the survey asked teachers about their emotions, feelings, beliefs, and experiences; the teacher's demographic and educational background; teaching experience; mental and physical health; and the impact of the COVID-19 pandemic on their experiences both as teachers and as parents or caregivers in their own families. In spring 2022, the survey asked teachers about their emotions, feelings, beliefs, and experiences, as well as about supports received for teacher wellbeing. Additionally, the survey asked about classroom activities and their classroom environment, including the frequency of mathematics and language- and literacy-oriented activities in the classroom; parent involvement; languages used in the classroom; use of curricula; primary assessment tool; and the teacher's demographic and educational background, teaching experience, and professional development activities. Teachers also reported on the overall behavior of the children in their classroom and about children with developmental concerns in their classroom.

1. Administration of teacher survey

The 2021–2022 study Head Start teachers completed surveys in fall 2021 and spring 2022 via web-based questionnaires, although they had the option of completing a paper version of the survey.[32] The fall survey took about 10 minutes to complete. In the spring, the survey took about 41 minutes for teachers who had completed a fall 2021 survey, and about 50 minutes for teachers completing a survey for the first time. In Chapter IV, we provide details on how the survey was administered, teachers' choice of web-based versus paper surveys (including the procedures for collecting web-based and paper surveys, and the percentage of teachers completing each type), and their response rates.

2. Learning activities

A teacher's influence in the classroom is evident through the variety of learning materials provided to stimulate both fine and gross motor development, creative and dramatic play, language and literacy, mathematics and science skills, and appreciation of cultural diversity. High-quality classrooms may adopt a planned approach that provides for both small-group activities and opportunities for individualized learning through free play and structured activities.

The classroom activities and stimulating environment provided by the teacher can have both direct and indirect effects on children's development (Duncan and Magnuson 2013). The 2021–2022 study teacher survey asked several questions about classroom activities. For example, teachers were asked to report on the scheduled learning activities in their classroom and to estimate the amount of time they spent on teacher-directed and child-selected activities in a typical day. The survey also asked how often the children participated in various reading, language, and mathematics activities. It also captured more broadly the amount of time spent across domains to include language and literacy, mathematics, science, social studies, the arts, and social-emotional development. These items were all included in previous FACES studies.

[32] Home visitors in the 60 programs with child-level data collection also completed the teacher survey and treated their caseload as a classroom.

The teacher survey captured the classroom's language environment to determine the languages spoken by the staff and children. Head Start's program performance standards require programs to support children's progress in learning their home language while recognizing the cultures represented in the classroom. The survey asked teachers about (1) the number of children in the classroom who spoke a language other than English; (2) the percentage of children who spoke each of the languages other than English; (3) whether the teacher spoke a language other than English and, if so, what languages they used in the classroom; (4) the languages used for instruction; (5) which staff in the classroom spoke the languages used for instruction; (6) the language of print materials; (7) whether the teacher completed a course on how to work with children who spoke a language other than English; and (8) how the teacher communicated with families who spoke a different language than the teacher. These items were all included in previous FACES studies. Composite variables were created to indicate whether a teacher's classroom included children who are dual language learners, what non-English languages were spoken by children in the classroom, and whether instruction occurred in English and in any non-English languages; see Chapter VII for details.

3. Classroom environment

Other factors, such as group size and the child-to-adult ratio, can indirectly affect a child's experience in Head Start by influencing the availability of stimulating resources and shaping the teacher's behavior in directing and facilitating the child's learning (ACF 1998, 2003). With more children and fewer adults in a classroom, a teacher is less able to give a child individual attention, create opportunities for learning during children's play, and prevent negative behavior.

The research literature provides evidence on the important contribution of structural factors to children's outcomes in early childhood classrooms (Hayes et al. 1990; NICHD Early Child Care Research Network 1998, 2000, 2006; Shonkoff and Phillips 2000; Whitebook et al. 1989; Zaslow 1991, Zaslow et al. 2010a). Items in the teacher survey collected information about the number of children in the classroom, classroom staffing, and the overall behavior of the children in the classroom. The survey also asked about the age and racial and ethnic composition of the children in the classroom, and the skill level of children in various domains. These items were all included in previous FACES studies. Composite variables were created to capture class size, child/teacher ratio, and child/adult ratio; see Chapter VII for details. The additional background on children in the classroom allows users to develop a picture of the children across all 176 programs in the Program and Staff Study (given that child-level data were collected only in the 60 programs participating in the Program, Staff, and Family Study).

4. Planning, curricula, and assessment

The 2021–2022 study instrument asked Head Start teachers what curricula they used, including those for teaching math and literacy. Teachers also answered questions about frequency of use, whether they received training and support in using the curricula, and whether a tool or checklist was used to assess curricula implementation. In addition, teachers responded to a series of questions on their main curriculum and whether they received training on and support for its use. If they indicated using Creative Curriculum, they were asked which edition they used and whether they used any of the curriculum's additional resources. For the 2021–2022 study, teachers described their main assessment tool, the methods used for assessment, the main tool's use in planning, and training on and support for using the tool. These items had all been included in previous FACES studies. Composite variables were created to capture whether the teacher uses a curriculum with an available aligned assessment tool, the main classroom curriculum, whether the teacher uses an aligned curriculum and assessment tool, and the count

of curricula used. Additional composite variables were created related to comprehensive curricula, math curricula, literacy curricula, social-emotional curricula, and other curricula. See Chapter VII for details on these composite variables.

5. Professional development

Professional development for staff can support program quality, improve teachers' practice, and support children's school readiness (Bloom et al. 2013; Douglass 2017; Whalen et al. 2016; Zalsow et al. 2010b). Teachers' ongoing professional development can provide a source of information about best practices. For the 2021–2022 study, teachers provided information about whether different professional development supports were offered and whether mentor teachers and coaches were available (informally or formally). In the case of mentors or coaches, teachers indicated whom their mentors or coaches were, how often they worked with them, and the methods their mentors or coaches used to support them. These items had all been asked in previous FACES studies. No summary scores were created from these items.

Newly as part of the 2021-2022 study, the survey gathered information about trauma-informed care, including whether teachers received training on providing such care and who provided the training.

6. Parental involvement

The relationship between parent and teacher is an important aspect of program quality. Communication and collaboration between parent and teacher support parent engagement (Spielberg 2011), and ultimately, child outcomes (Blue-Banning et al. 2004). In particular, research has demonstrated that the degree of attunement between child care practices at home and the care delivered in other environments is an important factor in child outcomes (van IJzendoorn et al. 1998). Questions in the 2021–2022 study assessed how often teachers met with parents (including meeting to discuss the progress or status of children with and without developmental concerns), and how the program promoted interactions between parents and staff; these items were also included in previous FACES studies. No summary scores were created from these items.

7. Teachers' beliefs about teaching practices

Teachers' knowledge and beliefs about developmentally appropriate teaching practices may affect classroom quality and the type and number of learning activities made available to children. FACES 2000 found that teachers' beliefs and knowledge about early childhood teaching practices constituted a conduit between their educational level and observed classroom quality. Teachers with higher degrees had higher scores on the teacher beliefs scale and higher observed quality in their classrooms (ACF 2003). FACES 2006 showed that teachers' beliefs were positively related to their observed classroom quality (based on the ECERS–R Provisions for Learning and CLASS Language Modeling) and children's social skills (Aikens et al. 2010). FACES 2009 found that teacher credentials—education and experience—were associated with the emotional climate of a classroom, as well as teacher satisfaction and attitudes (Moiduddin et al. 2012).

The assessment of Head Start teachers' knowledge and beliefs about early childhood teaching practices in the 2021-2022 study was based on 15 statements representing opinions on how Head Start children should be taught and managed (Burts et al. 1990). For example, two such statements posit that "Head Start classroom activities should be responsive to individual differences in development" and "Students should work silently and alone on seatwork." Teachers indicated whether and to what extent they agreed or disagreed with each of the 15 statements along a 5-point Likert-type scale (1 = strongly disagree, 5 =

strongly agree). These items have been used in previous FACES studies. Mathematica created summary scale scores for teachers' beliefs about developmentally appropriate practice, teachers' didactic beliefs, and teachers' beliefs about child-initiated classroom activities. In FACES 2009, 2014, and 2019, the internal consistency reliability (Cronbach's alpha) for the developmentally appropriate attitudes, child-initiated subscale ranged from 0.66 to 0.79 and, for the developmentally appropriate attitudes, didactic subscale, from 0.70 to 0.78. For more information about these subscales, see Chapter VII.

8. Teachers' likelihood of continuing to work for Head Start

High rates of teacher turnover are a persistent concern in Head Start and other early care and education programs (Bassok et al. 2013; Phillips et al. 2019; Whitebook et al. 2014). To understand the factors affecting such turnover, the survey gathered information about teachers' likelihood of continuing to work for Head Start. Teachers reported on statements about their enjoyment of teaching and the likelihood that they would continue working for Head Start through the next year. For teachers who reported they were very likely or somewhat likely to continue working for Head Start, the survey gathered information about the reasons they would stay with Head Start. For teachers who reported they were very unlikely or somewhat unlikely to continue working for Head Start, the survey gathered information about the reasons they would leave Head Start. These items have been included in previous FACES studies. No summary scores were created from these items.

9. Teachers' mental health and current health status

Teachers' self-reports about their mental health provide critical information about the environment of Head Start classrooms and teachers' interactions with the children in them. In fact, research has documented links between teachers' psychological well-being—including depressive symptoms and workplace stress—and the quality of care received by children in their classrooms (Gerber et al. 2007; Hindman and Bustamante 2019; Whitaker et al. 2015). A recent study showed that teachers who reported more depressive symptoms also reported less positive family-teacher relationships, which in turn was negatively associated with children's approaches to learning skills and mathematics achievement (Jeon et al. 2021). Another study showed that teachers who reported more depressive symptoms were less likely to have high-quality classroom learning environments, which led to a negative effect on children's mathematics achievement (McLean and Connor 2015).

In keeping with the approach of past FACES studies, the 2021–2022 study included the short form of the CES–D, with 12 items rated on a 4-point frequency scale (1 = rarely or never, 4 = most or all of the time). The items (Radloff 1977; Ross et al. 1983) provide information on respondents' levels of depressive symptoms. Based on its use in prior FACES studies, the scale has demonstrated variability (with about 20 percent of teachers reporting moderate or severe level of depressive symptoms; Doran et al. 2022). Also, teachers who report higher depressive symptoms exhibit lower observed language modeling behaviors in the classroom (Aikens et al. 2010). The internal consistency reliability (Cronbach's alpha) of teachers' depressive symptoms scores ranged from 0.80 to 0.90 in FACES 2006 through 2019. We created a summary score for teachers' depressive symptoms and a categorical indicator of teachers' depressive scores; see Chapter VII for details.

New in the 2021-2022 study, the teacher survey also included the seven-item GAD-7. These items (Spitzer et al. 2006) provide information on respondents' level of anxiety. Items were rated on a 4-point frequency scale (1 = not at all over the last two weeks, 4 = nearly every day over the last two weeks). It has established reliability (Cronbach's alpha = 0.92) and has demonstrated relationships with other mental

health outcomes (Löwe et al. 2008; Spitzer et al. 2006). We created a summary score for teachers' anxiety level and a categorical indicator of teachers' anxiety level; see Chapter VII for details.

The teacher survey also gathered information about teacher stress for the first time as part of the 2021-2022 study. It included the 4-item stress subscale from the Texas Christian University Survey of Organizational Functioning (TCU SOF). These items (Institute of Behavioral Research 2005) provide information about teachers' level of agreement with statements about job-related stress at the center; they are rated on a 5-point frequency scale (1 = strongly disagree, 5 = strongly agree). The survey also included four items that gathered information about teachers' current job stress due to the COVID-19 pandemic. These items asked about teachers' level of agreement with such statements as "You worry about your own potential exposure to COVID-19 while at work." Items were rated on a 5-point agreement scale (1 = strongly disagree, 5 = strongly agree). We created summary scores for teachers' COVID-19 job stress and general job stress; see Chapter VII for details.

Understanding Head Start teachers' physical health can provide a more complete picture of their well-being (Hascher and Waber 2021; Hall-Kenyon- et al. 2014). This issue is particularly important in light of research showing that the COVID-19 pandemic affected early childhood educators' mental and physical health (Swigonski et al. 2021). For the first time, the 2021–2022 study teacher survey asked about the teacher's current health status. Teachers were asked to rate their health on a 5-point scale (1 = excellent, 5 = poor).

10. Supports for staff wellness and overall well-being

For the first time, the 2021–2022 study gathered information about the supports for staff wellness and overall well-being offered to and used by teachers. The teacher survey asked about the supports for staff wellness and overall well-being offered by teachers' programs (such as professional mental health consultations or resources to support teachers' physical health and safety) and those supports teachers used or received. The teacher survey also asked whether the supports were offered at a convenient location and time, whether supports met teachers' needs, and if there were supports that would have been useful but were not offered by their program.

11. Impact of the COVID-19 pandemic

For teachers with children for whom they were the primary parent or caregiver, the teacher survey asked five items about this aspect of their experiences during the COVID-19 pandemic. The survey asked about teachers' stress or anxiety as a parent or caregiver, including parenting behaviors and stress, and changes in this stress or anxiety since March 2020. The survey also gathered information about whether teachers' children attended virtual or in-person school or programming, and whether teachers needed to use strategies to meet child care needs outside of their regular child care arrangements. We created composite variables related to teachers' experiences as caregivers; see Chapter VII for more details.

For all teachers, the survey gathered information about the challenges they faced due to the COVID-19 pandemic and the types of supports they found useful. The survey included open-ended questions about the biggest challenges teachers and their families faced during the pandemic, and the biggest challenges they experienced as a teacher. The survey also included open-ended questions about the supports from Head Start that were the most useful during the pandemic.

These items had not been included in any previous FACES studies.

12. Teachers' backgrounds

Head Start teacher characteristics—a class of structural factors related to early care and education quality—include education and training. The Head Start teacher survey asked about the classroom teacher's teaching experience (number of years teaching in Head Start and overall); educational background (the highest grade or year of school completed); teaching credentials (for example, Child Development Associate or state-awarded certificate or license); current enrollment in training or education programs; and demographics (age, gender, race, and ethnicity). The survey also collected information about teacher salaries. Research has demonstrated that average teacher salary levels in a program, which are a measure of program resources, are related to gains in measures such as letter knowledge and cooperative behavior (ACF 2003; Zill et al. 2005), though studies have not shown that salary is associated with child outcomes when accounting for observed quality (Aikens et al. 2010). These items had been included in previous FACES studies. We created composite variables for teacher race and teacher's years teaching Head Start; see Chapter VII for details.

D. Program director survey

In this section, we present the survey of Head Start program directors, which provides important information about the context of services for families and children, as well as factors that promote high-quality services, including information about leadership training, center licensing, and revenue outside of Head Start funding. Program-level factors—both the characteristics of program leaders and the policy and processes within programs—are central to classroom quality and to supporting child outcomes. Other important elements of program policy and process include professional development offered to staff and elements of data use for guiding decisions. In spring 2022, the 2021–2022 study surveyed program directors to address these and other topics. In that survey, items gathered information about directors' mental health and the types of staff compensation, benefits, and supports for staff well-being provided by programs. In addition, because the COVID-19 pandemic had potential impacts on Head Start program operations, the survey included items regarding programs' emergency preparedness and lasting changes to programs as a result of the pandemic.

1. Administration of program director survey

Program directors completed the survey in spring 2022 via web-based questionnaires, which took an average of 59 minutes for respondents to complete. In Chapter IV, we provide details on how the survey was administered and the response rate.

2. Program characteristics

In spring 2022, program directors confirmed some information from the most recent version of the Head Start PIR submitted by their agency—for example, whether the program had a management information system and about the program's funding and enrollment. These items had been included in previous FACES studies. We created composite variables related to program funding; see Chapter VII for details.

For the first time, the spring 2022 program director survey also gathered information about whether programs made an effort to recruit or start recruiting specific groups of families (for example, families experiencing unemployment or underemployment), and whether it was difficult for them to recruit any such groups.

3. Staff education and training

Leaders play a role in staff productivity, empowerment, proactivity, and approaches to implementation (Boyatzis and McKee 2005; Greenhalgh et al. 2005; Kruse 2001; Wang et al. 2017). Important elements of program policy and process include professional development offered to staff (Aarons et al. 2011; Whitworth and Chiu 2015). The survey asked program directors about efforts to promote staff education and training. It also asked about staff education initiatives, including the types of assistance offered to staff members (including family child care [FCC] providers and home visitors) working toward a college degree and eligibility for assistance. In addition, the survey asked about the coaching and mentoring offered to staff members. Program directors also indicated whether they participated in various types of professional development and identified areas in which they needed more support to lead their program more effectively. These items were included in previous FACES studies.

4. Program directors' mental health

Studies indicate that program leaders' working conditions can affect their mental health. Research on program leaders across a variety of industries shows that their mental health is related to inadequate resources and budget demands (Copeland and Kirsch 1995; Fry 1995), long hours (Hsu et al. 2016), work overload (Cooper 1984; Rogers et al. 1994; Worrall and Cooper 1995), and organizational pressures (Olinske and Hellmen 2017). Yet there is limited information available on the mental health of Head Start program directors. In spring 2022, the survey asked about program directors' mental health for the first time. The 2021–2022 study included the short (12-item) form of the CES–D (Radloff 1977; Ross et al. 1983) to gather information about respondents' depressive symptoms and the seven-item GAD-7 (Spitzer et al. 2006) to gather information on their level of anxiety. The CES-D has demonstrated reliability with Head Start teachers from previous FACES studies. The GAD-7 has established reliability (Cronbach's alpha = 0.92) and has demonstrated relationships with other mental health outcomes (Löwe et al. 2008; Spitzer et al. 2006). We created summary scores for program directors' depressive symptoms and anxiety level, as well as categorical indicators of program directors' depressive scores and anxiety level; see Chapter VII for details.

The survey also included four items that gathered information about program directors' current job stress due to the COVID-19 pandemic. The items asked about directors' level of agreement with statements about job-related stress due to COVID-19, such as "You worry about your own potential exposure to COVID-19 while at work." Items were rated on a 5-point agreement scale (1 = strongly disagree, 5 = strongly agree). These items had not been used with program directors in early FACES studies. We created a summary score for program director COVID-19 job stress; see Chapter VII for details.

5. Program activities

Leaders and managers create environments that promote success, including positive outcomes for children (Marzano et al. 2005), which are reflected in program activities and processes. In the program director survey, respondents answered questions about the main child assessment tool used in their program. They also reported on their program's communication with elementary schools attended by children from the program when they enter kindergarten, including the number of elementary schools into which the program feeds and the types of communication it uses with those schools. Program directors also reported on how much time they spent on some of their responsibilities over the course of the Head Start year and reported on the responsibilities of their program's educational coordinator. These items had been used in previous FACES studies.

6. Use of program data and information, systems, and resources

The use of data to guide decisions is an important element of program policy and process (for example, Ammerman et al. 2007; Caspe and Lopez 2006; Marsh et al. 2006; Spillane 2012). The program director survey inquired about the use of program data and information—particularly how the program's data were stored and how its management information system was set up. The survey asked whether the program relied on a designated staff member to analyze program data and, if so, whether that person focused only on analysis tasks and was trained in data analysis. Program directors also noted sources of revenue for their program other than Head Start, as well as licensing requirements and participation in their quality rating and improvement system. These items had been used in previous FACES studies.

7. Substance use issues in the community

Data from FACES 2019 show that a majority of program directors report that substance use and related problems, such as lack of resources for treatment, is either somewhat of a problem or a major problem in program communities (Doran et al. 2022). In spring 2022, the survey asked program directors whether substance use issues were a problem in their community, the types of supports they provided for staff who worked with children and families with a substance use problem, and if any of those staff supports were specific to the opioid epidemic.

8. Program directors' backgrounds

Research shows that the characteristics and qualifications of ECE directors, such as years of experience and level of education, are associated with program quality and processes, such as program and classroom quality ratings (McCormick Center for Early Childhood Leadership 2010; Talan et al. 2014) and supports for professional development provided to staff (Rohacek et al. 2010; Smith et al. 2019; Talan et al. 2014). Program directors replied to questions about their years of experience in various positions and early childhood programs (including their current Head Start program) and their salary, educational background, and demographics. These background questions had been included in previous FACES studies. We created a composite variable for program director race/ethnicity; see Chapter VII for details.

9. Staff compensation, benefits, and supports for staff well-being

Research shows that compensation and working conditions are key drivers of staff turnover in Head Start and other early care and education programs (Grant et al. 2019; Schaak et al. 2020; Wells 2015). To understand how programs are supporting their staff, the program director survey asked new questions about the types of compensation, benefits, and other supports programs provided to staff. Program directors replied to questions about staff compensation, including the types of compensation offered to staff (such as paid sick days and paid holidays); whether programs had added or increased staff compensation in the past 12 months; and whether programs had increased wages for specific positions (such as custodians and service staff) in the past 12 months. Respondents also replied to new questions about the types of supports for staff well-being offered (such as regular check-ins with staff and offering professional mental health consultations), and whether programs had added or increased supports for staff well-being in the past 12 months. In addition, the survey asked new questions about the types of activities programs had implemented to support staff in the past 12 months (such as increasing wages or hiring additional staff) and whether any of these activities were implemented with new funding for quality improvement. We created several composite variables related to staff compensation and well-being supports; see Chapter VII for details.

10. COVID-19 impact and emergency preparedness

Finally, the program director survey gathered information about changes to programs as a result of the COVID-19 pandemic. The survey asked program directors about their emergency preparedness, such as topics included in programs' emergency management and disaster preparedness and response plans, and the types of emergency management and disaster preparedness activities programs had conducted in the past 12 months. It also asked program directors open-ended questions about the greatest lasting change to their program as a result of the COVID-19 pandemic. These items had not been included in previous FACES studies.

E. Center director survey

In this section, we present the content for the survey of Head Start center directors, which provide important information about the context of services for families and children, as well as factors that promote high-quality services. Similar to the program director survey, this information was related to leadership training, center licensing, and revenue outside of Head Start funding. Center-level factors—both the characteristics of center leaders and the policy and processes within centers—are key to both classroom quality and supporting child outcomes. Important elements of center policy and process also include professional development offered to staff, use of activities and tools related to curriculum and assessment, and elements of data use for guiding decisions. In spring 2022, the 2021–2022 study surveyed center directors to address these and other topics. In the center director survey, items gathered information about center directors' mental health, the staffing challenges experienced by centers, and the supports for staff well-being provided by programs. In addition, because the COVID-19 pandemic had potential impacts on Head Start center operations, items regarding lasting changes resulting from it were included in the director surveys.

1. Administration of center director survey

Center directors completed the survey in spring 2022 via web-based questionnaires, which took on average 51 minutes for respondents to complete. In Chapter IV, we provide details on how the survey was administered and the response rate.

2. Center characteristics

Data from the 2021–2022 study center director surveys provide information needed for developing a multifaceted understanding of Head Start centers and families. Center directors reported on the schedule available for center-based enrollment slots, the number of currently employed lead teachers, and the number of newly hired lead teachers. They also described the language environment, including whether families and children spoke a language other than English, whether the center had any bilingual lead or assistant teachers, and how center staff communicated with families. We created composite variables for whether Spanish is spoken by teachers and families, the number of languages spoken by center families, the percent of family languages spoken by teachers, and the length of the Head Start year in months; see Chapter VII for details.

3. Staffing challenges

High rates of teacher turnover are a persistent concern in Head Start and other early care and education programs (Bassok et al. 2013; Phillips et al. 2019; Whitebook et al. 2014), and have been exacerbated by the COVID-19 pandemic (Farewell et al. 2022; Kim et al. 2022). The spring 2022 center director survey

asked about centers' challenges hiring and retaining staff. Center directors reported on lead teacher retention rates, the reasons lead teachers left the center, and challenges due to turnover among different types of staff (such as educational personnel and family services workers or child counselors or therapists). Center directors also reported on whether centers were able to find classroom coverage for teaching staff in the center and had enough staff to operate the center at full capacity. While previous FACES studies asked about staffing, recruitment, and turnover of lead teachers, the spring 2022 center director survey is the first time center directors were asked to report on problems due to staff turnover or these other staffing challenges. We created a composite variable for lead teacher turnover; see Chapter VII for more details.

4. Staff education and training

As described above for programs, the professional development offered to staff is an important aspect of center policy and process (Aarons et al. 2011; Whitworth and Chiu 2015). The spring 2022 center director survey asked about staff education and training. Items focused on training and professional development activities, as well as coaching or mentorship—for example, frequency of training, who conducts it, and training related to curriculum and assessment. The survey also asked center directors about the types of professional development services available to teachers, whether the center had consulted with training and technical assistance (T/TA) specialists, and the frequency of use of resources from national centers. In addition, the survey asked about the professional development activities offered to FCC providers and home visitors in addition to teachers. Directors also answered questions about how often teachers received a formal performance evaluation and whether teachers used data to guide instruction. Center directors indicated whether they participated in various types of professional development and identified areas in which they needed more support to lead their center more effectively. These items have been included in previous FACES studies.

5. Center directors' mental health and mental health supports available to staff

As described above for program directors, little evidence is available on the mental health of leaders in early care and education, despite evidence that aspects of working environments affect leaders' mental health (Cooper 1984; Copeland and Kirsch 1995; Fry 1995; Hsu et al. 2016; Olinske and Hellmen 2017; Rogers et al. 1994; Worrall and Cooper 1995). In spring 2022, the survey asked about center directors' mental health for the first time. The 2021–2022 study included the short (12-item) form of the CES–D (Radloff 1977; Ross et al. 1983) to gather information about respondents' depressive symptoms and the seven-item GAD-7 (Spitzer et al. 2006) to gather information on respondents' level of anxiety in the previous two weeks. The survey also included four items that gathered information about program directors' current job stress due to the COVID-19 pandemic. The items asked about center directors' level of agreement with statements about job-related stress due to COVID-19, such as "You worry about your own potential exposure to COVID-19 while at work." Items were rated on a 5-point agreement scale (1 = strongly disagree, 5 = strongly agree). The CES-D has demonstrated reliability with Head Start teachers from previous FACES studies. The GAD-7 has established reliability (Cronbach's alpha = 0.92) and has demonstrated relationships with other mental health outcomes (Löwe et al. 2008; Spitzer et al. 2006). We created summary scores for center directors' depressive symptoms and anxiety level, as well as categorical indicators of center directors' depressive scores and anxiety level; see Chapter VII for details.

Centers directors also responded to questions about supports for mental health available to staff, including whether centers offered services or supports to bolster staff wellness and overall well-being, the types of

supports offered, and whether centers offered training on trauma-informed care. These items were included for the first time as part of the 2021-2022 study.

6. Curriculum and assessment

The supports centers provide for teachers' use of assessment and curriculum tools can affect classroom environments (Penuel et al. 2007; Sarama et al. 2016; Weiland et al. 2018), which in turn can shape children's development (Duncan and Magnuson 2013). The center director survey asked about the use of assessment and curriculum tools. Center directors reported how often teachers made assessment results available to parents or program administrators, or recorded them in children's program records; how teachers assessed the English language abilities of DLL children who spoke a language other than English; and how data were reviewed. The survey also asked the directors whether the center used a parent education or support curriculum. These items were included in previous FACES studies.

7. Use of program data and information, systems, and resources

Similar to program directors, the use of data to guide decisions of center directors is an important element of program policy and process (for example, Ammerman et al. 2007; Marsh et al. 2006; Spillane 2012; Caspe and Lopez 2006). The center director survey inquired about the use of program data and information. It asked whether supervisors, coaches and mentors, or other specialists shared or reviewed individual children's data with teachers, and about the barriers to teachers' use of data to guide and individualize instruction. The survey also asked center directors about their center's licensing and monitoring, and its participation in their quality rating and improvement system. In addition, directors reported on their center's financial management, including who manages the center's finances, whether that person's primary responsibility was managing finances, and whether the director had training in financial management. These items were included in previous FACES studies.

8. Center directors' backgrounds

As described above for program directors, director qualifications and characteristics are associated with program quality and processes (McCormick Center for Early Childhood Leadership 2010; Rohacek et al. 2010; Smith et al. 2019; Talan et. 2014). Center directors responded to questions about their years of experience in various positions and early childhood programs (including their current Head Start program and center), and their salary and benefits, educational background, and demographics. These items were included in previous FACES studies.

9. COVID-19 impact

Finally, the center director survey gathered information about changes to centers as a result of the COVID-19 pandemic. It asked center directors open-ended questions about the greatest lasting change to their center as a result of the COVID-19 pandemic. These items were not included in previous FACES studies.

IV. Data collection procedures and response rates

To ensure uniformity and consistency across all modes of 2021–2022 study data collection, the study team developed, trained staff on, and implemented standardized procedures for each data collection step. This chapter provides an overview of the 2021–2022 study data collection procedures, describing the (1) team approach, (2) staff training, (3) sampling and recruitment procedures, (4) planning for and conducting data collection, (5) quality assurance, and (6) data collection response rates. Exhibit IV.1 lists the 2021–2022 study's data collection components, including modes and times.

The **2021–2022 study** included two components:

1. The **Program, Staff, and Family Study**, conducted in 60 programs, included the collection of parent surveys and TCRs in fall 2021 and spring 2022, as well as a teacher survey in fall 2021.
2. The **Program and Staff Study**, conducted in the 60 programs participating in the **Program, Staff, and Family Study** plus an additional 120 programs, included the collection of program director, center director, and teacher surveys in spring 2022.

Exhibit IV.1. Data collection for the 2021–2022 Study: administration characteristics of different components

Component	Fall 2021 Mode and time	Spring 2022 Mode and time
Sampling activities[a]	CADE with web	CADE with web
	About eight hours per program via remote contact	About four hours per program via remote contact
Parent survey[b]	Web and CATI	Web and CATI
	35 minutes	47 minutes for new respondents; 31 minutes for returning respondents
Teacher Child Report[b]	Web and paper	Web and paper
	10 minutes per child	7 minutes per child
Teacher survey	Web and paper	Web and paper
	10 minutes	41 minutes for new respondents; 50 minutes for returning respondents
Program director survey	NA	Web
	NA	59 minutes
Center director survey	NA	Web
	NA	51 minutes

[a]Study liaisons conducted remote sampling of teachers and children for all 60 programs in fall 2021. In spring 2022, liaisons conducted remote sampling of teachers in 114 of the additional 116 programs, plus one center in another program. Staff in the 60 programs participating in the Program, Staff, and Family Study collected information on all teachers at the selected centers and all children taught by the selected teachers in fall 2021, and updated this information in spring 2022. Staff in the 116 programs participating in the Program and Staff study collected information on all teachers at the selected centers in spring 2022.

[b]Information gathered from 60 programs participating in the Program, Staff, and Family Study; all other information (aside from sampling) was collected from all programs.

CADE = computer-assisted data entry; CATI = computer-assisted telephone interviewing; NA = not available.

In planning for and fielding the 2021–2022 study, the study team encountered several challenges. Broadly, the challenges related to a compressed timeline for recruitment relative to the timeline of

previous FACES studies, the ongoing pandemic and resulting nationwide child care crisis, and the time needed to retool the spring data collection to serve emerging initiatives. Given these challenges, we did not reach the target number of eligible and consented children for either study. With child-level instruments released only for consented children, the lack of completed consent forms limited our ability to reach the target number of cases needed for robust analyses of those instruments. Therefore, the findings from the 2021–2022 study are not nationally representative and do not lend themselves to comparison with any earlier rounds of FACES.

A. Team approach to mostly remote data collection

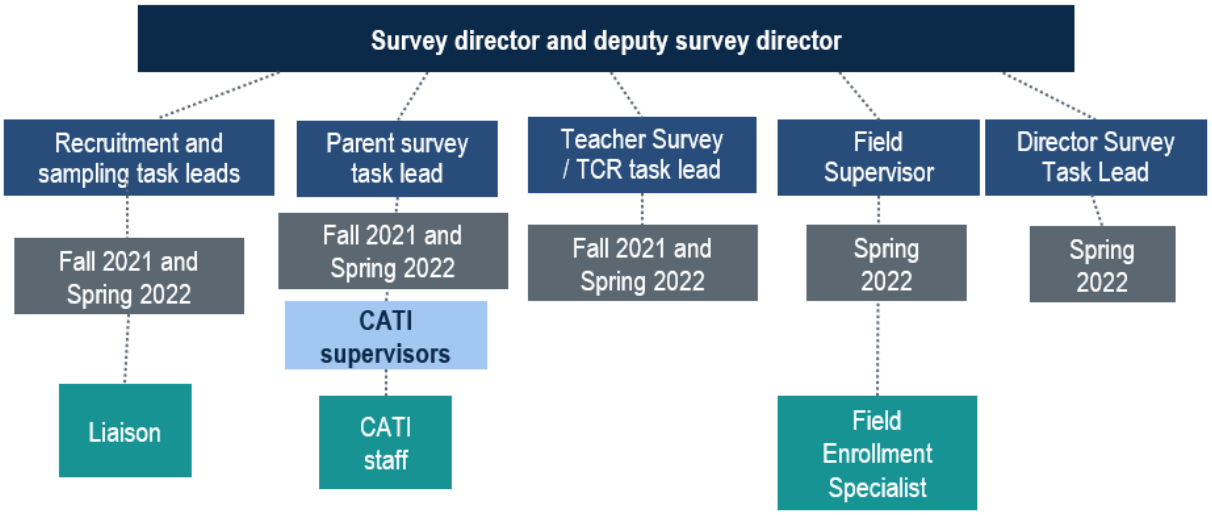
The evolving pandemic required a fully remote fall 2021 data collection, including sampling, distribution and collection of consent forms, parent surveys, teacher surveys, and TCRs. As discussed in Chapter II, the study did not reach its target number of consented children in the fall. As a result, we continued the consent collection process through the end of the spring 2022 data collection, adding in-person field enrollment specialist visits where programs allowed. To complete the remote recruitment, sampling, and data collection tasks efficiently, we employed a team of Mathematica liaisons, telephone interviewers, and field enrollment specialists. (Exhibit IV.2 shows the study team's structure.)

Although the criteria for working on the study team depended on the specific role, experience on previous FACES studies was preferred (but not required) for all staff. The criteria for selection of computer-assisted telephone interviewing (CATI) staff included customer service experience, as well as reading, presentation, listening, comprehension, and interpersonal skills. We also considered special skills, such as refusal avoidance and conversion. Because a large number of Head Start children and parents speak Spanish as their primary language, fluency in Spanish also was an important characteristic for hiring bilingual and CATI staff.[33] We selected field enrollment specialists based on their interpersonal skills, ability to work independently, and ability to operate a laptop computer and enter data quickly and accurately.

In Exhibit IV.2, we present the structure of the 2021–2022 study data collection teams and the relationships between and among staff. The team worked under the leadership of the survey director, deputy survey director, recruitment and sampling lead, instrument leads, and the field and CATI supervisors. The survey director and deputy survey director were directly involved with the data collection activities, and responsible for ensuring that all aspects of the data collection complied with contract requirements and Mathematica's professional standards; the field and CATI supervisors oversaw and supported the data collection team.

[33] Before collecting data, all bilingual field and telephone staff had to pass Alta Spanish language certification, a Spanish proficiency test.

Exhibit IV.2. The 2021–2022 study data collection: team structure



CATI = computer-assisted telephone interviewing; TCR = Teacher Child Reports

The main responsibilities of each role on the team are described below:

- The survey director and deputy survey director oversaw the work of all staff involved in recruitment, sampling, data collection, and data cleaning, meeting regularly with the task leads and supervisors described below, with senior project leadership, and with the project officers at OPRE. These staff were tasked with understanding how each part of the data collection process fit together and helping the task leads develop and implement creative solutions to challenges with recruitment, sampling, consent collection, and data collection. These staff monitored and reported on progress towards the study's consent collection and data collection goals, and worked to implement the enhanced consent collection strategies in spring 2022 (discussed further in Section G below).

- The recruitment and sampling task leads oversaw the work of the study liaisons to recruit programs and sample teachers and children. The recruitment and sampling task leads worked closely with FACES study team leadership and the lead sampling statistician and communicated with Regional Program Managers at the Office of Head Start for assistance in recruiting program. One lead oversaw recruitment, teacher sampling, and child sampling for the 60 programs in fall 2021. In spring 2022, that staff member continued to oversee teacher sampling for the 176 programs in the Program and Staff Study, while another staff member oversaw recruitment of those programs.

- Mathematica study liaisons served as the main point of contact between individual programs and the 2021-2022 Study. They worked with program directors to recruit programs and identify on-site coordinators, the person(s) the program director designated as the contact between the program and the 2021–2022 study team. Then, liaisons worked with the on-site coordinators to sample teachers and children, coaching on-site coordinators through the process of completing

the sampling spreadsheets and sharing them with Mathematica securely, and data entered information on teachers and children into the electronic sampling platform. After teachers and children had been selected, liaisons worked with on-site coordinators to coordinate the distribution and collection of parental consent forms. Study liaisons also followed up with on-site coordinators regarding outstanding staff surveys and (for programs participating in the Program, Staff, and Family Study only) Teacher Child Reports (TCRs). In spring 2022, liaisons contacted programs about the study's enhanced consent collection strategies and coordinated in-person consent follow-up visits where allowed by programs.

- The parent survey task lead oversaw programming and testing of the parent survey instrument, and monitored the progress of parent survey data collection in fall 2021 and spring 2022. They led the telephone interviewer trainings and worked with the CATI supervisors to resolve data collection issues. They worked with the Survey Operations Center to coordinate invitation, reminder, and incentive mailings for parents, and coordinated email invitations and reminders. They also performed data checks on incoming parent survey data and cleaned the parent survey data files after data collection ended (described in more detail in Chapter V).

- CATI supervisors recruited CATI interviewers and assisted in training CATI staff. They monitored telephone interviews conducted by the telephone interviewers and reported issues to the parent survey task lead.

- Telephone interviewers were trained to conduct computer-assisted telephone interviews (CATI) with parents in fall 2021 and spring 2022. In spring 2022, these staff also collected verbal consent from parents where allowed by programs.

- The teacher survey and TCR task lead oversaw programming and testing of the teacher survey and TCR instruments, and monitored data collection progress on those instruments. There were different task leads for fall 2021 data collection and spring 2022 data collection. They worked with the Survey Operations Center to coordinate mailings of invitation packets and incentives to teachers, and coordinated email invitations and reminders. They also performed data checks on incoming teacher survey and TCR data and cleaned the data files after data collection ended (described in more detail in Chapter V).

- The field supervisor recruited field enrollment specialists and assisted in their training. This supervisor made travel arrangements for site visits, reviewed data collection plans for site visits developed by the study liaisons, and maintained regular communication with field enrollment specialists during each visit, providing advice and reporting issues to the survey leadership team.

- Field enrollment specialists were trained to distribute and collect consent forms on-site at participating centers in spring 2022. These staff worked directly with Head Start staff, including the on-site coordinator, teachers, and other-center based staff and communicated directly with selected parents where allowed by programs. They showed parents how to complete the electronic consent form, collected paper consent forms from parents, and data entered paper consents into the electronic consent platform. They also followed up with teachers and center directors about outstanding spring 2022 instruments during their visits.

- The director survey task lead oversaw programming and testing of the center director and program director instruments, and monitored data collection progress on those instruments. This lead worked the Survey Operations Center to coordinate mailings of invitation packets and

incentives to teachers, and coordinated email invitations and reminders. They also performed data checks on incoming director survey data and cleaned the data files after data collection ended (described in more detail in Chapter V).

B. Liaison, CATI staff and Field Enrollment Specialist trainings

In fall 2021, Mathematica trained study liaisons on recruitment and sampling activities, and trained parent survey CATI staff on data collection activities. In spring 2022, the study team conducted a second parent survey CATI training and a second liaison training, as well as a Field Enrollment Specialist training. The second trainings in spring 2022 served as a refresher training for team members who had served in these roles in fall 2021 and as an introduction for new team members. This section discusses the training goals and procedures conducted for each data collection training.

1. Training goals

Due to the multifaceted nature of the 2021–2022 Study data collection, an important goal of the trainings was to ensure that trainees mastered a broad spectrum of tasks and successfully demonstrated the following:

- Fluency in the study goals, study design, and sample rationale so trainees could respond to questions from program staff and parents
- Mastery of the techniques required to carry out their assignments, such as sampling teachers and children, conducting interviews, or distributing and collecting consent forms
- Ability to use the specified computer systems for entering and transmitting data and doing administrative record keeping
- Mastery of data collection techniques, including contact and consent procedures, interviewing techniques, refusal avoidance and conversion, safety, record keeping, and maintaining privacy
- Ability to work smoothly and seamlessly with team members and the Head Start program's on-site coordinator— (field enrollment specialist training only)

Exhibit IV.3 shows the number of staff in each role in the fall and spring.

Exhibit IV.3. Number of staff in each data collection role, by data collection wave

Data collection wave	Field Enrollment Specialists	Liaisons	Parent CATI interviewers[a]
Fall 2021	n.a.	6	12
Spring 2022	9	10	15

[a]In fall 2021, eight of 12 CATI interviewers were bilingual. In spring 2022, five of 15 CATI interviewers were bilingual.

CATI = computer-assisted telephone interviewing; n.a. = not applicable.

2. Liaison training

a. Fall 2021 liaison training

In August 2021, six liaisons, all of whom were Mathematica staff, attended 10.5 hours of training over three days, focusing on an overview of the study, instructions in gaining rapport with program directors and other staff, paired practice with the recruitment materials, an overview of the computer-based sampling system, an overview of teacher and child sampling, an introduction to the Box secure file-sharing platform, practice in using the sampling system and Box, and best practices for sharing data securely with program staff. Project staff trained liaisons in establishing rapport with and gaining the cooperation of program directors and on-site coordinators by using recruitment scripts and refusal conversion techniques. Before the training, trainees received a manual via email that introduced them to the project and explained their role. During the training, project staff presented information in lecture format, followed by group discussion and hands-on practice to improve mastery.

To assess achievement of the training goals, trainers observed the trainees during paired practice sessions as they conducted a role-play exercise in which a trainer played the role of the center contact. The liaison spoke with the "center contact" and completed data entry during these role plays. After the training, each liaison participated in a certification session with a trainer who played the role of program staff. The certification session required the liaison trainee to simulate all standard liaison tasks—describing the 2021–2022 study and recruiting the program to participate, collecting and entering teacher and child data, selecting the classroom and child samples, and sharing information with the program using Box. All liaisons who received training passed the certification.

b. Spring 2022 liaison training

For spring 2022, Mathematica conducted a four-hour training for five new and five returning liaisons that incorporated paired practice. New liaisons also attended a one-hour pre-training session that focused on the goals of the study and gaining rapport with the program directors and staff. Certification procedures matched those used in fall 2021; all liaisons who were trained passed the certification.

3. Parent survey CATI staff training

Before each wave of data collection, the parent survey task lead and deputy survey directors trained Mathematica's Survey Operations Center (SOC) CATI staff on administering the parent survey. Before the training, all telephone interviewers participated in Mathematica's standard eight hours of general interviewer training, including training in how to conduct a CATI interview. In both fall and spring, interviewers received nine hours of training on the parent interview, which included an overview of the study and a question-by-question review of the instrument.

a. Fall 2021 CATI staff training

CATI staff training for the fall 2021 data collection took place in early October 2021. The training included 12 SOC staff interviewers, eight of whom were bilingual in both English and Spanish, plus three SOC telephone interview supervisors, who monitor the telephone interviewers and provide feedback and coaching to improve interviewer skills and ensure the quality of data collection. Two of the supervisors were bilingual.

Telephone interviewers took part in mock interviews monitored by the training staff. Interviewers demonstrated their ability to conduct the interview correctly by reading questions as worded, answering questions raised by respondents, recording responses accurately, and avoiding leading probes. During the paired practice, trainers and SOC supervisors listened to each interviewer read through a practice scenario, noting those who needed additional practice and individual follow-up before conducting interviews. We paired less experienced interviewers with experienced interviewers and monitors to ensure guided practice of key questions and immediate feedback.

b. Spring 2022 CATI staff training

We trained 15 interviewers to conduct the spring 2022 parent survey, including eight who conducted parent interviews as part of the 2021–2022 study in fall 2021. All 15 interviewers, along with three supervisors, received nine hours of training on the parent interview, which included an overview of the study and its goals, information on the study population, a question-by-question review of the parent instrument, and practice in pairs with mock scripts. Five of the 15 interviewers and two of the three supervisors were bilingual. As with the fall training, we paired experienced interviewers and supervisors with less experienced staff to allow for immediate and individualized feedback during the paired practice. SOC supervisors and trainers listened to each interviewer conduct a mock interview and noted those requiring additional training and practice before conducting interviews.

4. Field Enrollment Specialist training

In late April 2022, Mathematica conducted a two-day virtual training for the nine field enrollment specialists assigned to visit programs during the spring data collection. Six of the nine had worked on an earlier FACES study; the remaining three had worked in a similar capacity on other studies. Before the training, trainees received a manual via email that introduced them to the project and explained their role. The goal of the two-day training was to ensure that the field enrollment specialists mastered a broad range of skills related to their two main responsibilities: (1) distributing and collecting parent consent forms, and (2) following up on outstanding staff surveys. The deputy survey director led the training, with assistance from the survey director and sampling task lead. The training sessions covered (1) responding to questions from program staff and parents, (2) distributing and collecting parent consent forms, (3) conducting data entry and administrative record keeping, and (4) encouraging selected Head Start program staff to complete their surveys. The training also covered the background, purpose, and structure of the project and how to work with Head Start programs, staff, and parents. During the training, project staff presented information in lecture format, followed by group discussion and hands-on practice to improve mastery of consent form data entry. Trainers led group discussions on strategies for how best to distribute and collect the consent forms and outlined procedures for data transmission and administrative reporting.

To assess achievement of the training goals, after the training, each trainee participated in a remote certification test. The test had two components: (1) a knowledge assessment and (2) data entry of a mock consent form. The knowledge assessment consisted of 10 multiple choice questions; we required trainees to answer nine correctly. They were permitted up to one error on the data entry form. All field enrollment specialists who were trained passed the certification.

C. Recruitment procedures

As previously described, the study liaisons conducted the recruitment and sampling of programs selected for the 2021–2022 Study.[34] Liaison activities for the Program, Staff, and Family Study began in August 2021, with six liaisons assigned to work with the 60 programs. In December, we added an additional five liaisons to help with the recruitment and coordination effort for the additional 116 programs that agreed to participate in spring 2022 as part of the Program and Staff Study. All recruitment activities concluded in May 2022.

The liaisons' activities began with efforts to gain the cooperation of program directors. Liaisons worked with the on-site coordinator to collect lists of centers from which to select the sample of centers in each program. When centers were selected for the Program, Staff, and Family Study, liaisons worked with the on-site coordinators to sample teachers and children remotely. Similarly, in spring 2022, the liaisons conducted classroom sampling remotely for the programs in the Program and Staff Study. In fall 2021 and spring 2022, we offered the programs in the Program, Staff, and Family Study $500 in appreciation of their assistance with child sampling and consent gathering.[35] In spring 2022, we offered programs in the Program and Staff Study $250 for their assistance with remote sampling.

D. Center sampling

As the first step of the sampling process for returning programs in the Program, Staff, and Family Study, during the program recruitment process in fall 2021, liaisons confirmed whether the centers that had been selected to participate in FACES 2019 were still in operation and providing Head Start services. As described in Chapter II, once we recruited a returning program, we kept the originally sampled centers if they still planned to provide Head Start services at the time of program recruitment. If either of the two originally sampled centers was no longer in operation or providing Head Start services, we selected a new sample of two centers, with the center that participated previously having the same chance of selection as other eligible centers in the program. For all newly recruited programs, liaisons collected the full list of centers providing Head Start services as part of the program recruitment process, from which we selected a new sample of two centers. In total, 24 of the 60 Program, Staff, and Family Study programs required center sampling.

Similarly, for returning programs in the Program and Staff Study, during the program recruitment process in spring 2022, liaisons confirmed whether the centers that had been selected to participate in FACES 2019 were still in operation and providing Head Start services, and drew replacement programs as needed following the same procedures as for the Program, Staff, and Family Study. For all newly recruited programs, liaisons collected the full list of centers providing Head Start services as part of the program recruitment process, from which we selected a new sample of two centers.

[34] For spring 2022, a four-hour training was held for new and returning liaisons that incorporated paired practice. New liaisons also attended a one-hour pre-training session focusing on the goals of the study and gaining rapport with the program directors and staff.

[35] We originally planned to collect consents only during fall 2021, and to offer the increased honorarium payment in fall 2021 only. Because low consent rates in fall 2021 prompted the continued collection of consents through the spring 2022 data collection, we increased the spring 2022 honorarium payment for programs in the Program, Staff, and Family Study to match their fall payment.

E. Teacher sampling

In fall 2021, after confirming the existing center sample or selecting a new one as necessary, liaisons began the processes for remotely sampling teachers for the 60 programs in the Program, Staff, and Family Study. For each program, teacher sampling began with the liaisons contacting each on-site coordinator to explain the sampling process and the information needed. The liaisons then created custom teacher sampling spreadsheets for each program, which they uploaded to Box or emailed to the on-site coordinator in a password-protected, encrypted form. Liaisons provided guidance to on-site coordinators on how to create a Box account and how to download and upload files; liaisons used written instructions, telephone calls, and video conference meetings with screensharing to guide on-site coordinators through these processes.

The teacher sampling spreadsheets requested each lead teacher's or home visitor's name, email address, mode(s) of instruction (including in-person only, virtual/remote only, hybrid only, home visit, or other), and the number of children taught by the teacher (or served by the home visitor). Once they received this information, liaisons used a web-based sampling program to enter these data for each center. From this teacher list, the sampling program selected two teachers or home visitors for participation in the study. Procedures in spring 2022 for teacher sampling for the Program and Staff Study mirrored those used in fall 2021 for the Program, Staff, and Family Study.

F. Child sampling

After selecting the participating teachers in the 60 programs participating in the Program, Staff, and Family Study, the liaisons informed on-site coordinators of the selected teachers and requested information on all children instructed by each teacher. The requested information included each child's first and last name or other identifier, date of birth, date of first enrollment in preschool Head Start, whether the child had participated in Early Head Start, funding source(s), instruction type, sex, home language, and a sibling indicator. On-site coordinators provided the requested information via a custom child sampling spreadsheet for each program. Like the teacher sampling spreadsheets, these spreadsheets were either uploaded to Box or emailed to the on-site coordinator in a password-protected, encrypted form.

Once liaisons received this child information, they entered it into the web-based sampling program, which selected 12 children from each participating teacher's roster using a programmed sampling algorithm developed by senior statisticians at Mathematica. Because all children taught by a teacher were eligible for participation, children from different sections taught by the same teacher may have been selected. After identifying siblings within the center's selected child sample, the sampling program then randomly selected one child per family. (We define "family" and "sibling" as more than one child with the same primary caregiver; that is, parent or guardian.) We retained these samples of teachers and children for the spring 2022 wave of data collection.

In December 2022, due to low rates of program enrollment and parental consent, we released all eligible and previously nonselected children taught by participating Head Start teachers ("backup children") into the sample. (Chapter II has a more detailed discussion of these backup children.)

G. Parental consent procedures

After completion of the classroom and child sampling, we assembled parent consent packets—in English or Spanish, according to the home language the on-site coordinator provided on the child sampling

spreadsheet—for each child in the sample. The packets included a brightly colored brochure with a brief overview of the 2021–2022 Study, an introductory letter, and a consent form. The letter included a link to an online consent form. Each consent packet included a label with a customized username and password that parents could use to log in to the online consent form. We shipped on-site coordinators a bundle of consent forms, organized by teacher. The liaison then coordinated with the on-site coordinator to develop a distribution plan for the consent packets, which was to incorporate communication about the study into programs' typical parent communication methods, including speaking directly with parents at pick-up and drop-off or parent-teacher conferences; sending information home with students in their folders or backpacks; and sending reminders about the study via email, text message, social media, and virtual learning platforms. In a few programs, liaisons held virtual information sessions for teachers. The on-site coordinators collected the completed consent forms and sent them to Mathematica using pre-paid FedEx or USPS Priority Mail envelopes, or scanned and uploaded the forms to Box.

We distributed the consent forms for backup children in December 2021 and January 2022, in most cases immediately before or immediately following programs' winter holiday closures.

In spring 2022, if permitted by the program, we updated the consent process to improve reaching parents by email or telephone. We added questions to the parent consent form to accommodate verbal consent and assess parents' preferred mode for survey completion. Using verbal consent aimed to expedite parents' participation in the consent process by eliminating distribution and mail collection of the hard-copy consent form. Additionally, when programs met two requirements – (1) an overall parental consent rate below 75 percent and (2) 10 or more outstanding parental consent forms – and agreed to allow us to visit their centers, we used an approach successful in earlier FACES studies—we sent field staff to Head Start programs to distribute consent forms and provide parents with information about the study in person.

H. Planning and collecting data

Following the recruitment and sampling activities, the 2021–2022 study team engaged in extensive planning and preparation for the field teams' in-person data collection visits. This section describes the planning activities conducted before and during data collection in fall 2021 and spring 2022, including scheduling and field staff preparation. We also describe the data collection procedures for each wave.

1. Planning for fall 2021 remote data collection

Fall 2021 data collection began on a rolling basis as parents provided consent. We initially planned for data collection to begin in mid-October 2021 and continue until the end of December. However, because of the low consent rates described in Chapter II, we continued with data collection through the end of January 2022.

2. Fall 2021 data collection

Fall 2021 data collection included remotely collecting parent surveys, teacher surveys, and TCRs from the 60 programs participating in the Program, Staff, and Family Study between early October 2021 and late January 2022. Data collection began on a rolling basis as programs completed center, teacher, and child sampling, and program staff distributed consent forms to the parents of selected children.

a. Teacher Child Reports and teacher survey

To streamline the process for individual teachers, we released their TCRs and fall teacher surveys between October and December 2021, once 70 percent of a teacher's children had reached consented status. Due to the aforementioned low consent rates, we decided to change this process and release all teacher surveys and TCRs for consented children in early January 2022, giving the majority of teachers in the sample approximately three weeks to complete their fall instruments before the end of the fall 2021 data collection.

Teachers received invitations by email if an email address was provided during the teacher sampling stage. Additionally, we mailed hard-copy packets to each teacher's center. All teachers received a hard copy packet containing an introductory letter, web instructions, paper TCRs, a paper survey, a paper consent form, and postage-paid envelopes in which to return materials. The introductory packet also included the TCR instrument's web address and the teacher's log-in ID and password as an option for teachers wishing to complete their TCRs online, which most teachers chose to do. We did not offer a token of appreciation for completing the teacher survey, but teachers received $10 per completed TCR form when allowed by their Head Start programs. We sent weekly email reminders to nonresponders throughout the data collection period and added weekly telephone reminders in mid-January 2022 as the data collection period neared its end.

Of the completed teacher surveys (n = 193), 98 percent were completed online and 2 percent were completed on paper. Of the completed TCRs (n = 887), 93 percent were completed online and 7 percent were completed on paper. Exhibits IV.4 and IV.5 give a full description of completion and response rates.

b. Parent survey

We included a web-based parent survey option in addition to the telephone mode. Throughout the data collection field period, telephone interviewers in Mathematica's SOC scheduled weekend and evening interview appointments to accommodate parents' schedules.

Once the parent's consent was logged in the system (either directly by the parent or through SOC data entry for paper consents), they immediately received an email and, shortly afterward, a mailed letter inviting them to complete the survey. The invitations contained an internet web address, log-in ID, password, and web instructions for completing the survey online, as well as a toll-free telephone number for those wanting to complete it by phone. Parents who completed the consent form electronically could immediately continue to the web parent survey. After logging electronic receipt of parental consent, parents who did not provide an email address on their consent form received a call to remind them to complete the survey. Those who provided an email address received a similar reminder call within two days. All parents received follow-up reminder materials by email, mail, and phone, which described both online and telephone options for completing the survey. Reminder letters and emails were sent weekly starting approximately three weeks after parents completed their consent forms; outreach by phone continued every one to three days after the initial call. The parent survey was offered in English and Spanish.

We offered parents a token of appreciation ($30 gift card) for completing the survey. Of the completed parent surveys (n = 785), 48 percent were completed on the web and 52 percent by telephone. (Exhibit IV.4 shows completion and response rates.)

3. Planning for spring 2022 site visits and remote data collection

Before the spring 2022 data collection, the study team revised the spring 2022 instruments and made other procedural changes to highlight topics aligned with our measurement priorities[36] while keeping the instruments a reasonable burden. Additionally, we extended the consent collection for children selected in the fall. The study team wanted to encourage parents' and staff members' participation and raise awareness about the study among sample members. Additionally, we planned classroom observations across all programs for spring 2022 but had to cancel them in January 2022 due to concerns around the evolving pandemic.

a. Updates to consent forms and consent collection procedures

The study team made the following changes to consent forms and consent collection procedures to minimize burden on programs and facilitate consent distribution and collection:

- **Continued receipt of consents.** At the end of the fall data collection, we asked programs to continue sending us any completed consent forms they collected, explaining that we would accept consents beyond January 31, 2022, but without a specific end date.
- **Updates to consent forms.** We removed from the consent forms all references to fall activities and the canceled classroom observations.

In addition to the direct consent outreach and field enrollment specialist site visits described above in Section G, we shipped updated consent forms and flyers to programs beginning the week of April 4, 2022 and ending the week of June 6, 2022. In total, 44 programs received additional hard-copy consent forms by mail in spring 2022. We receipted hard-copy consent forms and scanned and uploaded them to Box or mailed them to Mathematica's SOC until data collection concluded on July 31, 2022.

b. Additional materials to raise awareness about the study

We created new flyers and additional reminder materials for flexible use (such as via email, social media, and bulletin board posts) by program staff to remind center directors, teachers, and parents about the study and encourage participation.

c. Increased honorarium for on-site coordinator

Because we extended the consent process past the fall wave, we increased the amount of the spring on-site coordinator honorarium from $250 to $500 to match that offered in the fall.

d. Roster updates

For the 60 programs participating in the Program, Staff, and Family Study, preparation for spring 2022 began with a roster update to determine which children selected in the fall were still eligible for the spring

[36] In early 2022, the Office of Management and Budget approached the study team about the possibility of revising some spring 2022 instruments to better address the role of Head Start as an employer, and to serve as a baseline to understand changes that might occur with potential new policy initiatives. To align with this request, the study team revised the spring 2022 instruments to gather additional information on topics already included in the surveys. Additionally, the study team made further revisions to the instruments to facilitate linking teachers and children to specific classrooms and to reduce the burden for new parent survey respondents in the spring who did not complete the fall 2021 survey (accounting for low consent and response rates in the fall).

data collection. The study team planned to survey all eligible parents and collect TCRs for all eligible children (those still enrolled in Head Start at the time of data collection). Children who left Head Start or moved to a Head Start center in a program not sampled for the 2021–2022 study were not eligible for the spring 2022 data collection. To establish children's eligibility for spring 2022, study liaisons recontacted the on-site coordinators in March 2022 and requested information on the sampled children to determine those who were still in the program and those who had left. Study liaisons also asked on-site coordinators to confirm the parent contact information contained in our SMS or provide updated information.

4. Spring 2022 data collection

Spring 2022 child-level data collection, including parent surveys and TCRs, continued for the 60 Program, Staff, and Family Study programs that participated in the fall. In support of the Program and Staff Study, we conducted teacher, program director, and center director surveys for all 176 programs.

The spring 2022 data collection components consisted of a parent survey administered by telephone or via the web; a TCR and teacher survey, both administered via the web with an option to complete on paper; and surveys of center and program directors, both administered via the web. For all instruments, data collection began in late April 2022 and concluded at the end of July 2022.

a. Parent survey

Spring procedures mirrored those used in the fall. Of the completed parent surveys (n = 928) 40 percent were completed via the web and 60 percent by telephone. (Completion and response rates are in Exhibit IV.5.)

b. TCR and teacher survey

Procedures for the spring teacher survey and TCR largely mirrored procedures used in the fall for teachers in the 60 programs participating in the Program, Staff, and Family Study. However, teachers in these programs were able to access their TCRs for consented children at the beginning of the field data collection period, regardless of how many consented children were enrolled in their class(es). Additionally, we added phone reminders for outstanding TCRs and teacher surveys; weekly phone reminders began in late June, with calls being made to all centers with outstanding teacher surveys and/or TCRs.

Procedures differed somewhat for teachers in the 116 Program and Staff Study programs. These teachers received a hard-copy study packet containing an informational letter with the teacher survey web address; log-in ID and password; web instructions; paper teacher surveys; paper consent forms; and return materials. Packets were mailed to the individual teachers using their center's address. We did not offer these teachers a token of appreciation for completing the teacher survey. Reminder procedures for these teachers mirrored those for teachers in the Program, Staff, and Family Study.

Of the completed teacher surveys (n = 363), 92 percent were completed online and 8 percent on paper. Of the completed TCRs (n = 1,251), 91 percent were completed online, and 9 percent were completed on paper. (Exhibit IV.5 gives a full description of completion and response rates.)

c. Program director survey

We offered the program director survey only as a web survey. We sent program directors a paper survey invitation by mail, containing the survey web address, a log-in ID and password, and web instructions.

They also received an email invitation requesting completion of the survey, along with weekly email reminders. In addition to these weekly email reminders, throughout the data collection period, study liaisons also contacted on-site coordinators on an ad hoc basis to ask them to follow up with program directors who had not yet completed their surveys. (Completion and response rates are in Exhibit IV.5.)

d. Center director survey

Like the program director survey, we offered the center director survey only as a web survey. We sent center directors a paper survey invitation along with the survey web address, log-in ID and password, and web instructions. They also received an email invitation requesting completion of the survey, along with reminders by email and phone. Email reminders were sent weekly beginning one week after the survey invitations were sent; weekly phone reminders began in late June, with calls being made to all centers with outstanding center director surveys. (Completion and response rates are in Exhibit IV.5.)

5. Monitoring consent and data collection in fall 2021 and spring 2022

In fall 2021 and spring 2022, liaisons received email notifications when programs uploaded scanned consent forms to Box. Following notification, liaisons would download the scanned consents and save them to Mathematica's secure network drive. Two to three times a week, the deputy survey director monitored the network drive location for new scanned consents and alerted the data entry team at the SOC of those ready for receipting. Hard-copy consent forms mailed to Mathematica by programs and families were received at the SOC and processed immediately upon receipt. Throughout fall and spring data collection, the recruitment and sampling task lead, deputy survey director, and survey director monitored processed consents daily via refreshable status reports.

Instrument leads and telephone supervisors (Exhibit IV.2) oversaw the progress of data collection. In fall 2021 and spring 2022, two telephone supervisors at Mathematica's SOC used daily status reports to monitor the work of the telephone staff who completed the parent surveys. Using daily status reports and weekly meetings with field supervisors, the FACES task leaders monitored completion rates of parent surveys, teacher surveys, and TCRs in fall 2021 and spring 2022, and staff surveys in spring 2022.

I. Quality assurance of parent telephone interviews in fall 2021 and spring 2022

During the fall 2021 and spring 2022 parent telephone survey data collections, SOC supervisors monitored interviewers' live calls to evaluate their performance and reviewed their notes in the CATI system for clarity. These processes ensured the professional conduct of all parent telephone interviews and the CATI staff's adherence to the standardized interview protocol. During the first week of telephone interviewing in fall 2021, and again in spring 2022, study staff and professional SOC monitors listened to the entire interviews (from dialing until completion) for 100 percent of the cases worked during this week. Monitors provided immediate feedback to the interviewers after each call. After the first week, the study team monitored 10 percent of all cases.

J. Response rates

We reached a parent consent rate of 44 percent, which resulted in a sample of 1,363 children for inclusion in the study in fall 2021 out of 3,105 sampled and study-eligible children (Exhibit IV.4). We collected completed surveys from 58 percent of participating families in fall 2021. We reached a 65 percent completion rate of teachers' ratings of these 1,363 eligible and consented children in fall 2021. As a result of our continued consent efforts and changes in children's eligibility, the spring 2022 sample included

1,837 children (Exhibit IV.5). We collected surveys from 51 percent of their parents in spring 2022. We reached a 68 percent completion rate for teachers' ratings of these spring-eligible children in spring 2022.

Exhibit IV.4. The 2021–2022 study fall 2021 target versus actual sample sizes and instrument completion and response rates

	Target		Actual	
Number of eligible and consented children	**2,400**		**1,363**	
	Number	**Percentage**	**Number**	**Percentage**
Completed parent surveys	2,040	85%	785	58%
Parent response rate (weighted)[b]	15.6			
Completed TCRs	2,040	85%	887	65%
TCR response rate (weighted)	17.7			
Number of selected teachers	**240**		**239**	
	Number	**Percentage**	**Number**	**Percentage**
Completed teacher surveys	204	85%	193	81%
Teacher survey response rate (weighted)	51.7			

Exhibit IV.5. The 2021–2022 study spring 2022 target versus actual sample sizes and instrument completion and response rates

	Target		Actual	
Number of eligible and consented children	**2,160**		**1,837**	
	Number	**Percentage**	**Number**	**Percentage**
Completed parent surveys	1,836	85%	928	51%
Parent response rate (weighted)[b]	23.8			
Completed TCRs	1,836	85%	1,250	68%
TCR response rate (weighted)	32.0			
Number of selected and eligible teachers	**720**		**631**	
	Number	**Percentage**	**Number**	**Percentage**
Completed teacher surveys	612	85%	358	57%
Teacher survey response rate (weighted)	40.5			
Number of recruited programs	**180**		**176**	
	Number	**Percentage**	**Number**	**Percentage**
Completed program director surveys	153	85%	132	75%
Program director response rate (weighted)	53.4			
Number of selected and eligible centers	**360**		**344**[37]	
	Number	**Percentage**	**Number**	**Percentage**
Completed center director surveys	306	85%	237	70%
Center director response rate (weighted)	45.5			

[37] Four of these centers declined to participate in the 2021-2022 study.

[b]The weighted cumulative response rate represents the percentage of the eligible population that, according to our estimate, we have captured with this instrument in this wave. It incorporates earlier stages of selection and participation (program, center, classroom); parental consent; estimated eligibility status for those children with unknown status; and child selection probabilities. It also uses a more stringent definition of eligibility, treating as eligible those children who are part of the Head Start population but who are out of scope according to the study's operational protocol. For purposes of this response rate, all children in the sample were assumed to have a known eligibility status in fall 2021, even if we did not obtain consent or response. In spring 2022, we continued the consent process and tracked who left the program between fall and spring, even if they had not consented in the fall.

n.a. = not applicable; TCR = Teacher Child Report.

Chapter VI includes a nonresponse bias analysis at the child and staff levels, and assesses whether the weighting adjustments appear to have mitigated the risk for nonresponse bias. The study team concluded that analysis weights adjust for the differential probabilities of selection and sample attrition, and can potentially help reduce the bias that may result from differential nonresponse.

Spring 2022 response rates for Head Start teachers, center directors, and program directors are shown in Exhibit IV.6. The study team completed teacher surveys for 57 percent of the sampled teachers, center director surveys for 70 percent of the centers, and program director surveys for 75 percent of the programs participating in the study.

Exhibit IV.6. The 2021–2022 study staff survey completion and response rates in spring 2022, overall and by programs with and without child-level data collection

	In all programs	In programs with child-level data	In programs without child-level data
Number of teachers who completed Head Start teacher surveys in spring 2022	358	173	185
Teacher spring survey completion rate[a]	56.7	73.6	46.7
Teacher spring survey response rate (weighted)[b]	40.5	47.6	35.9
Number of completed center director surveys	237	85	152
Center director survey completion rate[a]	69.7	75.2	67.0
Center director response rate (weighted)[b]	45.5	46.1	45.4
Number of completed program director surveys	132	42	90
Program director survey completion rate among participating programs[a]	75.0	70.0	77.6
Program director survey response rate (weighted)[b]	53.4	48.7	56.2

[a]The completion rate is the count of completes divided by the count of those who participated in that wave of data collection.

[b]Weighted cumulative response rate conveys the percentage of the eligible population that is represented by the respondents to a particular instrument. It incorporates earlier rounds of selection and participation when applicable (program and/or center), as well as selection probabilities. For the purposes of this response rate, all staff were assumed to have known eligibility status.

n.a.= not applicable.

V. Data preparation

This chapter describes the activities the study team used to collect and process data from electronic and paper surveys. The goal was to improve data quality during data collection and processing by limiting programming errors, data entry errors, and other correctable mistakes. Chapter VI describes how the study team modified the data to protect respondent confidentiality.

A. Electronic and paper documents

The primary mode of data collection for the 2021–2022 study was electronic, which greatly reduced data processing needs, as there was little need for data entry. The project's computer-assisted telephone interview (CATI) and web-based surveys were developed in Confirmit, a professional multimode platform for survey design and data collection. The electronic surveys contained built-in range checks, logic checks, and routing instructions that eliminated some types of respondent errors found with paper instruments, such as incorrectly skipping or responding to questions, or entering out-of-range responses. The electronic surveys' routing also allowed returning respondents to skip demographic questions they answered in the previous wave of the survey, which reduced respondent burden. Before fielding the instruments, the study team comprehensively checked all survey skip logic, validations, and question properties with a random data generator testing tool included in Confirmit that used about 500 test cases per instrument. The random data generator automatically produced test data sets of randomly generated instrument responses, simulating real surveys and interviews.

In fall 2021 and spring 2022, parents had the option to complete their survey on the web or via CATI. Upon completing the electronic study consent form or returning a paper consent form to the study team, parents received an email invitation to complete the survey on the web. Within days of receiving the email invitation, parents began receiving calls from Mathematica's SOC to complete the survey over the phone.

In fall 2021 and spring 2022, teachers had the option to complete their instruments via the web or paper, though the vast majority of teachers completed their instruments on the web (see Chapter IV for details on response rates and completion mode). The study team sent paper TCRs and teacher surveys, along with web links, to teachers' centers so teachers could pick up and return them directly to the SOC in prepaid envelopes. Throughout the field period, teachers entered most of their 2021–2022 study data electronically via the web-based instruments.[1] As discussed below (Section C), SOC staff entered data from paper instruments into the web-based instruments to produce electronic data throughout the field period, giving project staff the opportunity to review all incoming data from both web or paper surveys, and quickly identify emerging problems.

In spring 2022, the study team sent center and program directors a web survey link; no paper option was offered. Data processing began as each instrument was completed in Confirmit.

[1] Ninety-one percent of TCRs were completed via web and 9 percent by paper. Ninety-two percent of teacher surveys were completed via web and 8 percent via paper. Response rates by mode and other administration characteristics are discussed further in Chapter IV.

B. Frequency review

Throughout the field period, project staff monitored descriptive statistics for all instruments to identify emerging problems. Senior project staff regularly reviewed the electronic data to ensure no corrections to instruments were needed during or after data collection. Staff monitored survey data for respondent and interviewer errors, potentially indicated by high levels of item nonresponse ("don't know" and "refused"), nonresponse for an entire survey section, and survey completion time across all respondents to see whether the surveys took the estimated amount of time to complete.

In fall 2021, we performed frequency review after 25 teacher surveys and 200 TCRs had been completed; for all other instruments, after about 100 surveys had been completed, project staff members reviewed individual and aggregate data frequencies and descriptive statistics. At this point, no paper surveys had been received, so the frequency review included only web survey data.

We performed random data generator testing on the incoming respondent data for the program director survey, center director survey, and spring 2022 teacher survey. This was more efficient than performing manual data checks, as the larger sample of programs in spring 2022 (the 176 programs in the Program and Staff Study rather than just the 60 programs in the Program, Staff, and Family study), along with the increased parental consent rates resulted in these instruments having much larger numbers of completes than those from the fall (see Chapter IV for information on response rates). This testing showed that all instruments worked as expected, but also identified a few potential respondent errors. For example, some respondents entered the ***number*** of students who spoke a certain language when they were asked for the ***percentage*** of students. Instrument leads alerted the data processing teams of such issues, who corrected issues where possible based on responses to other items –for example, dividing the number of students who spoke a certain language by the number of students taught by the teacher in that class section to resolve the issue in the example in the previous sentence – or otherwise marked affected data as missing.

C. Data entry

SOC data entry staff entered the data from the paper TCRs and teacher surveys into the corresponding web-based instruments. This approach ensured that the data from paper TCRs and surveys underwent the same range, logic, and consistency checks built into the web-based instruments.

Project staff for the 2021–2022 study trained the data entry staff according to the specifications created for each instrument. For example, they trained staff on what to do when respondents provided several answers to a question that allowed for only one answer or answered questions they should have skipped based on their response to an earlier question. Data entry and project staff documented and resolved any identified issues.

When paper instruments arrived at the SOC, the data entry staff checked that an answer to each question was recorded and looked for any errors in administration (for example, respondents skipping a question they should have answered). The staff reviewed each paper survey for completeness, clarity, and adherence to routing and range rules, and informed SOC supervisors of errors. SOC supervisors in turn informed instrument leads of any errors, and the latter followed up with respondents via email and telephone to correct these errors as needed. In fall 2021, we did not need to follow up to correct any errors on the teacher survey or TCRs. In spring 2022, we followed up with seven respondents via email about potentially erroneous responses. One respondent provided updated information. The data processing team flagged and resolved other potential errors.

For final file production, SAS programs identified any deviation from final survey rules on skip patterns, allowable values, and item consistency, with case-based SAS code providing documentation for the editing process. As a final quality control check, a systems analyst reviewed all programming aspects of the editing process, including the SAS code.

D. Coding other-specify and open-ended responses

Some questions in the staff and parent surveys were open ended, allowing respondents to phrase their own answers and enter them directly, or requiring interviewers to record the responses to those items verbatim. In addition, on many closed-ended items, respondents could choose an open-ended response option ("other") and then write in their response if it did not fit into the provided response categories.

Whereas the staff surveys did not have a phone interview option, respondents to the parent survey could opt to complete it with an interviewer over the phone. Probes and help screens were built into the parent survey for the respondents and interviewers. Project staff trained the telephone interviewers on using appropriate probing techniques to ensure respondents provided a response that could be coded into existing categories. If a response truly did not fit into the given options, staff trained the interviewers to probe until they had enough information to record survey responses verbatim directly into the instrument.

1. Codebooks for other-specify items

All instruments in the 2021-2022 study contained other-specify items. The instrument leads created codebooks for each instrument and identified response categories for each of those items. The codebooks also included instructions and rules about which responses to back-code (by selecting the appropriate existing or newly defined response category) and which to leave as "other." The instrument design leads reviewed the codebooks to ensure all appropriate items were included, and that all instructions to coders were accurate before being finalized.

2. Back-coding and creation of new variables

Instrument leads provided specially trained SOC coders with an Excel file containing verbatim responses. The coders created an Ascribe coding database and then uploaded the data into it. The Ascribe program enabled coders to view the question number, question text, verbatim response, and respondent's project-specific identification number.

Coders received instructions about each instrument included in the 2021–2022 study and copies of both the questionnaire specifications and the codebooks created by the instrument leads as described above in Section D1. They also reviewed the verbatim responses to related questions to inform their back-coding decisions. If they found enough verbatim responses of a single type that could be coded to a new category (typically a similar response from about 10 percent of respondents), they would propose a new code to senior project staff for review.

After SOC coders finished back-coding by changing responses to the appropriate existing or newly defined code in the Ascribe coding database, instrument leads and other project staff manually reviewed all back-coded data to ensure accuracy and made updates where needed to improve codes. For example, the program director survey instrument lead reviewed back-codes for top reasons why lead teachers left a Head Start center. Several respondents entered "retirement" in the "other (specify)" text box, so SOC coders added this response option to the database. Senior project staff also reviewed verbatim responses to identify less common responses of possible substantive interest that would therefore warrant a new

code. Once senior project staff approved the proposed codes, coders added them to the database, and instrument leads added the new ones to the codebook. For items included in more than one survey, new codes created in one instrument were added to the others. As part of this process, any new categories or codes in the codebook that were not in the actual instruments were added to the codebook.

3. Open-ended items

Before making the files available to the public for restricted use access, study staff reviewed the verbatim text for all open-ended questions in the 2021–2022 study instruments and removed any personally identifiable information. Project staff stripped unique identifiers, such as names of people, Head Start programs, and Head Start centers; geographic markers, such as the names of cities, states, and regions; and other potentially identifiable data, such as the exact number of staff employed by a Head Start program and the names or acronyms of various state agencies. Each response was reviewed three times, first by the data cleaning lead, then by the deputy project director, and finally by a member of the study's senior leadership team. The study team inserted a description in brackets such as "[child's name]" or "[name of Head Start center]" to replace redacted text.

VI. Data file content, structure, and use

Mathematica has prepared a set of data files to use for conducting secondary analyses of data from the 2021–2022 study. In this chapter, we (1) describe the basic structure of the files and the conventions used to name the variables, and (2) offer guidance on how to use the data, including special instructions for SAS programmers using data files from the Child and Family Data Archive through the Interuniversity Consortium for Political and Social Research (ICPSR). We also describe the analysis weights that users should use. These weights adjust for the differential probabilities of selection and sample attrition, and can potentially help reduce the bias that may result from differential nonresponse. We include several weights with the data and offer some advice on which ones to use for various analyses. We describe when it may be appropriate to conduct within-round or cross-study analyses (that is, comparing the 2021–2022 study with one or more FACES studies) and approaches that data users may employ.[39] We conclude with a description of the variance estimation procedures appropriate for a complex sample design such as the one used in the 2021–2022 study.

A. Data files and data file structure

We organized the 2021–2022 study data collected in fall 2021 and spring 2022 into three data files (Exhibit VI.1):

1. **Spring 2022 center- and program-level file.** Contains identifiers to link the data to the other data files; program and center analysis weights; composite variables created from the program and center director surveys; and program and center director survey item-level data.
2. **Fall 2021–spring 2022 classroom- and teacher-level file.** Contains identifiers to link the data to the other data files, classroom- and teacher-level analysis weights, all classroom-/teacher-level composite variables, and teacher survey item-level data.
3. **Fall 2021–spring 2022 child-level file.** Contains identifiers to link the data to the other data files,[40] child-level analysis weights, indirect child assessment scores (for example, teacher-reported problem behavior scores), and composite variables developed using data from the 2021–2022 study staff instruments. The file also includes item-level data for all noncopyrighted items from the TCR and parent survey. The child-level file does not include item-level data available in the classroom-/teacher-level and center-/program-level files.

[39] Cross-*study* analyses may compare cross-sectional estimates from the 2021–2022 study with one or more FACES studies (for example, FACES 2014 or FACES 2006 through the 2021–2022 study) to examine changes in the characteristics of the Head Start population over two or more studies or trends in these characteristics over longer periods.

[40] The fall 2021–spring 2022 child-level file contains data only from the 60 programs in the Program, Staff, and Family Study. The spring 2022 center-/program-level and spring 2022 classroom-/teacher-level files contain data from all 176 Program and Staff Study programs. Therefore, when linking data from the child-level file to other data files, records will match only with those data from the 60 programs.

Exhibit VI.1. Fall 2021–spring 2022 data file descriptions

File and description	Survey instruments and collection date	Data file name	Number of variables/ records	Sort order of records (Sort ID)
Spring 2022 center- and program-level (including spring 2022 program director survey; spring 2022 center director survey; and spring 2022 center and program contextual data)	Program director survey (S22) and center director survey (S22)	ctrprog_21_22HSfaces_study	1010/340	D2_ID, C2_ID
Fall 2021–spring 2022 classroom- and teacher-level (including teacher survey and teacher survey contextual data)	Teacher survey (F21, S22)	classtchr_21_22HSfaces_study	657/651	CLS_ID
Fall 2021–spring 2022 child-level (including fall 2021 and spring 2022 Teacher Child Report [TCR], fall 2021 and spring 2022 parent survey data, fall 2021 and spring 2022 teacher survey contextual data, and spring 2022 center and program contextual data)	TCRs (F21, S22), parent surveys (F21, S22), teacher surveys (F21, S22), center director survey (S22), and program director survey (S22)	child_faces21_22HSfaces_study	1144/1945	ChildID

Note: F21 = fall 2021; S22 = spring 2022.

1. Organization of variables in data files

The structure of each of the 2021–2022 study files is similar. They generally contain ID variables; a set of flags indicating whether the case completed each instrument used as a data source (for example, completed parent survey in fall 2021); weight variables; composite variables or scores; item-level data from appropriate surveys or interviews; and the sampling variables[41] required to calculate standard errors, where the included sampling variables depend on the level of analysis.

Some instruments (such as the parent survey) were administered only to a subset of 60 programs participating in child-level data collection as part of the Program, Staff, and Family Study. To identify these programs, the center-/program-level and classroom-/teacher-level data files contain a cohort variable indicating whether a program completed child-level data collection or not. Below we describe the organization of data for these files.

a. Spring 2022 center-/program-level file

The file includes data from the spring 2022 program director and center director surveys. The data include one record for each of the 340 centers that contained any classrooms eligible for the 2021–2022 study sample in fall 2021 or spring 2022.[42] The data on the center-/program-level file are organized into a set of

[41] Sampling variables include those identifying strata and clusters (primary sampling units, or PSUs).
[42] Because there are typically two centers (and center director surveys) for each program, the program director data are replicated on each center record. However, some programs had only one center.

program and center identifiers. These identifiers are followed by **Cohort**, a flag indicating whether the program and center are part of the subset of 40 programs with child-level data collection that also participated in FACES 2019 (Cohort = 1), the 20 programs with child-level data that did not participate in FACES 2019 (Cohort = 4), the 82 programs without child-level data collection that also participated in FACES 2019 (Cohort = 2), or the 34 programs without child-level data collection that did not participate in FACES 2019 (Cohort = 5). Users wishing to conduct cross-study analyses between the 2021–2022 study and FACES 2019 should filter data to include only Cohorts 1 and 2. The following Cohorts are two larger blocks of variables (Exhibit VI.2).[43] The first block of variables includes spring 2022 weighting variables that support program- and center-level analyses of the 2021–2022 study data and composite variables from the program and center director surveys (for example, **D2MNTRCT**, which is a count of the mentors in the program, and **C2TCHTRN**, which is the lead teacher turnover in a center). The second block of variables contains data on items at the level of the program director and center director surveys (individual responses to each of the items on these surveys).

Exhibit VI.2. Spring 2022 center-/program-level data file structure

Identifiers
Unique identification numbers for program and center
Cohort indicator identifying programs with child-level data
Composite variables (spring 2022)
Data flags for each survey instrument (yes/no)
Analysis weights and clustering and stratification variables
Composite variables for the program director survey
Composite variables for the center director survey
Survey item-level data (spring 2022)
Items and responses for the program director survey
Items and responses for the center director survey

b. Fall 2021 and Spring 2022 merged classroom-/teacher-level file

The file includes data from the fall 2021 and spring 2022 teacher surveys. The file contains one record for each of the 631 teachers sampled and eligible for the data collection, including 236 teachers sampled in the 60 programs with child-level data (all but one of whom were still eligible in the spring) and 396 teachers sampled in the additional 116 programs with no child-level data (spring-only programs). The file also contains an additional record for each teacher who had two classrooms—six teachers in the 60 programs with child-level data and 26 in the additional 116 programs with no child-level data (spring-only programs). In addition, it contains five records for nonsampled teachers who began teaching selected children after those children were chosen for the sample and who completed a spring teacher survey. The original sampled teachers for these children either left the sampled center or were reassigned from teaching the sampled child; in spring 2022, teachers who no longer taught any sampled children but had been originally selected to participate in the study were invited to complete a spring 2022 teacher survey. Surveys collected from these five teachers do not have positive teacher weights and may be used only for

[43] Cohort 3 represents programs participating in a separate study conducted with Head Start programs in Region XI, discussed in a separate User's Manual: the 2021–2022 Study of Family and Staff Experiences in Head Start AIAN FACES Programs (2021–2022 Study) User's Manual.

child-level estimates. Weights for performing teacher-level analyses are set to adjust for the 32 teachers with two classrooms (described in Section B).

We organized the data on the classroom-/teacher-level file into a set of classroom-, teacher-, center-, and program-level identifiers, followed by two larger blocks of variables (Exhibit VI.3). **Cohort**, a flag, indicates whether the class was in a program—and thus a center—with child-level data collection. (programs in Cohorts 1 and 4 had child-level data collection). The first block of variables includes fall 2021 and spring 2022 teacher composite variables (for example, **T2CTRTIO**, which is the child/teacher ratio derived from the spring teacher survey). The second block contains data on items from the fall 2021 and spring 2022 teacher surveys.[44]

Exhibit VI.3. Fall 2021 and Spring 2022 classroom- and teacher-level data file structure

Identifiers
Unique identification numbers for classroom, teacher, center, and program
Cohort indicator identifying programs with child-level data
Constructed/derived variables (fall 2021 and spring 2022)
Data flags for the teacher survey (yes/no)
Analysis weights and clustering and stratification variables
Composite variables for the teacher survey (fall 2021 and spring 2022)
Survey data (fall 2021 and spring 2022)
Items and responses for the teacher survey (fall 2021 and spring 2022)

c. Fall 2021–spring 2022 merged child-level file

The merged child-level file includes fall 2021 and spring 2022 data from the TCR and parent survey collected for the children in the 60 programs with child-level data. It also includes identifiers that may be used to link data across the 2021–2022 study data files and composite variables from the teacher, center director, and program director surveys linked to each child. The file includes a single data record for each of the 1,837 eligible and consented children, regardless of whether there are data from the parent survey or TCR from fall 2021 or spring 2022.

We organized the data in the merged child-level file into a set of child-, classroom-, teacher-, center-, and program-level identifiers; demographic variables, such as the child's sex, age, and race/ethnicity; and two larger blocks of variables (Exhibit VI.4). The first block includes composite variables from the fall 2021 and spring 2022 TCR, parent survey, and teacher survey, and the spring 2022 program director and center director surveys. The second block of variables contains item-level data from the fall 2021 and spring 2022 TCR and parent survey.

The first block of composite variables begins with a series of data flags indicating whether there are completed instruments for the child, followed by weight variables for various types of analyses. The remaining composite variables in the block include the following: (1) indirect assessments derived from parent and teacher reports of children; (2) child and family characteristics; (3) parent processes and

[44] Given that some teachers taught more than one of the classes in the sample, a portion of the teacher survey data is replicated on both of the appropriate class records.

parenting; and (4) characteristics of Head Start classrooms, centers, and programs (organized by data source and data collection period).

The second large block of variables includes the item-level data from each child-level survey instrument. All items are included in the data file except child demographics and income variables from the parent survey, and copyrighted material from TCRs. Mathematica has negotiated agreements with different publishers and instrument developers. The terms of some agreements limit the distribution of item-level data. In addition, the child demographics and income variables are dropped from the parent survey to protect data quality. Users should use the composite variables, which have been cross-checked against other sources (child demographics) or imputed (income variables).

Exhibit VI.4. Fall 2021–spring 2022 merged child-level data file structure

Identifiers and demographic characteristics
Unique identification numbers for child, classroom, teacher, center, and program[a]
Child's age as of fall 2021 data collection
Child's sex/gender
Child's race/ethnicity
Flag to indicate whether child is new to Head Start or is returning for a second year
Composite variables (fall 2021 and spring 2022)
Data flags for each survey instrument (yes/no)
Analysis weights and clustering and stratification variables
Composite variables for the TCR (fall 2021 and spring 2022)
Composite variables for the parent survey (fall 2021 and spring 2022)
Composite variables for the teacher survey (fall 2021 and spring 2022)
Composite variables for the center director survey (spring 2022 only)
Composite variables for the program director survey (spring 2022 only)
Survey data (fall 2021 and spring 2022)
Teacher child rating items and (noncopyrighted) item responses (fall 2021 and spring 2022)
Parent survey items and (noncopyrighted responses (fall 2021 and spring 2022)

[a] The identification numbers associated with each child, classroom, center, and program remain the same across all waves. However, to allow for data linkage for children who moved to new classrooms, we provided separate classroom and teacher IDs in the child-level file for fall 2021 and spring 2022.

2. Variable names and labels

Variable names in the data files use a standard set of prefixes to indicate the source instrument for each data item (for example, parent survey, TCR, teacher survey, and so on) and data collection wave (fall 2021 and spring 2022). The first (and sometimes second) letter stands for the source instrument; the following number is the wave (Wave 1 for fall 2021 data and Wave 2 for spring 2022 data). For example, fall 2021 data items from the parent survey begin with P1; data items from the spring 2022 teacher survey begin with T2.

For survey item-level data, the remaining characters of the variable names correspond to the survey item numbers in the source instruments. For composite variables, the remaining characters consist of mnemonics. For example, the question in the fall parent survey about how often someone in the family

reads to the child is **P1D01**, whereas the composite variable from the same source on the number of people in the household is **P1HHSIZE**.[45]

We used a different naming system for analysis weighting variables, which require data from several data sources or waves. The variable names use the same letter prefixes to represent data sources, but in varying combinations. We also used the numbers 1 and 2, representing the fall 2021 and spring 2022 data collection waves.

Exhibit VI.5 lists the letter codes used to represent source instruments in the first character of the variable names.

Exhibit VI.5. Source codes for the 2021–2022 study instruments

Source instrument	Code (1st digit(s))
TCR	R
Parent survey	P
Teacher survey	T
Center director survey	C
Program director survey	D

Because many variables are repeated in the fall and spring, we included the first two characters of the variable name at the beginning of each variable label, affording users easy and rapid identification of the source and data collection wave. After the first two characters, a colon is followed by a general statement of the question/item content. For example, the label for the variable **P1A08** is "P1: Child participated in Early Head Start."

The 2021–2022 study surveys contain many multipart questions. For example, the household information section in the parent survey asks Questions B3 through B5 for each household member. Because the survey item number is part of the variable name, we developed a naming convention for multipart questions. The following are the rules used for naming variables in the 2021–2022 study data files:

- We always use two digits for the main numeric part of the question number, inserting leading zeros for Items 1 through 9 unless this approach creates names that exceed eight characters. For example, Question A8 in the fall 2021 parent survey is named **P1A08**.
- No underscores are used in variable names between question numbers and letters (for example, A.1.a is **A01a**). For example, Question D3 in the fall 2021 parent survey, which has 14 parts (a–n), contains the following names: **P1D03a**, **P1D03b**, . . ., **P1D03n**.
- Underscores are used in variable names to separate *numeric* parts of question numbers (for example, A1.1 is **A01_1**) unless the underscores make names exceed eight characters. For example, Question A23 in the spring 2022 teacher survey, which has six parts, contains the following names: **T2A23_1**, **T2A23_2**, . . ., **T2A23_6**.

[45] Many of the variables use an "n" in the second position of the name, indicating that the variables were created for more than one wave of the study. For example, **RnDISB2** indicates that the teacher-reported indicator for whether a child had a diagnosed disability is defined at each of the two waves of data collection (fall 2021 and spring 2022).

- If any list of numbered items in a question exceeds 10 items, leading zeroes are used for Items 1 through 9. For example, question F2 in the fall 2021 TCR, which has 12 parts, contains the following names: **R1F02_01**, **R1F02_02**, . . ., **R1F02_12**.
- On "select all that apply" item lists, where the separate items are coded 0/1, 0/2, … 0/n and the items are *not* listed in the questionnaire as a., b., . . ., the variables are named with the corresponding numeric code. For example, respondents were allowed to identify languages other than English spoken in their home (D8 in the parent survey). This set of items in the fall 2021 parent survey contains the following names: **P1D08_11**, **P1D08_12**, **P1D08_13**, . . ., **P1J01_31**.

3. Data processing

In the child-level data file, we adjusted the variables for certain questions in the parent survey that were asked only in the first survey conducted with the respondent. For most respondents, we asked these questions during the fall 2021 survey, but for parents who did not complete a fall survey, we first asked the questions in spring 2022. To adjust these variables for the respondents who did not take the fall 2021 survey, we merged the data from their spring survey with earlier data and dropped those spring 2022 items from the files. The affected variables appear in their original sequence in the data file among the fall 2021 variables, but the prefix on the variable names has changed from "P1" to "Pn" (for example, the mother's highest grade completed changed from **P1J24** to **PnJ24**). The data flag **Pn_first** is coded to show the source round of the "Pn" variables, with 1 = fall 2021 and 2 = spring 2022 (and missing for those who never completed a parent survey). The affected variables include demographic information on the child, mother, father, and respondent from Sections A, B, J, K, and L; information on activities done with the child in Section D; information on family food security from Section M; information on child health in Section P; and information on social support in Section T.

4. Data security and privacy

Personally identifying information (PII), such as center names and addresses or respondents' names and contact information, were dropped from the 2021–2022 study data files. In addition, we dropped verbatim responses from other-specify fields from all files. We included de-identified verbatim responses to open-ended items in the parent survey, teacher survey, program director survey, and center director survey related to COVID-19 experiences in each file. Finally, we replaced all ID variables with randomized ID versions.

The 2021–2022 study data are available as restricted-use files for the research community and therefore present a higher risk of deductive disclosure of the identity of a person or center. For this reason, these files went through additional "data coarsening" steps to protect respondent privacy. Data modifications include bottom- and/or top-coding and rounding of values, and combining small categories into the "other" category. In addition, to protect respondents' privacy, we rounded data collection dates to month and year. These data coarsening steps will cause these files to yield numbers slightly different than those presented in the report on the 2021–2022 study.

Even when we remove though PII and coarsened certain variables, certain combinations of variables may still allow deductive disclosure of the identity of a person or center. For this reason, the 2021–2022 study data files should be maintained securely to protect privacy in accordance with the restricted-use agreements.

5. Missing value codes

All variables in the data files use a standard scheme for identifying and explaining missing data. The following codes (Exhibit VI.6) identify data missing because of item nonresponse (missing data on items within a given instrument), legitimate skips, and unit nonresponse (an entire instrument is missing for the case).

Exhibit VI.6. The 2021–2022 study missing value codes

SPSS code	SAS code	Description
-1	.N	Not applicable, including legitimate skips based on earlier (screener) responses or conditions of who is or is not to answer a question or question set (for example, the center director survey asks certain questions only if the center serves children or families that speak a language other than English at home)
-7	.R	Refused (a type of item nonresponse)
-8	.D	Don't know (a type of item nonresponse)
-9	.M	Not ascertained (a type of item nonresponse), referring to items that were skipped but should have been answered (different from -1/.N above, which are skips or omissions by design)
SYSMIS	.	System missing (dot missing), including unit nonresponse where the full set of data for an instrument is missing, and some composite variables that could not be defined for certain cases

6. Identification numbers

The 2021–2022 study child-level data file contains **ChildID**, a six-digit child-level identifier, and ID numbers for the child's Head Start classrooms, teachers, centers, and programs at the fall 2021 and spring 2022 data collection points. The same identifiers are used on the separate data files for the fall 2021 and spring 2022 classroom-/teacher- and 2022 center-/program-level data. The classroom-level identifiers, **CLS1_ID** and **CLS2_ID**, are five-digit numbers indicating the child's Head Start classroom during fall 2021 and spring 2022. (In a large majority of cases, the values of **CLS1_ID** and **CLS2_ID** are the same.) The 2021–2022 child-level data file may be merged into the fall 2021 and spring 2022 classroom-/teacher-level data file by using **CLS1_ID** or **CLS2_ID**, renamed **CLS_ID**, which is the primary identifier in the 2021–2022 classroom-/teacher-level data file.

The teacher identifiers, **T1_ID** and **T2_ID**, are seven-digit numbers containing embedded information on classes taught by the teacher during fall 2021 and spring 2022. They are teacher-classroom combination identifiers, so the same teacher can have more than one teacher ID.[46] The first five digits contain a common value for classrooms taught by the same teacher, whether taught concurrently (for example, a morning and an afternoon class) or across time (fall and spring). The sixth digit indicates the number of classrooms associated with the teacher for that time period (fall or spring). When the sixth digit is zero, the teacher had only one classroom for that time period, whereas values of 1 or 2 are used in the separate class-level records for a teacher with data for two classrooms in that time period. The seventh digit has no specific meaning (it is a control digit), but the full seven-digit ID must be used when merging the data

[46] Some teacher IDs may have a leading zero, which may be lost when the file is converted from SAS if the ID is converted from a character to a numeric format.

with data on other 2021–2022 study data files. **T1_ID** and **T2_ID** are on both the 2021–2022 child-level and classroom-/teacher-level data files.

The center identifier, **C2_ID**, is a five-digit number identifying the Head Start center. **C2_ID** is in the spring 2022 center-/program-level, 2021–2022 classroom-/teacher-level, and 2021–2022 child-level data files.

The program identifier, **D2_ID**, is a four-digit number identifying the Head Start program in which the center is located. It is in the spring 2022 center-/program-level, 2021–2022 classroom-/teacher-level, and 2021–2022 child-level data files.

7. Merging data from different 2021-2022 Study files

Exhibits VI.1, VI.2, VI.3, and VI.4 in Section A provide lists of all 2021–2022 study data files. When working with the 2021–2022 data, certain weights and source data—center and program director surveys from spring 2022—are not included on the 2021–2022 child-level file, but only on the separate spring 2022 center-/program-level file. The child-level file does include all composite variables from all 2022 sources. In this section, we provide information on how to merge different 2021-2022 Study.

In conducting analysis, it is important to note that the center-/program-level files have one record per center, and contain both program- and center-level data. The file needs to be unduplicated to the program level if used as a program-level file. The classroom-/teacher-level file has one record per classroom and contains teacher survey data. It does *not* need to be unduplicated to the teacher level for teacher-level analysis. The teacher-level weights are only provided for the primary classroom for the affected teachers.

Common identifiers, provided to merge data for analyses that need such combined data sources across the 2021–2022 data files, are used as follows:

- To merge fall 2021 classroom-/teacher-level source data (all variables with the prefix T1 on the classroom-/teacher -level file) with the 2021–2022 child-level data file, use CLS_ID. On the child-level file, CLS1_ID should be renamed CLS_ID before the data merge.
- To merge spring 2022 classroom-/teacher-level source data (all variables with the prefix T2 on the classroom-/teacher -level file) with the 2021–2022 child-level data file, use CLS_ID. On the child-level file, CLS2_ID should be renamed CLS_ID before the data merge.
- To merge spring 2022 center director source data with the 2021–2022 child-level data file or perform analyses with child data aggregated to the center level (and appropriate weights and clustering/stratification information), use C2_ID.
- To merge spring 2022 program director source data with the 2021–2022 child-level data file or perform analyses with child data aggregated to the program level (and appropriate weights and stratification information), the program-level data from the 2022 center-/program-level file should first be unduplicated as described above. Then, use D2_ID to merge the two files.
- To merge spring 2022 center director source data with the 2021–2022 classroom-/teacher-level data file, use C2_ID (which is the primary identifier in the center-/program-level data file).
- To merge spring 2022 program director source data with the 2021–2022 classroom-/teacher-level data file, the program-level data from the center-level/program-level file should first be unduplicated as described above. Then use D2_ID to merge the files.

B. Analysis weights

Mathematica created nonresponse-adjusted sampling weights to account for variations in the probabilities of selection and differential response rates among those selected[47]. For each stage of sampling (program, center, teacher, and child) and within each explicit sampling stratum, we calculated the probability of selection. The inverse of the probability of selection is called the sampling weight, which takes into account the PPS (including certainty selections), equal probability sampling approaches, and the actual number of cases released. We assumed that the eligibility status of each sampled unit is known at each stage as a result of attempting to collect data. Then, at each stage and within each higher-level sampling unit (used as the weighting cells), we multiplied the sampling weight by the inverse of the weighted response rate so the respondents' weights account for both the respondents themselves and nonrespondents. The resulting products are called analysis weights.[48]

In this manner, the **program-level weights** adjust for the probability of program selection and response at the program level; **center-level weights** adjust for the probability of center selection and center-level response; **teacher-level weights** adjust for the probability of classroom selection and classroom-level response; **classroom-level weights** adjust the teacher weights for number of classrooms taught by the teacher; and **child-level weights** adjust for the probability of child selection and child-level response (accounting for parental consent as well as attrition from fall to spring). The formulas below represent the various weighting steps, where P is the probability of selection and RR the weighted response rate at that stage of selection.

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

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

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

$$W_{child} = W_{teacher} \cdot \frac{1}{P_{child}} \cdot \frac{1}{RR_{child}}$$

For the nonresponse adjustments, we used a weighting cell approach; that is, we inflated the weights for respondents within a cell to account for the eligible nonrespondents within the same cell. In other words, we applied an adjustment factor to the prevailing weights for the cell's respondents so their adjusted weight sum was the same as the sum of the prevailing weights for both respondents and nonrespondents

[47] As noted in Chapter II, due to higher than expected nonresponse, these data may not be fully representative of the population of Head Start children, families, teachers, centers, and programs even with sampling weights applied. See Appendix I for in-depth discussion of non-response bias analyses conducted on these data.

[48] In our reporting products, we use weights because when children are selected to be in the sample, they have different probabilities of being selected. For example, larger programs and centers are more likely to be selected than smaller ones, and we select a fixed number of classrooms per center. This approach means that some children may be more likely to be selected into our sample than others. Additionally, we use weights to adjust for changes in children's eligibility status and the effects of nonresponse.

in the cell. The cells were defined in such a way that we assumed those sample members in the same cell (whether respondents or not) were more homogeneous in what their responses would be than across cells. In addition, the information used to form the cells had to be available for both respondents and nonrespondents.

When adjusting for nonparticipation at the program level, we first tried to use the program pair as the weighting cell. As described in Chapter II, we selected an augmented sample of programs with more than twice the target number of participants. Among those sampled, we then formed program pairs within sampling strata based on program enrollment and randomly selected one member of each pair to be released initially, with the other to be released if the first one did not participate (due to ineligibility or refusal). If both members of a pair were released, both were eligible, and one responded, we weighted the participating program to account for its paired nonparticipating program. Otherwise, we used the program sampling strata as the weighting cell to account for program nonparticipation. We also used program sampling strata as the weighting cell for nonresponse adjustments for the program director survey. These strata were based on the program's census region, whether it was in a metropolitan statistical area, and whether it had more than 40 percent Black or Hispanic enrollment.

For nonresponse adjustments at the center level, we used the program from which they were sampled as the weighting cell. If neither sampled center within the program responded, we used the program sampling strata as the weighting cell to account for center-level nonresponse. Using the program as the primary weighting cell is based on the homogeneity of centers within a program as evidenced by the ICC for programs found in prior FACES studies. ICCs vary depending on the particular estimate, as some measures are more clustered than others. For classroom-level estimates, we found that about 20 percent of the total variance is explained by what program the classroom is in and about 20 percent by what center the classroom is in, leaving only 60 percent explained by differences between classrooms within the center. For child-level estimates, we found that about 10 percent of the total variance is explained by what program the child is in; about 5 percent by what center the child is in; and about 12 percent by what classroom the child is in, leaving less than 75 percent of the total variation explained by differences among children within a classroom (based on early 2019 design planning analyses with FACES 2014 spring 2015 data).

For nonresponse adjustments at the teacher and child levels (including nonconsent adjustments), we used the center from which they were sampled as the weighting cell. If there were no respondents within the teacher's or child's center (either to an individual instrument or a combination, depending on the analysis weight definition), we adjusted within the teacher's or child's program. If there were no respondents within the teacher's or child's program, we adjusted within the program's sampling stratum. Using the center as the primary weighting cell is based on the ICC for centers found in prior FACES studies.

In the 2021–2022 study, many analyses will focus on the teacher or classroom level and others on the child level. Analyses may also focus on higher levels (program or center) by using data collected at those levels to answer questions about Head Start programs and centers. The program-level weights are for analyses at the program level, including the program director survey, and are a component of weights in all subsequent stages of sampling. In addition, the center-level weights, which will be used for any analyses at the center level, including center director surveys, are a component of the teacher- and child-level weights. The teacher-level weights are for analyses of the teacher survey at the teacher level and are a component of the child-level weights. Given that some teachers teach two classrooms, we calculated two weights for the teacher survey: one at the teacher level and another at the classroom level. The child-

level weights that build on the program-, center-, and teacher-level weights are for analyzing data on child outcomes (teacher ratings) both alone and together with data from the parent survey.

Recognizing that only 60 of the 176 study programs participated in child-level data collection, we constructed two sets of program-, center-, and teacher-level analysis weights for fall 2021 and spring 2022, each one representing all study-eligible programs in Head Start to the extent possible. The first set includes all 176 participating spring 2022 Program and Staff Study programs, and is intended for program-, center-, and teacher-level estimates (and are included in the data files). The second set includes only the 60 subsampled Program, Staff, and Family Study programs, accounting for program subsampling. These weights are used only as building blocks for child-level estimates, and not for analysis (and are not included in the data files).

1. Fall 2021 weights

Exhibit VI.7 shows the number of sampling units selected and released for the fall 2021 sample, the number of eligible units, and the number of those units that agreed to participate in the study for the 60 programs with child-level data collection (those in the Program, Staff, and Family Study) only. (For children, this number comprises eligible children whose parents gave consent for them to participate through the end of fall data collection.) The sum of the child weights, which is our best estimate of the number of study-eligible children enrolled in Head Start in fall 2021, is 450,054. We used the program and center weights from the fall only as building blocks for teacher- and child-level weights, and did not include them in the file. We did not collect any data at those levels in fall 2021.

Exhibit VI.7. Fall 2021 sample: Program, Staff, and Family Study

	Program	Center	Teacher	Child
Selected and released	115	116	239	3,139
Eligible	104	116	236	3,105[a]
Participating/consented	60	114	236	1,363
Sum of fall 2021 weights	n.a.	n.a.	32,929	450,054

a. These children consisted of those still in the sampled Head Start program at the time of data collection. There were 30 additional children who had left but were eligible at the time of sampling, for a total of 3,135.

n.a. = not applicable.

In Chapter II, we describe what made a Head Start program eligible for participation in the 2021–2022 study. Accounting for the FACES 2019 original sample as the base that was then freshened (see Chapter II), 313 programs were sampled and released for recruitment. Thirty-two programs were deemed ineligible, and 105 eligible programs declined to participate, for an unweighted participation rate of 63 percent.[49]For a child to be eligible for the Program, Staff, and Family Study, they had to be enrolled in the selected Head Start center at the time of sampling.[50] A child moving to a teacher different from the one

[49] The unweighted response rate is the number of participating programs divided by the number of sampled and eligible programs.

[50] As described in Chapter IV, sampling took place remotely. A Mathematica liaison obtained lists of teachers and their child rosters from programs electronically, then selected probability samples of teachers and children by using built-in sampling routines in a web-based application.

the child was sampled from was still eligible for data collection if they remained enrolled in a sampled center.[51]

In addition, we considered 30 children who were ineligible according to the study's operational protocol to be part of the target population for weighting purposes. These children left their selected Head Start center between the time of sampling and the beginning of the spring 2022 data collection period, and we did not pursue data collection for them. They may have left Head Start altogether or moved to a different center or program that was not part of the study sample, or the program may not have known where the child went after dropping out of their center. Because they were part of the Head Start population at the time of sampling, they were deemed eligible. We considered 3,135 children to be part of the target population for weighting purposes, including these 30 children.

Exhibit VI.8 is a list of weights created for analyzing data from fall 2021 (baseline). We provide one weight for analysis of the fall 2021 teacher survey (**T1TCHWT**) and four weights for child-level analyses of fall 2021 data. **P1WT**, one child-level weight, is designed for users who will use data from the fall parent survey. **PR1WT**, another child-level weight, is designed for users who will use data from both child-level instruments: the parent survey and TCR. The PR1WT weight excludes from the analysis any child missing one or both of these components. There are 785 children with a fall 2021 parent survey and 537 children with data from both the parent survey and the fall TCR. We provide weight **R1WT** for analysis of the fall 2021 TCR data on their own. We have also included a child-level base weight (**CNST1WT**) that accounts for selection probabilities and parental consent for study participation (through the fall data collection period). However, it is not tied to any data collection instrument and may be used for child-level estimates not involving a particular instrument—for example, an examination of which children stay in Head Start for the full program year.

Exhibit VI.8. Fall 2021 cross-sectional weights

Weight name	Unit of analysis	Weight to be used for analysis of:	Records with positive weight
T1TCHWT	Teacher	Fall 2021 teacher survey	191
P1WT	Child	Fall 2021 parent survey data	785
PR1WT	Child	Fall 2021 parent survey data in combination with TCR data	537
R1WT	Child	Fall 2021 TCR	887
CNST1WT	Child	Fall 2021 child study participation	1,363

2. Spring 2022 weights

By spring 2022, the sample of children who were eligible and consented for child-level data collection increased from 1,363 to 1,837; 346 children became ineligible for the study because they left Head Start (including the 30 who had left the program before fall 2019 data collection), and 57 left the selected Head Start center for another one that was not part of the study (Exhibit VI.9). However, because we let the

[51] In fall 2021, the 2021–2022 study excluded children who were no longer in the sampled center at the end of the data collection period (as of January 31, 2021). In spring 2022, the study excluded children no longer in the sampled center at the beginning of the data collection period (as of April 29, 2022).

parental consent process continue beyond the fall 2021 data collection period, we got an additional 582 consents among those sampled children who were still eligible in spring 2022.

To provide greater precision for analysis of Head Start programs, centers, classrooms, and teachers, our goal for the spring 2022 staff-level data collection was to recruit an additional 120 programs in spring 2022 (mostly from spring 2020 participating programs) and sample their centers and teachers, to be combined with the 60 programs recruited in fall 2021 and their sampled centers and teachers. For the spring-only programs, we released 198 for spring 2020 data collection. Of them, 21 were ineligible for FACES and, among the 177 eligible programs, 61 did not agree to participate. Combined, we had 116 programs participating in the spring 2022 data collection in addition to the 60 from the fall 2021 collection.

Exhibit VI.9. FACES spring 2022 sample

	Program	Center	Teacher	Class	Child
From fall 2021 and eligible and consented in spring 2022	60	113	235	241	1,837
Eligible and recruited/consented for spring 2022	116	227	396	422	n.a.
Total eligible and recruited/consented spring 2022	176	340	631	663	1,837
Sum of spring 2022 weights	1,425	11,664	30,541	32,215	402,460

Some teachers taught more than one classroom during the spring 2022 data collection period. In such cases, we asked the sampled teacher to complete the teacher survey once but rotate twice through the section on classroom characteristics. When the teachers with two classes were selected, we constructed their associated classroom-level weights to account for their two classes. The sample included six teachers who had two classes in the 60 programs in the Program, Staff, and Family Study with child-level data, and 26 in the additional 116 programs with no child-level data (spring-only programs).

Nine teachers who had not originally been selected to participate in the 2021-2022 study had sampled children move into their classrooms by spring 2022. These "mover" teachers were considered ineligible for teacher- and classroom-level weights, but the children were still considered eligible for the child-level weights, and their teachers were asked to complete both a teacher survey and a TCR.

We created six weights for cross-sectional analysis of the spring 2022 data and three weights for longitudinal analysis for the program year (fall 2021 and spring 2022 combined). For analysis of the spring 2022 data, we created weights for staff surveys (program director, center director, teacher) and for data collected at the child level.

Exhibit VI.10 shows the 10 cross-sectional instrument weights for spring 2022. Note that for some of the child-level weights we consider to be cross-sectional, we allow for the parent survey to have been completed in either fall 2021 or spring 2022. The child weights among them allow for a fall 2021 parent survey *or* a spring 2022 survey (that is, for a fall 2021 parent survey to be used in lieu of a parent survey in spring 2022 that we did not obtain) and could be considered longitudinal in that respect only. The weights sum to the eligible population in spring 2022: the staff of programs in spring 2022, the classrooms in those programs, and children who were enrolled in Head Start in fall 2021 and still attending Head Start in spring 2022 (Exhibit VI.9). It is important to note that child-level analysis may

use data collected at the program, center, or class/teacher level as long as the data are appended to each child's record as contextual data.

We have also included in this Exhibit four base weights: a program-level base weight (**D_WT_S**); a center-level base weight (**C_WT_S**); a teacher-level base weight (**TCH_WT_S**); and a child-level base weight (**CNST2WT**). These weights account for selection probabilities and study participation through the spring data collection period but are not tied to any data collection instrument. These weights may be used for estimates not involving a particular instrument.

Exhibit VI.10. Spring 2022 cross-sectional weights

Weight name	Unit analysis	Weight to be used for analysis of:	Records with positive weight
D_WT_S	Program	Participating programs (for example, summarizing characteristics of centers, classrooms, or children in those programs)	176
D2WT	Program	Program director survey among participating programs	132
C_WT_S	Center	Participating centers (for example, summarizing characteristics of classrooms or children in those centers)	340
C2WT	Center	Center director survey among participating centers	237
CP2WT	Center	Center director and program director surveys in spring 2022	179
TCH_WT_S	Teacher	Teacher participation in study in spring 2022	631
T2TCHWT	Teacher	Teacher survey data in spring 2022	358
TCP2WT	Teacher	Teacher survey and center director and program director surveys in spring 2022	207
T2CLSWT	Class	Teacher survey data at classroom level in spring 2022	372
CNST2WT	Child	Participating children who were selected for the sample and consented in fall 2021 or spring 2022, and remained in the program in spring 2022	1,837
P1_2WT	Child	Parent survey data in fall 2021 or spring 2022	1,159
P21R2WT	Child	Parent survey data in fall 2021 or spring 2022 in combination with TCR in spring 2022	802
P21RT2WT	Child	Parent survey data in fall 2021 or spring 2022 in combination with spring 2022 TCR and teacher survey	778
R2WT	Child	TCR data in spring 2022	1250

TCR = teacher child report.

Users may also conduct analyses at higher levels (at the center or program level) by aggregating some data collected at the teacher or classroom level to answer questions about Head Start centers and programs in spring. For more information, see Section C, Choosing the best weight.

3. Fall 2021–spring 2022 longitudinal analysis

The program year child-level longitudinal weights sum to the eligible child population in spring 2022 (Exhibit VI.9)—that is, children who were enrolled in Head Start in fall 2021 and still enrolled in spring

2022. In Exhibit VI.11, we show one teacher-level weight and four child-level weights for the program year.

Exhibit VI.11. Fall 2021–spring 2022 longitudinal weights

Weight name	Unit Analysis	Weight to be used for analysis of	Records with positive weight
T12TCHWT	Teacher	Teacher survey data in fall 2021 and spring 2022	146
PR12WT	Child	Parent survey data in fall 2021 or spring 2022 in combination with TCR data in fall 2021 *and* spring 2022[b]	503
P12WT	Child	Parent survey data in fall 2021 *and* spring 2022	505
PR12T2WT	Child	Parent survey data in fall 2021 or spring 2022 in combination with TCR data in fall 2021 and spring 2022[*] *and* teacher survey data in spring 2022	486
R12WT	Child	TCR in fall 2021 and spring 2022	680

[*]Meaning that a parent survey had to be completed in the fall or spring and the TCR had to be completed in both fall and spring.

4. Nonresponse bias analysis

Nonresponse bias can occur when the responses of survey nonrespondents would have been different enough from those of survey respondents to make the survey results not representative of the full sample. Nonresponse bias analysis examines (1) whether estimates of important outcomes seem like they might be biased because certain people did not respond, based on comparing those who responded to the survey with those who did not, and (2) whether analysis weights lessen the severity of this bias. As response rates decrease, the risk of nonresponse bias for an estimate increases. Although bias cannot usually be directly measured, we can look for indications of the potential for bias for key outcomes and examine whether the nonresponse-adjusted weights appear to have mitigated the risk for bias. There is no rule of thumb for how large a bias is acceptable, but the larger it is, the more caution is merited in analysis.

As discussed in Chapter IV, participation and response rates for fall 2021 and spring 2022 were lower than anticipated due to the COVID-19 pandemic and associated challenges. To investigate whether the responses from Head Start programs, centers, staff, and families that participated in the 2021-2022 study differed from the responses we would have gotten from those that were unable to participate in the study, we conducted a nonresponse bias analysis (NRBA) on data collected in fall 2021 and spring 2022 using covariates. We analyzed program, center, and child-level covariates for indications of potential bias due to nonresponse. The covariates are drawn from the Head Start Program Information Report, information from the Office of Head Start, and information obtained as part of the center and child sampling process. Appendix I outlines the NRBA conducted in greater detail.

Although we examine a wide variety of covariates that are available for both respondents and nonrespondents, we are unable to test the potential for bias across all characteristics that might be important in assessing the risk for bias. For example, we collect information about center size (number of classrooms) from Head Start programs as part of the center sampling process and can therefore examine this in the NRBA. In contrast, although the 2021-2022 Study parent survey collects information about parent depression, we are unable to include parent depression in the NRBA because we do not know how nonrespondents would have answered those items.

Exhibits VI.12 through VI.14 summarize our conclusions from the NRBA at a very high level. Exhibit VI.12 outlines fall 2021 NRBA, Exhibit VI.13 outlines spring 2022 staff level NRBA, and Exhibit VI.14 outlines spring 2022 child level NRBA. The rows in each table list the weights for which we conducted an NRBA. The second column indicates if there are indicators of the risk for meaningful nonresponse bias for survey items remaining after weighting based on the covariates that were tested.[52] The remaining columns indicate whether users should control for the specified NRBA covariate in analyses using each weight. In a modeling context, potential bias due to nonresponse may be mitigated by controlling for any provided NRBA covariates for which there were large differences between the weighted respondents' and full sample estimates.

Users should note that **blank cells in these exhibits do not necessarily indicate a lack of meaningful nonresponse bias**, but rather that the weights mitigate *observable* differences on our covariates.

[52] Appendix I summarizes how we interpreted findings from the NRBA to draw these conclusions. Briefly, we considered the size of the difference between the weighted respondents' estimate and the full sample (based on the effect size and the difference in the covariate's original scale), and whether the covariate was correlated with key study outcomes in the responding sample.

Exhibit VI.12. Summary of fall 2021 nonresponse bias analyses

Fall participation or instrument	Weight name	Indicators of meaningful nonresponse bias after weighting?	Control for covariate in analysis?				
			Program in a metropolitan statistical area (METRO)	Program enrollment (A_15_CAT)	Child sex (CHGENDER2)	Child participation in Early Head Start (EHS_Participant)	Percentage of teachers and home visitors in the program who left that were replaced (PTEACHHVREPL_CAT)
Program participation	FALL21_PROGRAM_WEIGHT	Yes	√	√			
Child consent	CNST1WT*		√	√			
Parent survey	P1WT*		√	√			
Parent survey + Teacher Child Report	PR1WT	Yes	√	√	√	√	√
Teacher survey	T1TCHWT*		√	√			

* Program-level participation is a building block for other weights and nonresponse bias analyses, so we recommend users control for all program participation NRBA covariates for which indicators of meaningful nonresponse remained after weighting.

Exhibit VI.13. Summary of spring 2022 staff-level nonresponse bias analyses

Spring participation or instrument	Weight name	Indicators of meaningful nonresponse bias after weighting?	Control for covariate in analysis?	
			Percentage of lead teachers or home visitors who left the program (PTEACHHVLEFT_CAT)	Center's number of teachers (C2A01)
Program director survey	D2WT	Yes	√	
Center director survey	C2WT*		√	
Teacher survey	T2TCHWT*		√	
Teacher survey + program director survey + center director survey	TCP2WT	Yes	√	√

* We did not conduct a separate nonresponse bias analysis for program participation in the Program and Staff Study (of which the programs participating in both fall and spring are a subset). However, we did conduct an NRBA for program director survey response in the Program and Staff Study programs. Because this accounts for program participation in addition to program director response, we recommend that for spring staff survey data, users control for the one program director survey NRBA covariate for which indicators of meaningful nonresponse remained after weighting.

Exhibit VI.14. Summary of spring 2022 child-level nonresponse bias analyses

Spring participation or instrument	Weight name	Indicators of meaningful nonresponse bias after weighting?	Control for covariate in analysis?					
			Program in a metropolitan statistical area (METRO)	Program enrollment (A_15_CAT)	Percentage of teachers and home visitors in the program who left that were replaced (PTEACHHVREPL_CAT)	Child sex (CHGENDER2)	Child age group (P1RCAGE)	Percentage of lead teachers or home visitors who left the program (PTEACHHVLEFT_CAT)
Child consent (fall or spring)	CNST2WT*		√	√				
Fall or spring parent survey	P1_2WT*		√	√				
Fall and spring parent survey	P12WT*		√	√				
Parent survey (fall or spring) + spring TCR	P21R2WT	Yes	√	√	√			
Parent survey (fall or spring) + fall TCR + spring TCR	PR12WT	Yes	√	√		√	√	√

* Program-level participation is a building block for other weights and nonresponse bias analyses, so we recommend users control for all program participation NRBA covariates for which indicators of meaningful nonresponse remained after weighting. Child-level data is collected in the 60 programs participating in fall, so we recommend controlling for the same two covariates flagged for the fall data (Exhibit VI.12) when using spring 2022 child-level data.

For additional information on the procedures used to conduct these NRBA, more in-depth discussion of key takeaways and recommendations for data users, correlations between the covariates used in these analyses and key survey constructs, and tables with detailed NRBA results, see Appendix I: Nonresponse Bias Analyses.

C. Choosing the best weight

In the preceding sections, we have described six base weights and 18 analysis weights. In this section, we provide the user with guidance in choosing among the weights. First and foremost, the choice of weight depends on the research question and the data required to answer the question—for example:

- What is the target population for the analysis and the level of analysis?
- What are the behaviors or characteristics of interest (what data will be used, and from which sample and source instruments)?
- Does the analysis focus on a single time point or on changes over two time points (which wave or waves of data are involved)?

Many analyses using the 2021–2022 study data will focus on the child level. At the same time, analyses at higher levels (for example, Head Start programs or classrooms) will be required to answer particular types of questions, or data collected at several levels will be used to answer questions about Head Start children, families, programs, centers, classrooms, and teachers. Each of these levels depends on particular instruments, and, in Exhibits VI.14, VI.16, and VI.17, we identify the weights to be used when analyzing data from the various instruments. For example, the program weight **D2WT** is intended for use with data from the program director survey in spring 2022 and adjusts for each program's probability of selection and program refusals, as well as for program director survey nonresponse. The center weight **C2WT** is intended for use with data from the center director survey in spring 2022. The teacher-level weights are intended for use in analyses at the teacher level that include data from the Head Start teacher survey in fall 2021 (T1TCHWT) or spring 2022 (**T2TCHWT**) or both (T12TCHWT). Users interested in Head Start classrooms and their characteristics would use the spring teacher survey weight for classroom estimates (**T2CLSWT**).

For the 2021–2022 data, the child-level weights that build on the program-, center-, and teacher-level weights should be used to analyze data on child outcomes from TCRs and parent survey data. In choosing a child-level weight, the user must also determine whether the analysis requires more than one wave of data. The 2021–2022 study includes two sets of cross-sectional (one wave) weights (one for fall 2021 [Exhibit VI.8] and one for spring 2022 [Exhibit VI.10]), and one set of longitudinal weights for both the 2021 and 2022 waves (Exhibit VI.11). Aside from matching the weight to the wave(s) of data that will be used, the choice among the different child weights will depend on the types of data required to answer a specific research question or series of questions. For example, will data from the spring 2022 teacher surveys be required to examine the influence of classroom characteristics on children's gains in school readiness skills (from the TCR) over the 2021–2022 program year? If so, **PR12T2WT** might be the best choice. Users interested in how children's family and home life have changed over the 2021–2022 Head Start year would most likely select **P12WT**.

Almost all of the child weights require a completed parent survey in fall 2021 and/or spring 2022. The reason is that the parent survey is the source for many of the child characteristics (age, race/ethnicity, and sex) and those of the family (primary home language, family structure, household income) often used when describing the Head Start population and examining children's school readiness skills.

It is important to recognize that there is no perfect weight for all potential research questions that may be answered with the 2021–2022 study data. In many cases, the user must decide between two or more weights, basing the decision on several factors, including balancing the risks of bias and variance for estimates while also considering ease of use. **A more restrictive weight** (more requirements for

completed instruments, such as **PR12T2WT**) will result in fewer cases for analysis and might mean being unable to use some collected data, but the weight should reduce the bias among the retained cases. **A less restrictive weight** (fewer requirements for completed instruments, such as **P21R2WT**) will have more cases with a positive weight, but if the requirements for the weight do not match the variables in an estimate, those cases with noncompleted instruments will simply drop out of the estimate with no adjustment, treating the dropped cases as if they were missing completely at random (which is unlikely), thus reducing the analytical sample size at the same time potential bias is being introduced.

For the spring 2022 Program and Staff Study, we created two program-level cross-sectional weights:

1. **D_WT_S**, which accounts for a program's probability of selection, eligibility, and study participation
2. **D2WT**, which accounts for whether the program director completed a survey

Users are more likely to use the second weight; however, if the user is interested in all participating programs and not responses to the program director survey—for example, for purposes of aggregating and summarizing data collected at the classroom level to the program level—the first weight would be preferable.

Similarly, for the spring 2022 Program and Staff Study, we created two center-level cross-sectional weights:

1. **C_WT_S** for analysis of all participating centers
2. **C2WT** for making estimates from the center director survey

As with the program-level D_WT_S, for aggregating data from the classroom level to the *center* level without using the center director survey data, **C_WT_S** is the appropriate weight. For analyses of classroom characteristics at the center level, the user could calculate the weighted proportion of such classrooms in each center, merge it to the center-level record, and then conduct the analysis at the that level by using the base center-level weight.

Users also need to make another choice—whether to conduct analysis at the classroom or teacher level. As with programs and centers, a general teacher-level (**TCH_WT_S**) or classroom-level (**CLS_WT_S**) weight is not dependent on the completion of any particular instrument and is to be used when analyzing all participating teachers or classrooms. When analyzing data from both the fall and spring teacher surveys (using the longitudinal T12TCHWT weight), we recommend controlling for the covariates recommended for fall 2021 (because only the teachers from programs participating in fall data collection would have a positive longitudinal weight); that is, whether the program was in a metropolitan statistical area (METRO) and program enrollment (A_15_CAT). This is because only the teachers from programs participating in fall data collection would have a positive longitudinal weight, and in the nonresponse bias analyses for the cross-sectional teacher weights, we saw no observable evidence of nonresponse bias.

Many users will use multilevel modeling to examine the associations between classroom or program characteristics and child outcomes. Given that FACES child weights take into account earlier stages of sampling (programs, centers, teachers), a multilevel model requires only weighting data with a single child weight at Level 1. That is, the user does not need to use a child weight for Level 1 and a different weight (for example, a classroom weight) for Level 2. If users are doing a center- or teacher-level analysis, they need only to weight the data with the appropriate center or teacher weight, respectively.

D. Cross-round and cross-study analyses

Data from the fall 2021 teacher survey support cross-study analyses with data from FACES 2019 or earlier rounds of FACES. However, users should note that these data were collected at a different point in the program year than past FACES teacher surveys, as FACES 2019 and earlier FACES studies only collected teacher surveys in the spring. Additionally, users should note that because the fall 2021 teacher survey was a shorter instrument focused mainly on teacher wellbeing, there are a limited number of variables available for cross-study analyses.

Due to lower than expected response rates, associated concerns about nonresponse bias (see section 5e below), and the resulting conclusions we can draw about who the 2021-2022 study sample represents, we do not recommend conducting cross-study analyses with data from any of the other 2021-2022 study instruments.

E. Variance estimation

Most standard procedures in commonly used statistical software packages, such as SAS, Stata, IBM SPSS, or R, can estimate specific characteristics. For example, they can construct weighted estimates of children's mean social skills score, percentage of children demonstrating particular learning skills, percentage of classrooms of particular sizes, and teachers with particular educational backgrounds and credentials. However, standard errors or variances will not be correct in their standard procedures because they do not account for sample design complexities (multistage clustered sample with unequal probabilities of selection) of surveys such as the 2021–2022 study. Instead, the procedures assume that the data or observations come from a simple random sample design and thus would underestimate the true variance.

These software packages all now provide design-based variance estimation methods such as replication and Taylor Series expansion account properly for the sample design. Specialized complex survey procedures in SAS, Stata, R, and IBM SPSS (as well as specialized packages such as SUDAAN) may generate the proper standard errors of the weighted estimates. The procedures use either the Taylor Series linearization method or the replication method to estimate the proper variance. In the 2021–2022 study, we do not provide weights to use with replication methods and therefore assume users will be using the Taylor Series method.

Users of survey data analysis software must specify the weight variable and the design variables, such as the first-stage sampling strata (STRAT) and—for analyses conducted at the child level—the primary sampling units (PSUs). When analyzing data at the classroom, teacher, or center levels, the user should use two slightly different versions of these variables to account properly for programs selected with certainty: STRAT_C and PSU_C. However, when running estimates based on the center director survey, statistical software designed for survey data (such as SDUAAN, SAS Survey or Stata Svy) will notice can be strata containing only one observation, which poses a problem for estimating the variance. For center-level analysis, STRAT_C = 888 is a singleton stratum. For teacher/class-level analysis, STRAT_C = 555 and 666 are singleton strata. Each package has its own options for dealing with this scenario. In SAS Survey, one can specify "/NOCOLLAPSE" at the end of the STRATA statement. In Stata Svy, one can specify "SINGLEUNIT(CERTAINTY)." In SUDAAN, one can specify "/MISSUNIT" at the end of the NEST statement. Please refer to the software documentation for other alternatives. Different options may slightly under- or overestimate standard errors. In general, the optimal approach is to collapse two similar

strata to avoid singleton strata. We have found that the choice of method has little impact on the standard errors.

As a cautionary note, for subgroup analysis, the full set of records should be read into the procedure with a statement to identify the subgroup to be analyzed. If the file is subsetted, the software may not be able to compute a sampling variance (that is, the program will stop); if it does compute a variance, it may substantially underestimate the sampling variance. In Stata Svy, the statement is subpop; in SAS, it is DOMAIN; in SUDAAN, it is SUBPOPN.

Below we present the values for stratum primary sampling unit and the weight that should be specified when running survey data analysis software using the Taylor Series method.

The following are the specifications for **program**-level analysis:

Sample Stata code (continuous variables)

```
use "[program data file]",clear
svyset [pweight = D2WT], strata(STRAT) singleunit(certainty)
svy: mean [continuous variable(s)]
```

Sample SAS code (continuous variables)

```
proc surveymeans data=[program data file];
    strata STRAT;
    weight D2WT;
    var [continuous variable(s)];
    title '[title]';
```

Sample Stata code (categorical variables)

```
use "[program data file]",clear
svyset [pweight = D2WT], strata(STRAT) singleunit(certainty)
svy: tab [categorical variable]
```

Sample SAS code (categorical variables)

```
proc surveyfreq data=[program data file];

    strata STRAT;
    weight D2WT;
    tables [categorical variable(s)] / deff nototal;
    title '[title]';
```

The following are the specifications for **classroom-, teacher-, and center**-level analysis:

Sample Stata code

```
use "[ classroom, teacher, or center data file]",clear
svyset PSU_C [pweight = T2TCHWT OR C2WT], strata(STRAT_C) singleunit(certainty)
svy: tab [categorical variable]
```

Sample SAS code

```
proc surveyfreq data=[ classroom, teacher, or center data file];
```

strata **STRAT_C**;
cluster PSU_C;
weight [**T2TCHWT** OR **C2WT**];
tables [categorical variable(s)] / deff nototal;
title '[title]';

The following are the specifications for **child**-level analysis:

Sample Stata code

use "[child data file]",clear
svyset PSU [pweight = **PRA1WT** OR **PR12WT**], strata(STRAT) singleunit(certainty)
svy: tab [categorical variable]

Sample SAS code

proc surveyfreq data=[child data file];
strata **STRAT**;
cluster **PSU**;
weight [**PRA1WT** OR **PR12WT**];
tables [categorical variable(s)] / deff nototal;
title '[title]';

This page has been left blank for double-sided copying.

VII. Indirect child assessment scores and composite variables

The 2021–2022 study measured Head Start children's development, home and family life, and classroom and program experiences. We constructed composite variables and indirect child assessment scores to facilitate using the 2021–2022 study data to answer important questions about Head Start children, their families, classrooms, staff, and programs. In this chapter, we describe the composite variables and scores included in the data files.

Chapter III details the survey items and indirect child assessments, including psychometric information from developers or publishers and FACES studies, as applicable. Appendix A also notes the indirect child assessments used in the 2021–2022 study and the direct and indirect assessments in FACES studies. This chapter describes the approach to composite construction, followed by a series of exhibits (Exhibits VII.1 through VII.5) that list the scores and composite variables in the data files. These files also include some data flags not noted in these exhibits. The flags identify cases with income data based on imputed information (discussed in this chapter) and other related data characteristics.

The 2021–2022 study instruments lent themselves to the development of a substantial number of scores and composite variables. Therefore, we limited the series of scores and composite variables we created and included on the data files to those most crucial for answering questions about Head Start children, families, staff, classrooms, and programs. We also prioritized creating variables based on data from several items, or that require considerable effort to construct. For this reason, the data files do not contain variables that may be simple recodes or transformations of a survey item or question, such as transforming a continuous survey item into a categorical variable. Many of the scores and composite variables for the 2021–2022 study have been used in FACES studies; others are new to the 2021–2022 study. Finally, unless specified otherwise, we created scores and composite variables across all of the 2021–2022 study instruments only for cases with valid responses on at least 75 percent of the necessary source items.

Each section of this chapter includes exhibits that present information on the variable or score name, label, source items used to construct the variable, values and possible response ranges, reliability (internal consistency using Cronbach's alpha) if applicable, and data collection waves for which the variable or score is available. The "n" used in a variable name is a placeholder for the wave number, with "1" referring to fall 2021 and "2" to spring 2022. Variables constructed in more than one wave are listed with the "n" in the text. In Appendix H, we provide additional information about the construction of these variables/scores.

A. Indirect assessment scores

1. The 2021–2022 study included indirect assessment scores only.

Indirect assessment scores are based on teacher or parent ratings of children's social-emotional behaviors and teacher ratings of approaches to learning and academic and nonacademic skills. We also used teacher reports of concerns about children's development, including disability status and type. In Exhibit VII.1, we list scores we derived from the teacher and parent reports.

Exhibit VII.1. 2021–2022 study composite variables—child indirect assessments

Instrument	Variable name	Variable label	Instrument/survey items	Values and possible response ranges	2021–2022 study reliability (Cronbach's alpha) fall 21, spring 22
Social skills					
Teacher Child Report	R1SSRS, R2SSRS	Teacher-reported social skills score	Copyrighted[a]	0–24	0.9, 0.91
Behavior					
Teacher Child Report	R1BAGGR, R2BAGGR	Teacher-reported problem behaviors: disruptive/aggressive score	Copyrighted [a]	0–8	0.86, 0.87
Teacher Child Report	R1BHYPE2, R2BHYPE2	Teacher-reported problem behaviors: hyperactive score	Copyrighted [a]	0–6	0.75, 0.79
Teacher Child Report	R1BWITH, R2BWITH	Teacher-reported problem behaviors: withdrawn score	Copyrighted [a]	0–12	0.75, 0.78
Teacher Child Report	R1BPROB2, R2BPROB2	Teacher-reported problem behaviors: total score	Copyrighted [a]	0–28	0.86, 0.88
Parent survey	P1BEHCHG, P2BEHCHG	Count of parent-reported changes in child behaviors since March 2020	PnG02a–PnG02d	0–4	n.a.
Parent survey	P1BANYBEHCHG, P2BANYBEHCHG	At least one parent reported child behavior change since March 2020	PnG02a–PnG02d	0–1	n.a.
Approaches to learning					
Teacher Child Report	R1APROHR, R2APROHR	Teacher-reported Approaches to	RnH01a–RnH01f	1–4	0.92, 0.93

Exhibit VII.1. *(continued)*

Instrument	Variable name	Variable label	Instrument/survey items	Values and possible response ranges	2021–2022 study reliability (Cronbach's alpha) fall 21, spring 22
		Learning – Revised Scale			
Parent survey	P1APROHR, P2APROHR	Parent-reported Approaches to Learning – Revised Scale	PnG03a–PnG03f	1–4	0.73, 0.74
Current learning skills					
Teacher Child Report	R1TPELS, R2TPELS	Teacher-reported child literacy behaviors	RnB01, RnB03, RnB05a, RnB05b, RnB05g	0–7	0.71, 0.66
Developmental conditions and concerns					
Teacher Child Report	R1DISB2, R2DISB2	Child Disability Status (Teacher)	RnF02_01…RnF02_10	0, 1	n.a.
Teacher Child Report	R1MLIMP2, R2MLIMP2	Multiple Disability (Teacher)	RnF02_01…RnF02_10	0, 1	n.a.
Teacher Child Report	R1SNSIMP, R2SNSIMP	Sensory Disability (Teacher)	RnF02_01, RnF02_02, RnF02_03, RnF02_04	0, 1	n.a.
Teacher Child Report	R1PHYIMP, R2PHYIMP	Physical Disability (Teacher)	RnF02_05	0, 1	n.a.
Teacher Child Report	R1LNGIMP, R2LNGIMP	Speech Disability (Teacher)	RnF02_06	0, 1	n.a.
Teacher Child Report	R1COGIMP, R2COGIMP	Cognitive Disability (Teacher)	RnF02_07, RnF02_08, RnF02_09	0, 1	n.a.
Teacher Child Report	R1BEHIMP, R2BEHIMP	Behavioral Disability (Teacher)	RnF02_10	0, 1	n.a.
Teacher Child Report	R1CHIEP, R2CHIEP	Child has IEP (Teacher)	RnF05_05	0, 1	n.a.

[a] Mathematica's agreement with the publisher/developer of these items does not allow us to share the items publicly without prior written approval.

IEP = individualized education program; n.a. = not applicable.

The "n" used in a variable name is a placeholder for the wave number, with "1" referring to fall 2021 and "2" to spring 2022. If a particular variable is listed with the "n" in the text, it was constructed in more than one wave.

We derived raw scores[53] of children's social skills, problem behaviors, approaches to learning, and language and literacy skills from the TCR, and of children's approaches to learning and changes in behavior from the parent survey. Raw scores are indicators of absolute performance calculated as the sum or mean of items, depending on the type of assessment. All scores indicate the extent to which given statements reflect a child's behavior or skills.

We calculated composite scores about children's behaviors from the TCR as follows:

- Teacher-reported social skills score (**RnSSRS)** is a sum of 12 items with a total of 24 possible points, all related to children's cooperative behavior and social skills. The items come from the Personal Maturity Scale (Entwisle et al. 1987) and the Social Skills Rating Scale (Gresham and Elliott 1990). Higher scores indicate the child exhibits cooperative behavior more frequently.
- Teacher-reported problem behaviors score (**RnBPROB2**) is a sum of 14 items that contains three subscale scores—Aggressive Behavior (**RnBAGGR**, four items), Withdrawn Behavior (**RnBWITH**, six items), and Hyperactive Behavior (**RnBHYPE2**, three items). The number of items in the three subscales add up to 13. One item not included in the subscales is included in the total score for problem behaviors. Therefore, there are 14 items in the total score for problem behaviors. The items come from the Behavior Problems Index (Peterson and Zill 1986), the Personal Maturity Scale (Entwisle et al. 1997), and the Social Skills Rating Scale (Gresham and Elliott 1990). Higher scores indicate the child exhibits negative behavior more frequently.
- Teacher-reported approaches to learning (**RnAPROHR**) is a mean (average) of six items that make up the Approaches to Learning Scale from the ECLS–K Approaches to Learning Scale (U.S. Department of Education 2002). Higher scores indicate the child exhibits positive approaches to learning behaviors more frequently.
- Teacher-reported child literacy behaviors (**RnTPELS**) is a sum of five items about children's early literacy skills. These items are adapted from the National Household Education Survey (U.S. Department of Education 2000). Higher scores indicate the child exhibits literacy skills more frequently.

We also constructed teacher-reported indicators for whether a child had a diagnosed disability (**RnDISB2**) and, if so, the nature of the disability (**RnCOGIMP**, **RnLNGIMP**, **RnPHYIMP**, **RnBEHIMP**, **RnSNSIMP**), whether the child had multiple impairments (**RnMLIMP2**), and whether the child had an IEP or IFSP (**RnCHIEP**).

We calculated composite scores about children's behaviors from the parent survey as follows:

- Parent-reported approaches to learning (**PnAPROHR**) is a mean (average) of six items similar in content to those for teachers using the teacher version of the Approaches to Learning Scale from the ECLS–K (U.S. Department of Education 2002). The 2021–2022 study team adapted the language of the items for use with parents and caregivers in the home setting to assess a child's motivation, attention, organization, persistence, and independence in learning. Higher scores indicate the child exhibits positive approaches to learning behaviors more often.

[53] Raw scores are an indicator of absolute performance based on the items the child received. They are calculated as a sum or items or a mean of items, depending on the type of assessment. Raw scores are used to calculate other scores such as standard scores.

- Parent report of changes in child behaviors since March 2020 (**PnBEHCHG**) is the sum of parent-reported behaviors since this time; the binary variable (**PnBANYBEHCHG**) indicates whether at least one parent reported that the child experienced any behavior changes since this time.

B. Composite variables

We constructed four additional groups of variables linked to the logic model (Exhibit I.2): (1) child and family characteristics, including demographic characteristics of children, mothers, and fathers; household income; household composition and characteristics; and family economic risk index; (2) family processes and parenting, including family-child activities, parental mental health, parent-child relationship, social support, household food security, household financial strain, and household material hardship; (3) characteristics of Head Start teachers and classrooms, including class size and ratios, main curriculum used and alignment of curriculum and assessment, classroom language environment, and teacher characteristics; and (4) characteristics of Head Start programs and centers, including teacher turnover, center language environment, Head Start year length, program schedules, staff compensation and benefits, and center director and program director mental health. In the sections that follow, we describe the development of a subset of composite variables, highlighting those that may be commonly used in analyses or useful for providing context to findings.

1. Child and family characteristics

We developed several composite variables for describing child and family characteristics. In Exhibit VII.2, we indicate the instrument, name, label, survey items, and values associated with each variable. We briefly describe the development of the composites for child characteristics, household income, household composition and characteristics, and family economic risk index.

a. Child characteristics

We created several composite variables to describe child gender, race and ethnicity, age, the primary language spoken to the child, and whether the child was newly entering Head Start versus returning for a second year or previously participated in Early Head Start. The composites describing child gender, race and ethnicity, age, and primary language spoken to the child are based on fall 2021 data, and, in some instances, spring 2022 data (if fall data were missing).

We derived the child gender[54] composite (**CHGENDER2**) from the gender recorded on the parent consent form. If such information was not available, we used the gender indicated in the parent survey (PnA01). If both parent-reported items were missing, we used the child sex recorded from the child sampling form. It is accompanied by a flag identifying children whose spring 2022 data were used for the composite, **CHGENDER_F**.

We derived the child race and ethnicity composite (**CRACE**) from items PnA03 and PnA05 in the parent survey. This categorial variable describes whether parents indicated that the child belongs to one or more race category and is ethnically Hispanic. The categories include (1) White, non-Hispanic; (2) Black or African American non-Hispanic; (3) Hispanic, Latino/a, or Chicano/a/x; (4) American Indian or Alaska Native, non-Hispanic; (5) Asian or Pacific Islander, non-Hispanic; (6) Multi-racial/bi-racial, non-

[54] We updated the previous composite used in FACES 2019 and earlier, child sex (CHSEX), to reflect the child's gender identity rather than sex.

Hispanic; and (7) Other race. The variable was missing if the race and ethnicity source variables (PnA03, PnA05) were missing. It is accompanied by a flag identifying children whose spring 2022 data[55] were used for the composite, **CRACE_F**.

The child age composite (**P1RCAGE**) reflects the child's age as of September 1, 2021. We derived the child age composite from the date of birth recorded on the parent consent form or child sampling form. If such information was not available, we used the date of birth indicated in the parent survey (PnA02). It is accompanied by a flag identifying children whose spring 2022 data were used for the composite, **P1FCAGE**.

We derived the composite describing the primary language spoken to the child in the home (**P1RHHLNG**) from items PnD07 and PnD10 in the parent survey. It is accompanied by a flag identifying children whose spring 2022 data were used for the composite, **P1FHHLNG**.

We derived the composites describing whether the child was newly entering Head Start (**NewToHSR**) and whether they had participated in Early Head Start (**EHS_Participant**) from center records. All centers provided information on each child's enrollment date. If information about the child's participation in Early Head Start could not be obtained from center records, we set **EHS_Participant** to missing.

b. Household income (PnINCOME) and poverty (P1POVRTY and P1POVRTO)

We derived the household income composite from items PnM03amt, PnM03per, PnM04, PnM05, and PnM06 in the fall 2021 parent survey. To convert all responses to annual income, we created a continuous income variable (**CINCOME**) from these survey items by multiplying the amount (PnM03amt) by the appropriate factor based on the period (P1M03per). We used values of this continuous income variable to categorize cases for **PnINCOME**, the categories for which are provided in Exhibit VII.2.

In fall 2021, when **PnINCOME** could not be constructed because of out-of-range or missing values, we imputed the continuous income variable by employing multiple imputation by chained equations (MICE) (van Buuren & Groothuis-Oudshoorn, 2011) that uses a prediction model with many predictors.[56] Using R software with the fall 2021 data, MICE produces several imputed data sets. We chose MICE because it (1) incorporates a large number of important variables in the model, (2) appropriately models (given the predictors) the regression of the imputed variable depending on the variable's type, and (3) provides stable imputed data by performing imputation iteratively.[57]

[55] If the parent did not complete a child survey in fall 2021, spring 2022 data (collected at first interview) are used.
[56] Candidate covariates used in the R mice package for imputation of the continuous income variable are those variables that are statistically significant ($\alpha = 0.05$) in generalized linear models. Candidate covariates include presence of the biological mother and father in the household, family structure, father's and mother's employment status, father's and mother's age, mother's race and ethnicity, father's and mother's highest education, receipt of several types of public assistance, receipt of particular assistance (for example, welfare/Temporary Assistance for Needy Families [TANF], unemployment insurance, food stamps, WIC, child support, Supplemental Security Income [SSI]/Social Security retirement, payments for foster care, energy assistance), parent marital status, and whether an adult in the household reduced or skipped meals. We retained only the five predictors most strongly associated with income in the final imputation model.
[57] In general, another important consideration in using several imputations is that they provide better variance estimation if the correct variance formula based on multiple imputations is implemented in the data analysis. In our application here, however, we retain only one set of imputed data.

Starting with the variable that has the fewest number of missing values, the imputation process begins by imputing all covariates with missing values. The sequential process continues and updates itself every time a variable is imputed from the previous step. It then continues until all missing variables have been imputed and a complete nonmissing data set is created. We implemented MICE for 10 iterations in performing imputation five times. The 2021–2022 study used the first imputed data set value to construct **PnINCOME**. Although we ended up with five imputed data sets, our imputation strategy is not a strict multiple imputation method—**it is more of a hybrid approach**. This is particularly true for the current round of FACES data, because the majority of cases missing the continuous income variable (CINCOME) were imputed at the second stage using a single imputation method (unbounded). Therefore, the analytical framework for MI does not directly apply to the current strategy. After reviewing the five distributions of CINCOME in the five replicate files, we found them to have similar distributions, so we recommend that researchers randomly choose one imputed data set for analysis.

We carried out imputation of income in two steps. First, when categorical income was available, we constrained the imputed values to that income range. Second, when categorical income was missing, we did not constrain the imputed values to any income range; instead, we let the imputed values be dependent on empirical distribution of the data, treating the imputed values from the first step as if they were true values from the parent survey. A flag (**PnINCIMP**) identifies children whose **PnINCOME** was based on imputed values.

P1POVRTY is a dichotomous variable identifying a family whose income is below the poverty line based on federal poverty thresholds for 2020.[58] Poverty status is based on a household's income relative to the number of household members. We derived the poverty status composite using item M03 in the parent survey, a continuous variable that represents annual household income, and the total number of household members (**P1HHSIZE**).

We also created a categorical variable describing the ratio of income to poverty threshold (**P1POVRTO**) based on how far the household income falls below or above the federal poverty threshold, in conjunction with household size. **P1POVRTO** was created using items P1M03 and P1J21 in the parent survey, a continuous variable that represents annual household income, **P1HHSIZE**, and **P1POVRTY**.

c. Household characteristics and composition

We derived the mother's employment status composite (**PnMOMEMP, PnMOMEMP2** for households with a second mother present) from items PnJ17, PnJ18, PnJ19, and PnJ21 in the parent survey, and the father's employment status composite (**PnDADEMP, PnDADEMP2** for households with a second father present) from items PnK17, PnK18, PnK19, and PnK21 in the parent survey. This categorical variable describes whether the mother and father are working full or part time, looking for work, or not in the labor force. It includes all mothers and fathers, regardless of whether they live with the child.

We derived the mother's highest education composite (**P1RMOMED, P1RMOMED2** for households with a second mother present**)** from item PnJ24 in the parent survey and the father's highest education composite (**P1RDADED, P1RDADED2** for households with a second father present) from PnK24 in the parent survey. All are categorical variables with four classes of highest education: (1) less than high school diploma; (2) high school diploma or GED; (3) vocational/technical diploma, associate's degree, or

[58] See https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html for the annual thresholds. As an example, the poverty threshold for a family of four in 2020 was $26,496.

some college; and (4) bachelor's degree or higher. These composites were derived using fall 2021 data and, in some instances, spring 2022 data (if fall data were missing). Each of these composites are accompanied by a flag identifying mothers and fathers whose spring 2022 data were used for the composite, **P1FMOMED**, **P1RMOMED2**, **P1RDADED**, and **P1FDADED2**.

We derived the household size composite (**PnHHSIZE**) to describe the total number of household members, as reported by parents. **PnHHSIZE** is created by counting the number of responses to PnB05 in the parent survey and equals the number of rows created in the household roster.

We derived the family structure composites (**PnFMSTRC**, **PnFMSTRC2**) by using items Pn_SC9, PnB05, PnB05a1, and PnB05a2 from the parent survey. We derived categories from survey items to indicate who was present in the household. We used the values on these variables and PnB05, PnB05a1, and PnB05a2 to categorize the **PnFMSTRC** variable. We classified as "other" any family structures that did not include the presence of biological or adoptive parents, stepparents, or biological or adoptive grandparents. The **PnFMSTRC2** variable derives categories from the same survey items as **PnFMSTRC**, but it considers household composition in a gender-neutral way to be more inclusive of a variety of family structures.

We derived the parent marital status composite (**P1RMARTL**) from items PnB09 and PnJ15 in the parent survey and the family structure composite (**PnFMSTRC**). We created a categorical variable from the survey items to indicate the parents' marital status (not on the data file). We used the values on parent marital status and **PnFMSTRC** to create the **P1RMARTL** variable. **P1RMARTL** was derived using fall 2021 data and, in some instances, spring 2022 data (if fall data were missing). It is accompanied by a flag identifying mothers and fathers whose spring 2022 data were used for the composite, **P1FMARTL**.

d. Family economic risk index (P1ECRISK)

The family economic risk index provides information on the cumulative risks of children's families. Children whose families had higher risks are considered more disadvantaged than those with fewer risks. The family economic risk index (**P1ECRISK**) includes in its calculation the number of economic risks experienced by the family, including single parenthood (based on **P1FMSTRC**), low maternal education (based on **P1RMOMED**), and household poverty (based on **P1POVRTY**).

4. **Single parent household.** We defined a dichotomous variable based on **P1FMSTRC** (discussed above) to identify children whose family structure is either a single mother or father (**P1FMSTRC** = 2 or 3).
5. **Mother's education less than high school diploma.** Using **P1RMOMED** (discussed above), we created a dichotomous variable to identify children whose mothers do not have a high school diploma (**P1RMOMED** = 1).
6. **Household income below federal poverty threshold.** We used the dichotomous **P1POVRTY** variable (discussed above) to identify whether a family's income is below the federal poverty line.
7. We calculated the risk index by summing the above three dichotomous variables. If any variable was missing, we coded **P1ECRISK** as missing.

Exhibit VII.2. 2021–2022 study composite variables—child and family characteristics

Instrument	Variable name	Variable label	Instrument/survey items	Value Labels[a]
Child's demographic characteristics				
Parent survey	CRACE	Child race/ethnicity	PnA03, PnA05	White, non-Hispanic; Black or African American, non-Hispanic; Hispanic, Latino/a/, or Chicano/a/o; American Indian or Alaska Native, non-Hispanic; Asian or Pacific Islander, non-Hispanic; Multi-racial/bi-racial, non-Hispanic; Other race, non-Hispanic
Child sampling form, parent survey	CHGENDER2	Child gender	Child sampling form, otherwise PnA01	Male; female; another gender identity
Child sampling form, parent survey	P1RCAGE	Child age	Child sampling form, otherwise PnA02	n.a.
Child sampling form	NewToHSR	Newly entering Head Start	Child sampling form	Yes; No
Child sampling form	EHS_Participant	Participated in Early Head Start	Child sampling form	Yes; No
Child sampling form	HS_Fund	Child funded by Head Start	Child sampling form	Yes; No
Child sampling form	PK_Fund	Child funded by state or local pre-K	Child sampling form	Yes; No
Child sampling form[b]	CS_Fund	Child funded by child care subsidies	Child sampling form	Yes; No
Child sampling form	O_Fund	Funded by other source(s)	Child sampling form	Yes; No
Child sampling form	No_Fund	Child funding information unknown	Child sampling form	Yes; No
Mother's demographic characteristics				
Parent survey	P1RMAGE, P1RMAGE2[b]	Mother's age	PnSC7, PnSC9, PnJ08, PnB04, PnB05	Continuous

Exhibit VII.2. (*continued*)

Instrument	Variable name	Variable label	Instrument/survey items	Value Labels[a]
Parent survey	P1MBrnUS, P1MBrnUS2[b]	Whether mother was born in the U.S.	P1J13	Born in U.S.; not born in U.S.
Parent survey	P1MYrsUS, P1MYrsUS2[b]	Years in U.S. if born elsewhere for mother	P1J14	< = 5 years; 6 to 10 years; > 10 years
Parent survey	P1MOrign, P1MOrign2[b]	Country/area of birth for mother if born elsewhere	P1J13	Mexico; Central America; South America; Caribbean; Africa, Asia; Other
Parent survey	P1MOMEMP, P2MOMEMP, P1MOMEMP2[b], P2MOMEMP2[b]	Mother's employment status	PnJ17–PnJ19, PnJ21	Working full-time; working part-time; looking for work; not in labor force
Parent survey	P1RMOMED, P1RMOMED2[b]	Mother's highest education	PnJ24	Less than high school diploma; high school diploma or GED; vocational/technical diploma, associate's degree, or some college; bachelor's degree or higher
Parent survey	P1MEMPCV, P1MEMPCV2[b]	Change in mother's employment due to COVID-19	P1J21a-P1J21g	At least one change in mother's employment; No changes in mother's employment
Parent survey	P1MEMPLS, P1MEMPLS2[b]	Mother working less hours, has lost job, or has been furloughed	P1J21c, P1J21f	Yes; No
Parent survey	P1MEMPMR, P1MEMPMR2[b]	Mother working more hours or more jobs	P1J21b, P1J21d	Yes; No
Father's demographic characteristics				
Parent survey	P1RFAGE, P1RFAGE2[c]	Father's age	PnSC7, PnSC9, PnB04, PnB05, PnK08	n.a.
Parent survey	P1FBrnUS, P1FBrnUS2[c]	Whether father was born in the U.S.	P1K13	Born in U.S.; not born in U.S.
Parent survey	P1FYrsUS, P1FYrsUS2[c]	Years in U.S. if born elsewhere for father	P1K14	< = 5 years; 6 to 10 years; > 10 years
Parent survey	P1FOrign, P1FOrign2[c]	Country/area of birth for father if born elsewhere	P1K13	Mexico; Central America; South America; Caribbean; Africa, Asia; Other
Parent survey	P1DADEMP, P2DADEMP,	Father's employment status	PnK17–PnK19, PnK21	Working full time; working part-time; looking for work; not in labor force

Exhibit VII.2. (*continued*)

Instrument	Variable name	Variable label	Instrument/survey items	Value Labels[a]
	P1DADEMP2[c], P2DADEMP2[c]			
Parent survey	P1RDADED, P1RDADED2[c]	Father's highest education	PnK24	Less than high school diploma; high school diploma or GED; vocational/technical diploma, associate's degree, or some college; bachelor's degree or higher
Parent survey	P1DEMPCV, P1DEMPCV2[c]	Change in father's employment due to COVID-19	P1K21a-P1K21g	At least one change in father's employment; no changes in father's employment
Parent survey	P1DEMPLS, P1DEMPLS2[c]	Father working less hours, has lost job, or has been furloughed	P1K21c, P1K21f	Yes; No
Parent survey	P1DEMPMR, P1DEMPMR2[c]	Father working more hours or more jobs	P1K21b, P1K21d	Yes; No
Household composition and characteristics				
Parent survey	P1HHSIZE, P2HHSIZE	Household size	PnB05	A count of the number of rows for this item
Parent survey	P1FMSTRC, P2FMSTRC	Family structure	PnSC9, PnB05	Bio/adoptive mother and bio/adoptive father; Bio/adoptive mother only; bio/adoptive father only; non-biological mother and non-biological father; non-biological mother and bio/adoptive father; bio/adoptive mother and non-biological father; bio/adoptive grandparent(s); other
Parent survey	P1FMSTRC2, P2FMSTRC2	Family structure (non-gendered version)	PnSC9, PnB05	Two bio/adoptive parents; one bio/adoptive parent; one bio/adoptive parent and one non-bio/adoptive parent; two non-bio/adoptive parents; bio/adoptive grandparent (s); other
Parent survey	P1RMARTL	Parent marital status	PnFMSTRC, PnB09, PnJ15	Married; divorced; separated; registered civil union/domestic partnership; living as partners in a committed relationship; not married; not a two-parent household
Parent survey	P1SINGLE	Single-parent household	P1FMSTRC	Single-parent household; not a single-parent household
Parent survey	P1PBRNUS	Whether both parents were born in the U.S.	P1MBrnUS and P1FBrnUS	Both parents born in U.S.; one parent born outside U.S.; both parents born outside U.S.
Parent survey	P1RHHLNG	Primary language spoken to child	PnD07, PnD10	Non-English; English

Exhibit VII.2. (*continued*)

Instrument	Variable name	Variable label	Instrument/survey items	Value Labels[a]
Parent survey	PnHSNGOT	Parent lives in housing shared with one or more families or in transitional housing or a homeless shelter	PnM07	Yes; no
Household income				
Parent survey	P1ASSIST, P2ASSIST	Multiple public assistance	PnM01a–PnM01i	Receives multiple public assistance; does not receive multiple public assistance
Parent survey	P1POVRTY	Poverty status	P1M03, P1HHSIZE, P1INCOME	Below poverty threshold; at or above poverty threshold
Parent survey	P1POVRTO	Ratio of income to Poverty threshold	P1M03, P1J21, P1HHSIZE, P1INCOME , P1POVRTY	Below 50% of poverty threshold; 50–100% of poverty threshold; 101–130% of poverty threshold; 131–185% of poverty threshold; 186–200% of poverty threshold; at or above 200% of poverty threshold
Parent survey	P1INCOME, P2INCOME	Household Income	PnM03amt, PnM03per, PnM04, PnM05, and PnM06	$5,000 or less; $5,001 to $10,000; $10,001 to $15,000; $15,001 to $20,000; $20,001 to 25,000; $25,001 to $30,000; $30,001 to $35,000; $35,001 to $40,000; $40,001 to $50,000; $50,001 to $75,000; or more than $75,000
Parent survey	P1ECRISK	Family economic risk index	P1FMSTRC, P1POVRTY, P1RMOMED	0–3

Note: The “n” used in a variable name is a placeholder for the wave number, with “1” referring to fall 2021 and “2” to spring 2022. If a particular variable is listed with the “n” in the text, it was constructed in more than one wave.

SMS = Sample management system.

[a] In this column, we note the potential response range that could be obtained on the variable across all available data collection waves. For the data values associated with these labels, see Appendix G: Fall 2021 – Spring 2022 Child Level File Codebook.

[b] Variable created for households in which there is a second mother present.

[c] Variable created for households in which there is a second father present.

n.a. = not applicable.

2. Family processes and parenting

Next we describe the development of composite variables for family-child activities, parent mental health, parent-child relationship, social support, household food security, household financial strain, and household material hardship. In Exhibit VII.3, we note the instrument, variable name, variable label, source survey items, values and possible response ranges, and reliability (Cronbach's alpha) for these and other family processes and parenting variables.

a. Family-child activities

We created several composites to describe parent-child activities. We derived the composite describing whether the parent or someone in the family had read to the child three or more times in the past week (**PnREADS**) by dichotomizing item PnD01 in the parent survey ("yes" if parent reports either three times a day or every day).

b. Parent mental health

We calculated the parent's depressive symptoms score (**PnDEPSCO**) from responses to 12 items (PnU01a through PnU01l) on the parent survey (drawn from the short form of the CES–D). The CES-D is a screening tool and not a diagnostic tool, but scores have been correlated with clinical diagnosis (Radloff 1977). Items used a 4-point scale and ranged in value from 1 ("rarely or never") to 4 ("most or all"). First, we recoded the 12 items so responses ranged from zero to 3 instead of 1 to 4. Second, we calculated **PnDEPSCO** from the sum of the recoded interview items. If any one of the items was missing, we calculated **PnDEPSCO** by multiplying the average of the other 11 items by 12. If any two items were missing, we calculated **PnDEPSCO** by multiplying the average of the other 10 items by 12. If three or more items were missing, we coded **PnDEPSCO** as missing. Continuous depressive symptoms scores range from zero to 36.

We used values from the continuous score (**PnDEPSCO**) to categorize the level of depressive symptoms (**PnDEPCAT**). We categorized cases with values of **PnDEPSCO** from zero to 4 as "no to few symptoms" (**PnDEPCAT** = 1); from 5 to 9 as "mild" (**PnDEPCAT** = 2); from 10 to 14 as "moderate" (**PnDEPCAT** = 3); and 15 or higher as "severe" (**PnDEPCAT** = 4). We coded any cases for which **PnDEPSCO** was missing as missing on **PnDEPCAT**.

We calculated the parent's anxiety level score (**PnAnxSco**) from responses to seven items (PnU02a through PnU02g) on the parent survey (drawn from the GAD–7). The GAD–7 is a valid and reliable tool to screen for anxiety (Spitzer et al. 2006). Items used a 4-point scale and ranged in value from zero ("not at all") to 3 ("nearly every day"). We calculated **PnAnxSco** from the sum of these items. If any one of the items was missing, we calculated **PnAnxSco** by multiplying the average of the other six items by seven. If any two items were missing, we calculated **PnAnxSco** by multiplying the average of the other five items by seven. If 3 or more items were missing, we coded **PnAnxSco** as missing. Continuous anxiety symptoms scores range from zero to 21.

We used values from the continuous score (**PnAnxSco**) to categorize the level of anxiety symptoms (**PnAnxCat**) (Spitzer et al. 2006). We categorized cases with values of **PnAnxSco** from zero to 4 as "minimal" (**PnAnxCat** = 1); from 5 to 9 as "mild" (**PnAnxCat** = 2); from 10 to 14 as "moderate" (**PnAnxCat** = 3); and 15 or higher as "severe" (**PnAnxCat** = 4). We coded any cases for which **PnAnxSco** was missing as missing on **PnAnxCat**.

Exhibit VII.3. 2021–2022 study composite variables—family processes and parenting

Instrument	Variable name	Variable label	Instrument/survey items	Value labels[a] and possible response ranges	2021–2022 study reliability (Cronbach's alpha) fall 21, spring 22
Family-child activities					
Parent survey	P1READS, P2READS	Read to child 3+ times in past week	PnD01	Yes; No	n.a.
Child care during COVID-19					
Parent survey	P1CCVrtl, P2CCVrtl	Children have at least some virtual or remote instruction	PnN26	Virtual or remote instruction only/hybrid of in-person instruction and virtual or remote instruction; in-person instruction only/homeschooled	n.a.
Parent survey	P1VrtlWk, P2VrtlWk	Assists child with online learning during working hours	PnN28, PnN29	Assists during work hours; assists outside of working hours; does not assist	n.a.
Parent survey	P1CCNdCt, P2CCNdCt	Count of strategies used to meet child care needs outside of regular arrangements	PnN33a–PnN33f	0–6	n.a.
Parent survey	P1CCNeed, P2CCNeed	One or more strategies used to meet child care needs outside of regular arrangements	PnN33a–PnN33f	Yes; No	n.a.
Health and health care access					
Parent survey	P1ERReg, P2ERReg	Goes to emergency room for routine medical care	PnP05	Yes; No	n.a.

Exhibit VII.3. (*continued*)

Instrument	Variable name	Variable label	Instrument/survey items	Value labels[a] and possible response ranges	2021–2022 study reliability (Cronbach's alpha) fall 21, spring 22
Parent survey	P1COVHHF	Someone in household or a close friend or family member had COVID-19	P1P49	Yes; No	n.a.
Parent survey	P1COVHH	Someone in household had COVID-19	P1P49	Yes; No	n.a.
Parent survey	P1COVRsp	Respondent had COVID-19	P1P49a	Yes; No	n.a.
Parent survey	P1COVChd	Child had COVID-19	P1P49a	Yes; No	n.a.
Parent survey	P1COVHsp	Someone in household or a close friend or family member has been hospitalized due to COVID-19	P1P50	Yes; No	n.a.
Parent survey	P1CVHspH	Someone in household has been hospitalized due to COVID-19	P1P50	Yes; No	n.a.
Parent survey	P1CVHspC	Child has been hospitalized due to COVID-19	P1P50	Yes; No	n.a.
Parent survey	P1COVDied	Someone in household or a close friend or family member died of COVID-19	P1P51	Yes; No	n.a.

Exhibit VII.3. (*continued*)

Instrument	Variable name	Variable label	Instrument/survey items	Value labels[a] and possible response ranges	2021–2022 study reliability (Cronbach's alpha) fall 21, spring 22
Parent survey	P1CVDiedH	Someone in household died of COVID-19	P1P51	Yes; No	n.a.
Parent mental health					
Parent survey	P1DepSco, P2DepSco	Parent's depressive symptoms, CES-D short form	PnU01a–PnU01l	0–36	0.88, 0.88
Parent survey	P1DepCat, P2DepCat	Parent's depressive score, CES-D short form, categories	PnDepSco	Not depressed; mildly depressed; moderately depressed; severely depressed	n.a.
Parent survey	P1AnxSco, P2AnxSco	Parent anxiety level score, GAD-7	PnU02a–PnU02g	0–21	0.9, 0.91
Parent survey	P1AnxCat, P2AnxCat	Parent anxiety level category, GAD-7	PnAnxSco	No anxiety; mild anxiety; moderate anxiety; severe anxiety	n.a.
Parent survey	P1AnxInc, P2AnxInc	Parent stress and anxiety increased compared to pre-COVID-19	PnU03	Yes; No	n.a.
Parent survey	P1ParBeh, P2ParBeh	Parenting stress and behavior	PnU04a–PnU04f	1–5	n.a.
Parent-child relationship					
Parent survey	P1ParWrm, P2ParWrm	Parenting warmth score	PnG4a-PnG4f	1–4	n.a.
Social support					
Parent survey	P1HelpAl	Number of types of help parent can always get	T01a-TI01h	0–6	n.a.

Exhibit VII.3. (*continued*)

Instrument	Variable name	Variable label	Instrument/survey items	Value labels[a] and possible response ranges	2021–2022 study reliability (Cronbach's alpha) fall 21, spring 22
Family engagement					
Parent survey	P2Cultre	Strengths-Based Practices Inventory cultural competency score	P2W05g, P2W05k, and P2W05m	1–5	n.a.
Household food security					
Parent survey	P1FDntLst	Food did not last, no money for more	P1M10a	Often/sometimes true; never true	n.a.
Parent survey	P1NoBalM	Could not afford balanced meals	P1M10b	Often/sometimes true; never true	n.a.
Parent survey	P1CtMls	Cut or skipped meals	P1M11	Yes; No	n.a.
Parent survey	P1FCtMls	Frequency of cut or skipped meals	P1M12	Often/sometimes true; never true	n.a.
Parent survey	P1AteLss	Ate less than should	P1M13	Yes; No	n.a.
Parent survey	P1Hungry	Hungry because could not afford food	P1M14	Yes; No	n.a.
Parent survey	P1FdSec	Total food security status level	P1FDntLst, P1NoBalM, P1CtMls, P1FCtMls, P1AteLss, P1Hungry	0–6	n.a.
Parent survey	P1FdSCat	Total food security status level categories	P1FdSec	High food security; marginal food security; low food security; very low food security	n.a.
Household financial strain					
Parent survey	P1MtNd	Material needs	P1M15a–P1M15d	Reported a significant financial strain; did not report a significant financial strain	n.a.

Exhibit VII.3. (*continued*)

Instrument	Variable name	Variable label	Instrument/survey items	Value labels[a] and possible response ranges	2021–2022 study reliability (Cronbach's alpha) fall 21, spring 22
Parent survey	P1MtNdM	Material needs–mean	P1M15a–P1M15d	1–5	n.a.
Parent survey	P1MtNdS	Material needs–sum	P1M15a–P1M15d	4–20	n.a.
Parent survey	P1NOMTNM	Can't make ends meet–mean	P1M16, P1M17	1–5	n.a.
Parent survey	P1NOMTNS	Can't make ends meet–sum	P1M16, P1M17	1–10	n.a.
Housing, utilities, and medical hardships					
Parent survey	P1HSec	Count of housing insecurity	PnM18d, PnM18e	0, 1, 2	n.a.
Parent survey	P1UtSec	Count of lack of basic utilities	PnM18c, PnM18f, PnM18g	0, 1, 2, 3	n.a.
Parent survey	P1MedSec	Count of unmet medical needs	PnM18a, PnM18b	0, 1, 2	n.a.
Parent survey	P1RRmRto	Ratio of people to rooms in the household	PnM9c, PnHHSIZE	Continuous	n.a.

Note: The "n" used in a variable name is a placeholder for the wave number, with "1" referring to fall 2021 and "2" to spring 2022. If a particular variable is listed with the "n" in the text, it was constructed in more than one wave.

[a]Noted "value labels" reflect the actual response range obtained on the variable. For the data values associated with these labels, see Appendix G: Fall 2021 – Spring 2022 Child Level File Codebook.

CES-D = Center for Epidemiological Studies Depression Scale; GAD = General Anxiety Scale–7; n.a. = not applicable.

c. Parent-child relationship

We created a measure of the warmth of the parent-child relationship by using six items from the Early Childhood Longitudinal Study, Kindergarten Class of 2022–23 2020 field test preschool parent survey (U.S. Department of Education 2020). We calculated a parenting warmth score (**PnParWrm**) from responses to items PnG04a, PnG04b, PnG04c, PnG04d, PnG04e, and PnG04f in the parent survey. We first reverse coded PnG04e and PnG04f so that all items had high values (ranging from 1 to 4), indicating a positive parent-child relationship. We then calculated the mean of the six items to create a parenting warmth score.

d. Social support

We derived the composite variable describing the number of types of help parents can always get (**P1HelpAl**) using items P1T01a, P1T01b, P1T01c, P1T01g, P1T01e, and P1T01h in the parent survey. For each item, parents indicated whether help was received never, sometimes, or always. When the parent reported having always received a specific type of help, we counted that help toward the total number of types of help the parent can always get. When the parent reported having received a specific type of help sometimes or never, we did not count that help toward the total number of types of help the parent can always get.

e. Household food security

We created food security scale scores by using the U.S. Department of Agriculture (USDA) "Guide to Measuring Household Food Security, Revised 2000" (Bickel et al. 2000) and USDA's 2006 updates to the security labels. We calculated the food security scale score (**P1FdSec**) from responses to six items (P1M10a, P1M10b, P1M11, P1M12, P1M13, and P1M14) on the parent survey. First, we recoded the six items into a series of dichotomous indicators[59] of whether food in the household ever did not last (**P1FDntLst**), the family could not afford balanced meals (**P1NoBalM**), the parent ever cut or skipped meals (**P1CtMls**), the parent sometimes or often cut or skipped meals (**P1FCtMls**), the parent ate less than they should (**P1AteLss**), and the parent was ever hungry because they could not afford food (**P1Hungry**). Second, we calculated **P1FdSec** from the sum of these dichotomous variables. If one of the items was missing, we calculated **P1FdSec** by multiplying the average of the other five items by six. If two or more items were missing, we coded **P1FdSec** as missing. Continuous food security scale scores ranged from zero to 6.

We used values from the continuous score (**P1FdSec**) to categorize the level of food security (**P1FdSCat**). We categorized cases with a **P1FdSec** value of zero as "high food security" (**P1FdSCat** = 0) and a value of 1 as "marginal food security" (**P1FdCat** = 1). We categorized cases with **P1FdSec** values from 2 to 4 as "low food security" (**P1FdSCat** = 2); and from 5 to 6 as "very low food security" (**P1FdSCat** = 3). We coded any cases for which **P1FdSec** was missing as missing on **P1FdSCat**. High food security indicates no indications of food-access problems or limitations; marginal food security indicates minimal indications of food-access problems or limitations, suggesting little anxiety over food sufficiency or shortage of food in the house and little or no indication of changes in diet or food intake. Low food security indicates reports of reduced quality, variety, or desirability of diet but little or no

[59] The items included response categories denoting how frequently something was true for the family and whether statements were true for families (yes/no).

indication of reduced food intake. Very low food security points to reports of several indications of disrupted eating patterns and reduced food intake.

f. Household financial strain (P1MtNd)

We created financial strain scores by using four items from the Economic Strain Questionnaire (Conger et al. 1993) that measure the extent to which a parent feels they have enough money to afford the type of home (P1M15a), clothing (P1M15b), food (P1M15c), and medical care they need (P1M15d). To calculate the financial strain score (**P1MtNd**) and align it with earlier use of the same four items (for example, Raver et al. 2013), we categorized a family as having "reported a significant financial strain" (**P1MtNd** = 1) if the parent disagreed or strongly disagreed that they had enough money to afford any of the four items (home, clothing, food, or medical care). Otherwise, a family "has not reported a significant financial strain" (**P1MtNd** = 0). We coded any cases for which any of the parent survey source items was missing as missing on **P1MtNd**.

g. Household material hardship: Housing insecurity (P1HSec), lack of basic utilities (P1UtSec), and unmet medical needs (P1MedSec)

The 2021–2022 study includes a set of composite variables on household material hardship, which provides information on family housing, utilities, and medical hardships. To measure these hardships, we used parents items that measured whether anyone in the family had experienced a housing, utility, or medical hardship in the past 12 months (six items adapted from the Health Profession Opportunity Grants [HPOG] Impact Study based on the Survey of Income and Program Participation [SIPP] Adult Well-Being topical module [U.S. Bureau of the Census 2011] and one parallel, newly developed item). We categorized families' reported hardships as follows: (1) housing insecurity if the parent responded "yes" to either of the two housing insecurity items, (2) lack of basic utilities if the parent responded "yes" to any of the three basic utility items, and (3) unmet medical needs if the parent responded "yes" to either of the two medical need items. We then created three indexes reflecting the count of hardships experienced separately for housing insecurity (**P1HSec**), lack of basic utilities (**P1UtSec**), and unmet medical needs (**P1MedSec**).

3. Head Start teacher and classroom characteristics

Composite variables for describing Head Start teacher or classroom characteristics come from the teacher surveys conducted in fall 2021 and spring 2022. We describe the approach to constructing variables for class size and child/adult ratios and then explain a subset of composite variables from the teacher survey on the alignment of curriculum and assessment tools used in the classroom, whether a child's home language is used for instruction (this variable also uses information from the Head Start parent survey), and teacher characteristics. In Exhibit VII.4, we note the instrument, variable name, variable label, source survey items, values and possible response ranges, and data collection wave associated with each variable, along with reliability information (that is, internal consistency) on scales constructed from items included in the teacher survey.

a. Class size and ratios

We used the teacher survey to construct variables to describe Head Start class size and child/adult ratios.

We defined class size (**T2CSIZE**) as equal to T2A00_1. If T2A00_1 was missing, we coded **T2CSIZE** as missing. We derived two measures of Head Start class ratios: (1) child/teacher ratio (**T2CTRTIO**) and

(2) child/adult ratio (**T2CARTIO**). We derived the child/teacher ratio by dividing the number of children (T2A00_1) by the number of lead teachers and assistant teachers (T2A00_2 and T2A00_3) in the classroom.[60] We derived the child/adult ratio by dividing the number of children by the number of lead teachers, assistant teachers, and paid aides in the classroom (T2A00_2, T2A00_3, and T2A00_4). For both measures of class ratio, if any source item was missing or the ratio was below 1, we coded **T2CTRTIO** and **T2CARTIO** as missing.

b. Main curriculum used and alignment of curriculum and assessment

We constructed several variables to capture whether teachers use aligned curriculum and assessment tools. The first is a categorical variable that captures the main curriculum reported by teachers (**T2CRRCLA**). It draws on items T2A06, T2A07_XX,[61] and T2A08 from the teacher survey. The categories include (1) Creative Curriculum, (2) HighScope, (3) locally designed, (4) widely available curriculum, (5) other, and (6) uses multiple curricula equally. The variable was missing if teachers reported not using a curriculum (T2A06 = 3) or the curriculum source variables (T2A06, T2A07_XX, T2A08) were missing.

The second variable (**T2CRRAT**) captures whether teachers use a curriculum that has an available, aligned assessment tool; it is based on teachers' reports of their main curriculum (T2CRRCLA, T2A07_XX, T2A08). Curricula considered to have an aligned assessment include Creative Curriculum, which aligns with the Teaching Strategies Gold (GOLD) Assessment; HighScope, which aligns with the HighScope Child Observation Record (COR); Montessori, which we consider to have alignment when a program designs its own assessment; and Galileo, which now offers both a curriculum and an assessment. The variable was missing if the source variables were missing.

The third variable (**T2ACAT**) is a dichotomous variable that captures whether teachers— among those who use a curriculum with an aligned assessment available—use the assessment that aligns with their main curriculum as defined above. The variable is based on **T2CRRAT**, **T2CRRCLA**, and T2A21. It was missing if the teachers used a curriculum without an aligned assessment tool (**T2CRRAT** = 0) or no curriculum/curricula were reported in T2A07_XX or T2A08 (that is, **T2CRRAT** was missing).

The fourth variable (**T2ACAT2**) is a categorical variable that reflects whether teachers used an assessment that aligns with their main curriculum or a secondary curriculum. The variable is based on **T2CRRAT**, **T2CRRCLA**, T2A06, T2A07_XX, T2A08, and T2A21. Unlike **T2ACAT**, the variable is defined even if the curriculum does not have an available, aligned assessment tool. Categories include (1) aligned (Creative Curriculum + GOLD), (2) aligned (HighScope + COR), (3) aligned (Montessori), (4) aligned (Galileo), (5) not aligned (Creative + nonaligned assessment), (6) not aligned (HighScope + nonaligned assessment), (7) not aligned (Montessori), (8) not aligned (Galileo), (9) not aligned (uses curriculum with no aligned assessment available), and (10) not aligned (multiple curricula including Creative but no assessment). According to the curricula and assessments reported in the teacher survey, some of these categories do not include any observations. The variable was missing if no curriculum/curricula were reported in T2A07_XX or T2A08.

[60] A lead teacher is the head or primary teacher in the classroom. An assistant teacher is a teacher who supports the lead teacher in the classroom.

[61] Where respondents can select all that apply from a long list of items, we use "_XX" to denote the full list of response options for a given item.

We also created a variable to assess the number of hours of training teachers have received on the main curriculum (**T2CRRHT**) based on the number of curricula that teachers selected in T2A07 in the teacher survey (**T2CRRCT**).

We created the last set of variables to categorize the content of the curriculum. We constructed five dichotomous variables to reflect whether a teacher uses a comprehensive curriculum (**T2CRCOMP**), a math curriculum (**T2CRMATH**), a literacy curriculum (**T2CRLIT**), a social-emotional curriculum (**T2CRSOC**), or other curriculum (**T2CROTH**). We defined a curriculum as comprehensive if it covered a broad range of content. Comprehensive curricula include Creative Curriculum, Creating Child-Centered Classrooms – Step By Step, HighScope, Learn Every Day, Montessori, and Tools of the Mind. We categorized math curricula if the content was specific to math. Examples of math curricula include Building Blocks, DLM Early Childhood Express, Everyday Mathematics, Frog Street, Number Worlds, and Scholastic Curriculum. Examples of literacy curricula include DLM Early Childhood Express, Frog Street, Fundations, Let's Begin with the Letter People, Open Circle, Opening the World of Learning (OWL), Preschool PATHS (Promoting Alternative Thinking Strategies), Scholastic Curriculum, and Zoophonics. Social-emotional curricula include Preschool PATHS, Pyramid Model for Supporting Social Emotional Competence, and Second Step. Similarly, we categorized curricula as literacy curricula if the content was specific to literacy, such as Let's Begin with the Letter People, Scholastic Curriculum and Frog Street. Finally, we created an "other curriculum" category to encompass curricula not categorized as comprehensive, math, literacy, or social-emotional. For example, this category includes Handwriting without Tears, a curriculum that focuses on visual motor skills.

We also constructed six additional variables on curricula teachers were using to teach math that had a comprehensive (**T2MATHCO**), math (**T2MATHMA**), or other (**T2MATHOT**) curriculum type. Similarly, we constructed variables on literacy curriculum type that had a comprehensive (**T2LITCO**), literacy (**T2LITLI**), or other (**T2LITOT**) curriculum type. These dichotomous variables are based on curricula reported in the teacher survey (T2A07_XX) and type as defined in the prior paragraph.

c. Classroom and home language environments

We constructed 13 variables to capture the classroom language environment. We created five dichotomous variables to reflect whether a classroom includes DLL children (**T2DLLCls**), any children who speak English (**T2ChEng**), any who speak Spanish (**T2ChSpan**), any who speak Asian languages (**T2ChAs**), and any who speak other non-English languages (**T2ChOth**). These variables are based on the teacher survey item T2A03_FX.

We also constructed four variables on the language a teacher uses for classroom instruction, drawing from teacher survey item T2A04_XX. The variables include English (**T2InEng**), Spanish (**T2InSpan**), Asian languages (**T2InAs**), and other languages (**T2InOth**) used for classroom instruction.

The final four dichotomous variables we constructed were to categorize the staff using non-English languages in classrooms. The variables include lead teacher speaks non-English language in classroom (**T2LTSpk**), assistant teacher speaks non-English language in classroom (**T2ATSpk**), class aide speaks non-English language in classroom (**T2CASpk**), and volunteer/non-staff speaks non-English language in classroom (**T2VSpk)**. These dichotomous variables are based on teacher survey items T2A04A02_X to T2A04A15_X.

Even though we constructed several variables to capture the classroom language environment (Exhibit VII.4), in spring 2022, we constructed one variable to capture whether a child's home language (as

reported by parents) is used for classroom instruction (as reported by teachers) (**P2LngMch**). We constructed this dichotomous variable at the child level. The variable compares teacher reports of languages used for instruction in the classroom by the teacher or other adults (T2A04_XX) with parent reports of the language they usually speak to the child at home (P1D10 or P2D10). If source items were missing, we coded **P2LngMch** as missing.

d. Teacher characteristics

We briefly describe fall 2021 and spring 2022 constructed variables concerning teacher race and ethnicity, number of years teaching, developmentally appropriate attitudes, satisfaction, and mental health.

Race and ethnicity. We constructed **T2RACE** to describe Head Start teachers' race and ethnicity, drawing on items T2D21 and T2D23 from the teacher survey. The categories include (1) White, non-Hispanic; (2) Black or African American non-Hispanic; (3) Hispanic, Latino/a, or Chicano/a; (4) American Indian or Alaska Native, non-Hispanic; (5) Asian or Pacific Islander, non-Hispanic; (6) Multi-racial/bi-racial, non-Hispanic; and (7) Other race. The variable was missing if the race and ethnicity source variables (T2D21, T2D23) were missing.

Years teaching Head Start. We constructed **T2YRSHS** to describe the number of years teachers have taught Head Start, using the item T2D02 from the teacher survey. The four categories include less than 1 year, 1–2 years, 3–4 years, 5–9 years, and 10+ years of teaching Head Start. If the source item (T2D02) was missing, the variable was coded as missing.

Developmentally appropriate attitudes. We constructed a series of summary and subscale scores on teachers' developmentally appropriate attitudes. The 15 source items for the scores come from the Teacher Beliefs Scale (Burts et al. 1990) and consist of statements that reflect positive attitudes and knowledge of generally accepted practices in preschool settings or, conversely, a lack of these attitudes and knowledge. Four subscales are typically identified: Didactic, Child-initiated, Teach Letters, and Explicit Rewards. However, in FACES 2006, a factor analysis of the data resulted in a different factor-loading pattern across the 15 items. We constructed three variables based on the items identified by the FACES 2006 factor analysis: (1) **T2DASCO2** is a summary scale score of teachers' beliefs about developmentally appropriate practice, (2) **T2DISCO2** is the mean Didactic subscale, and (3) **T2INSCO2** is the mean Child-initiated subscale. For all scores, all source items use a 5-point Likert scale from 1 ("strongly disagree") to 5 ("strongly agree"), with negatively worded items reverse-scored. Consistent with earlier FACES studies, **T2DASCO2** scores start at a value of 1 and then increment by one point for certain responses to each of nine items, forming a sum score ranging from 1 to 10. If one item was missing, we calculated **T2DASCO2** by multiplying the sum of the other eight items by 9/8. If two items were missing, we calculated **T2DASCO2** by multiplying the sum of the other seven items by 9/7. If three or more items were missing, we coded **T2DASCO2** as missing. Meanwhile, **T2DISCO2** and **T2INSCO2** scores range from 1 to 5 and are calculated as the mean of relevant source items (five and six items, respectively). If any of the source items was missing, we coded **T2DISCO2** and **T2INSCO2** as missing. In Exhibit VII.4, we list the source items used to construct each subscale.

Teacher satisfaction. We constructed **T2SATISF**, which is a summary scale formed from three satisfaction items asked of Head Start teachers: how much teachers enjoy their present teaching job, how much they feel they are making a difference in the lives of children they teach, and whether they would

choose teaching again as a career. Scores range from 1 (strongly disagree) to 5 (strongly agree). The variable was coded as missing if one of the source items (T2B03a, T2B03b, T2B03c) was missing.

Teacher mental health. We constructed two variables (**T2DEPSCO** and **T2DEPCAT**) to describe Head Start teachers' depressive symptoms; the variables were similar to those constructed for parent depressive symptoms described above. We calculated Head Start teachers' depressive symptoms score (**T2DEPSCO**) from responses to 12 items (T2C01a through T2C01l) on the teacher survey, which were drawn from the short form of the CES–D. The construction of **T2DEPSCO** was equivalent to the construction of **PnDEPSCO**.

The categories identified in **T2DEPCAT** are equivalent to those in **PnDEPCAT**: values of zero to 4 on **T2DEPSCO** were categorized as "no to few symptoms" (**T2DEPCAT** = 1); values of 5 to 9 indicated "mild" (**T2DEPCAT** = 2); values of 10 to 14 indicated "moderate" (**T2DEPCAT** = 3); and values of 15 or higher indicated "severe" (**T2DEPCAT** = 4). We coded any cases for which **T2DEPSCO** was missing as missing on **T2DEPCAT**.

We calculated the teacher's anxiety level score (**TnAnxSco**) from responses to seven items (TnU02a through TnU02g) on the teacher survey drawn from the GAD–7, a valid and reliable tool to screen for anxiety (Spitzer et al. 2006). The items used a 4-point scale and ranged in value from zero ("not at all") to 3 ("nearly every day"). We calculated **TnAnxSco** from the sum of these items. If any one of the items was missing, we calculated **TnAnxSco** by multiplying the average of the other six items by seven. If any two items were missing, we calculated **TnAnxSco** by multiplying the average of the other five items by seven. If three or more items were missing, we coded **TnAnxSco** as missing. Continuous anxiety symptoms scores range from zero to 21.

We used values from the continuous score (**TnAnxSco**) to categorize the level of anxiety symptoms (**TnAnxCat**) (Spitzer et al. 2006). We categorized cases with values of **TnAnxSco** from zero to 4 as "minimal" (**TnAnxCat** = 1); from 5 to 9 as "mild" (**TnAnxCat** = 2); from 10 to 14 as "moderate" (**TnAnxCat** = 3); and 15 or higher as "severe" (**TnAnxCat** = 4). We coded any cases for which **TnAnxSco** was missing as missing on **TnAnxCat**.

Teacher job stress. We constructed **T2GnJbSt** by averaging four items (T2C13a through T2C13d) from the TCU SOF. The items (Institute of Behavioral Research 2005) provide information about teachers' level of agreement with statements about job-related stress at the center, and were rated on an agreement scale from 1 ("strongly disagree") to 5 ("strongly agree"). We multiplied the average by 10 to rescale the score (Knight et al. 2012; Lehman et al. 2002). Scores range from 10 to 50. The variable was coded as missing if two or more of the source items were missing.

Exhibit VII.4. 2021–2022 study composite variables—Head Start teachers and classroom characteristics

Instrument	Variable name	Variable label	Instrument/survey items	Value labels and possible response ranges	2021–2022 study reliability (Cronbach's alpha)
Classroom characteristics					
Child sampling form	CLS_HS_Fund[b]	Proportion of children funded by Head Start in classroom	Child sampling form	Continuous	n.a.
Teacher survey	T2CSIZE	Class size	T2A00_1	4–34[a]	n.a.
Teacher survey	T2CTRTIO	Child/teacher ratio	T2A00_1–T2A00_3	1.67–20[a]	n.a.
Teacher survey	T2CARTIO	Child/adult ratio	T2A00_1–T2A00_4	1.06–20[a]	n.a.
Teacher survey	T2CRRAT	Teacher uses curriculum with available aligned assessment tool	T2CRRCLA, T2A07_XX, T2A08	Yes; No	n.a.
Teacher survey	T2CRRCLA	Main classroom curriculum	T2A06, T2A07_XX, T2A08	Creative Curriculum; HighScope; locally designed; widely available; other; uses multiple curricula equally	n.a.
Teacher survey	T2ACAT	Teacher uses aligned curriculum and assessment tool	T2CRRAT, T2CRRCLA, T2A21	Yes; No	n.a.
Teacher survey	T2ACAT2	Teacher uses aligned curr and assmt, all teachers	T2CRRAT, T2CRRCLA, T2A06, T2A07_15, T2A07OthA, T2A07OthB, T2A08, T2A21	Aligned Creative Curriculum + GOLD; aligned High/Scope + COR; aligned Montessori; aligned Galileo; NOT aligned Creative + GOLD; NOT aligned High/Scope + COR; NOT aligned Montessori; NOT aligned Galileo; uses curriculum without aligned assessment tool; uses multiple curricula equally,	n.a.

Exhibit VII.4. (*continued*)

Instrument	Variable name	Variable label	Instrument/survey items	Value labels and possible response ranges	2021–2022 study reliability (Cronbach's alpha)
				including Creative (and Reggio) but no assessment	
Teacher survey	T2CRRCT	Count of curricula used	All T2A07_XX	1–7[a]	n.a.
Teacher survey	T2CRCOMP	Teacher uses a comprehensive curriculum	T2A07_11, T2A07_12, T2A07_15, T2A07_17, T2A07_30, T2A07_37	Yes; No	n.a.
Teacher survey	T2CRMATH	Teacher uses a math curriculum	T2A07_18, T2A07_24, T2A07_25, T2A07_26, T2A07_27, T2A07_31	Yes; No	n.a.
Teacher survey	T2CRLIT	Teacher uses a literacy curriculum	T2A07_14, T2A07_18, T2A07_24, T2A07_26, T2A07_28, T2A07_32, T2A07_33, T2A07_34, T2A07_38	Yes; No	n.a.
Teacher survey	T2CRSOC	Teacher uses a social-emotional curriculum	T2A07_34, T2A07_35, T2A07_36	Yes; No	n.a.
Teacher survey	T2CROTH	Teacher uses other curriculum	T2A07_29	Yes; No	n.a.
Teacher survey	T2MATHCO	Math curriculum type – comprehensive	T2A07a_11, T2A07a_12, T2A07a_15, T2A07a_17, T2A07a_30, T2A07a_37	Yes; No	n.a.
Teacher survey	T2MATHMA	Math curriculum type – math	T2A07a_18, T2A07a_24, T2A07a_25, T2A07a_26, T2A07a_27, T2A07a_31	Yes; No	n.a.
Teacher survey	T2MATHOT	Math curriculum type – other	T2A07a_14, T2A07a_18, T2A07a_24, T2A07a_26, T2A07a_28, T2A07a_29, T2A07a_32, T2A07a_33, T2A07a_34, T2A07a_35, T2A07a_36, T2A07a_38	Yes; No	n.a.
Teacher survey	T2LITCO	Literacy curriculum type – comprehensive	T2A07b_11, T2A07b_12, T2A07b_15, T2A07b_17, T2A07b_30, T2A07b_37	Yes; No	n.a.

Exhibit VII.4. (*continued*)

Instrument	Variable name	Variable label	Instrument/survey items	Value labels and possible response ranges	2021–2022 study reliability (Cronbach's alpha)
Teacher survey	T2LITLI	Literacy curriculum type – literacy	T2A07b_14, T2A07b_18, T2A07b_24, T2A07b_26, T2A07b_28, T2A07b_32, T2A07b_33, T2A07b_34, T2A07b_38	Yes; No	n.a.
Teacher survey	T2LITOT	Literacy curriculum type – other	T2A07b_18, T2A07b_24, T2A07b_25, T2A07b_26, T2A07b_27, T2A07b_29, T2A07b_31, T2A07b_34, T2A07b_35, T2A07b_36	Yes; No	n.a.
Teacher background characteristics					
Teacher survey	T1RACE, T2RACE	Teacher race	TnD21, TnD23	White, non-Hispanic; Black or African American, non-Hispanic; Hispanic, Latino/a/, or Chicano/a/o; American Indian or Alaska Native, non-Hispanic; Asian or Pacific Islander, non-Hispanic; Multi-racial/Bi-racial, non-Hispanic; Other race, non-Hispanic	n.a.
Teacher survey	T1YRSHS, T2YRSHS	Teacher's Years Teaching Head Start	TnD02	0 years, 1–2 years, 3–4 years, 5–9 years, 10+ years	n.a.
Classroom language environment					
Teacher survey	T2DLLCls	Classroom includes DLL children	T2A03_E	Yes; No	n.a.
Teacher survey	T2ChEng	Any children speak English	T2A03_F1	Yes; No	n.a.
Teacher survey	T2ChSpan	Any children speak Spanish	T2A03_F2	Yes; No	n.a.
Teacher survey	T2ChAs	Any children speak Asian languages	T2A03_F3, T2A03_F4, T2A03_F5, T2A03_F6,	Yes; No	n.a.

Exhibit VII.4. (*continued*)

Instrument	Variable name	Variable label	Instrument/survey items	Value labels and possible response ranges	2021–2022 study reliability (Cronbach's alpha)
			T2A03_F7, T2A03_F12, T2A03_F15		
Teacher survey	T2ChOth	Any children speak other non-English languages	T2A03_F8, T2A03_F9, T2A03_F11, T2A03_F13, T2A03_F14	Yes; No	n.a.
Teacher survey	T2InEng	Instruction occurs in English	T2A04_1	Yes; No	n.a.
Teacher survey	T2InSpan	Instruction occurs in Spanish	T2A04_2	Yes; No	n.a.
Teacher survey	T2InAs	Instruction occurs in Asian languages	T2A04_12, T2A04_4, T2A04_7, T2A04_15, T2A04_5, T2A04_6, T2A04_3	Yes; No	n.a.
Teacher survey	T2InOth	Instruction occurs in other non-English languages	T2A04_8, T2A04_9, T2A04_10, T2A04_11, T2A04_13, T2A04_14, T2A04_other	Yes; No	n.a.
Teacher survey, parent survey	P2LngMch	Child's home language used for classroom instruction	T2A04_01- T2A04_07 (teacher), T2A04_11- T2A04_15 (teacher), T2A04_31 (teacher), P1D10 or P2D10 (parent)	Yes; No	n.a.
Teacher survey	T2LTSpk	Lead teacher speaks non-English language in classroom	T2A04A02_1–T2A04A15_1	Yes; No	n.a.
Teacher survey	T2ATSpk	Assistant teacher speaks non-English language in classroom	T2A04A02_2–T2A04A15_2	Yes; No	n.a.
Teacher survey	T2CASpk	Class aide speaks non-English language in classroom	T2A04A02_3–T2A04A15_3	Yes; No	n.a.
Teacher survey	T2VSpk	Volunteer/non-staff speaks non-English language in classroom	T2A04A02_4–T2A04A15_4	Yes; No	n.a.
Teacher beliefs					
Teacher survey	T2DASCO2	Developmentally Appropriate Attitudes Scale	T2B06c, T2B06d, T2B06e, T2B06f, T2B06g, T2B06h, T2B06k, T2B6l, T2B06o	1–10	0.6
Teacher survey	T2DISCO2	Didactic Subscale	T2B06e, T2B06g, T2B06j, T2B06k, T2B06l, T2B06o	1–5	0.73

Exhibit VII.4. (*continued*)

Instrument	Variable name	Variable label	Instrument/survey items	Value labels and possible response ranges	2021–2022 study reliability (Cronbach's alpha)
Teacher survey	T2INSCO2	Child-Initiated Scale	T2B06a, T2B06c, T2B06d, T2B06f, T2B06i	1–5	0.7
Teacher survey	T2SATISF	Teacher Satisfaction Scale	T2B03a, T2B03b, T2B03c	1–5	0.83
Teacher mental health					
Teacher survey	T1DEPSCO, T2DEPSCO	Teacher's depressive symptoms, CES-D Short Form	TnC01a–TnC01l	0–36	0.9, 0.92
Teacher survey	T1DEPSCAT, T2DEPSCAT	Teacher's depressive score, CES-D Short Form, categories	TnDEPSCO	Not depressed; mildly depressed; moderately depressed; severely depressed	n.a.
Teacher survey	T1AnxSco, T2AnxSco	Teacher anxiety level score, GAD-7	TnC03a-TnC03g	0–21	0.93, 0.91
Teacher survey	T1AnxCat, T2AnxCat	Teacher anxiety level category, GAD-7	TnAnxSco	Minimal anxiety; mild anxiety; moderate anxiety; severe anxiety	n.a.
Teacher survey	T1CVJbSt, T2CVJbSt	Teacher COVID-19 job stress	TnC04a-TnC04d	1–5	0.81, 0.79
Teacher survey	T2GnJbSt	Teacher general job stress	T2C13a-T2C13d	10–50	0.89
Teacher survey	T2SupOff	Number of supports for staff wellness and overall well-being offered by program in the past year	T2C14	0–17	n.a.
Teacher survey	T2SupUse	Number of supports for staff wellness and overall well-being used or received in the past year	T2C15	0–17	n.a.
Teacher survey	T2SupNO	Number of supports for staff wellness and overall well-being that would have been useful and were not offered	T2C18, T2C19	0–17	n.a.
Teachers as caregivers					
Teacher survey	T1CCVrtl	Teacher's child/children had at least some virtual or remote	T1C05, T1C08	Virtual or remote instruction only/hybrid of in-person instruction and	n.a.

Exhibit VII.4. (*continued*)

Instrument	Variable name	Variable label	Instrument/survey items	Value labels and possible response ranges	2021–2022 study reliability (Cronbach's alpha)
		instruction at the start of 2021–2022 school/program year		virtual or remote instruction; in-person instruction only/homeschooled	
Teacher survey	T1CCVrtlS	Teacher's child/children had at least some virtual or hybrid instruction at the start of 2021–2022 school/program year and school/program plans have not changed	T1C09	Virtual instruction only and school/school district/provider has not changed plans; not virtual instruction only or school/school district/provider has changed plans	n.a.
Teacher survey	T1CCNdCt	Count of strategies used to meet child care needs outside of regular arrangements	T1C11a–T1C11f	0–6	n.a.
Teacher survey	T1CCNeed	One or more strategies used to meet child care needs outside of regular arrangements	T1C11a–T1C11f	One or more strategies; no strategies	n.a.

Note: The "n" used in a variable name is a placeholder for the wave number, with "1" referring to fall 2021 and "2" to spring 2022. If a particular variable is listed with the "n" in the text, it was constructed in more than one wave.

[a] Noted "values" reflect the actual response range obtained on the variable. For the data values associated with these labels, see Appendix F: Fall 2021 – Spring 2022 Classroom-/Teacher-Level File Codebook.

[b] Data users interested in using CLS_HS_Fund should consider using the fall classroom participation weight, CLS1_WT, because this information is available only for the subset of classrooms participating in fall 2021.

CES-D = Center for Epidemiological Studies Depression Scale; DLL = dual language learner; GAD=7 = General Anxiety Scale-7; n.a. = not applicable.

4. Head Start program and center characteristics

Composite variables for describing Head Start programs, centers, and directors' backgrounds came from one of three sources: (1) the program director survey, (2) the center director survey, and (3) the Head Start Program Information Report (2020-2021 program year). We describe the approach to constructing a subset of variables from all three instruments. In Exhibit VII.5, we note the instrument, variable name, variable label, source survey items, values and possible response ranges, and data collection wave associated with each variable.

a. Teacher turnover

We used center director reports to calculate lead teacher turnover. Turnover is defined as the number of lead teachers who left and had to be replaced in the last 12 months (C2A04) divided by the total number of lead teachers currently employed at the center (C2A01). The variable may underestimate the level of turnover if the director chose a response of three or more, which was top coded as three for the calculations. If source items were missing, we coded **C2TCHTRN** as missing.

b. Center language environment

Center directors reported on the language environment of centers. They reported the languages other than English spoken by children and families (**C2NMFLNG**; C2A12i11 through C2A12i20, C2A12iOth) and the languages other than English spoken by lead or assistant teachers (C2A12k11 through C2A12k20, C2A12kOth). Within each center, we compared the languages other than English spoken by children and families with those spoken by teachers; we then used the resultant information to calculate (1) the percentage of centers with Spanish-speaking families that also had Spanish-speaking lead or assistant teachers (**C2SPNTCH**) and (2) the percentage of the total number of languages other than English spoken by children and families in a center also spoken by that center's lead or assistant teachers (**C2PCTFLN**). We defined these variables only if the center director reported that any families and children speak a language other than English; these variables were set to missing if the only language spoken by families was English (C2A12h = 0). If source items were missing, we coded the composite variables as missing.

*c. Head Start year length **(C2YRLGTH)***

Center directors reported on the start and end dates of their Head Start year (C2A001). We took the difference between the two dates and rounded to the nearest month to calculate the length in months of the program year for Head Start-funded center-based slots.

d. Program schedules

C2PGMSCD indicates the number of days per week that the centers in a program operate; we constructed it at the program level. Center directors reported the number of days per week that Head Start-funded slots in their centers could receive services (D2C2A002a); the options were four days per week, five days per week, or both. We then combined responses from both center directors in a program to create a variable showing the full range of center-based schedules in a program.

e. Center director years of experience

Center directors reported their years of experience in Head Start programs and in their current center in response to instrument items C2I02a and C2I02b, respectively. They entered their years of experience,

which we used to construct two categorical variables—center director's years of experience in any Head Start program (**C2ANYHSP**) and their years of experience in their current Head Start center (**C2CURHSP**)—with the following categories: three years or less, four through nine years, 10 through 19 years, and 20 or more years.

f. Largest sources of funding other than Head Start

Program directors also reported on their two largest sources of funding other than Head Start (D2O03). We derived a set of variables from these funding items, indicating whether each was one of the two largest sources of funding other than Head Start:

- **D2PNTFND**: parent tuition or fees[62]
- **D2STFND**: tuition or fees paid by state government
- **D2LCLFND**: local government
- **D2FEDFND**: federal government other than Head Start
- **D2GNTFND**: grants or community organizations
- **D2GFTFND**: fundraising, gifts, bequests, or special events
- **D2PKFND:** state or local pre-K funds from state or local government
- **D2CCSFND:** child care subsidies
- **D2OTHFND**: another source of funding

D2REVSRC is the total count of program revenue sources other than Head Start. We also developed **D2GVTFND**, which indicates any government source other than Head Start as one of the two largest sources of a program's funding.

g. Coaches or mentors in the program

Program directors indicated the number of coaches or mentors by four types: (1) employees or staff hired by the program to serve as mentors or coaches and who devote most or all of their time to that role, (2) consultants or contractors hired by the program to serve as mentors or coaches, (3) other program employees or staff who serve as mentors or coaches but spend less than half of their time in that role, and (4) individuals from other organizations or agencies who provide free coaching or mentoring services to early childhood programs. We took the sum of these four types to arrive at the number of mentors or coaches in the program, which we calculated three different ways.

First, we constructed **D2MNTRCTR** for the number of mentors or coaches in the program and included values of zero for respondents who reported they had mentors or coaches (D2B06h) but entered zero for the number of staff of these four types (D2B24). Respondents not asked how many mentors/coaches were in their program (D2B24) (because they reported not having mentors or coaches in their program at all in item D2B06h) were treated as logical skips for this variable.

[62] Many Head Start programs serve non-Head Start children through other funding sources (including tuition). In addition, because many programs also serve Head Start families for longer than the Head Start day, they may require additional funds.

Second, we also constructed **D2NCMALL** for the number of staff who serve as a mentor or coach; this variable includes programs with no mentors or coaches by using values of zero both for those respondents who reported having mentors or coaches (D2B06h) but entered zero when asked how many (D2B24), and those who reported not having any mentors or coaches in their program (D2B06h).

Third, we constructed **D2NCMANY**, the number of staff who provide mentoring or coaching among programs with mentors or coaches; programs with no mentors or coaches would not be included in this variable because for respondents who reported they did not have mentors (D2B06h) or had mentors but entered zero for how many (D2B24), we treat these values as logical skips rather than values of zero. For this reason, **D2NCMANY** has the most restrictive sample size.

h. Data types linked electronically to child assessment information

Program directors were asked which of the following data and information types their program links electronically to child assessment information (D2N05c): child or family demographics; vision, hearing, developmental, social, emotional, and/or behavioral screenings; child attendance data; school readiness goals; family needs; service referrals for families; services received by families; parent or family attendance data; parent or family goals; Classroom Assessment Scoring System-Pre-K (Pre-K CLASS; Pianta et al. 2008) results or other quality measures; staff and teacher performance evaluations; personnel records; none of the above; and not applicable because the program does not store child assessment information in an electronic data system. We constructed a variable (**D2DTALNK**) that counts each type of data indicated by the program director.

i. Program director years of experience

Program directors reported their years of experience in Head Start programs generally and in their current program in response to instrument items D2I02a and D2I02b, respectively. They entered their years of experience, which we used to construct two categorical variables—program director's years of experience in any Head Start program (**D2ANYHSP**) and program director's years of experience in their current Head Start program (**D2CURHSP**) with the following categories: three years or less, four through nine years, 10 through 19 years, and 20 or more years.

j. Staff compensation and well-being supports

For different types of staff, program directors reported the types of compensation provided (for example, paid sick days, paid holidays, health benefits, or retirement benefits). Using items D2Z07a through D2Z07j, we constructed a variable indicating the number of types of compensation provided for each type of staff: educational personnel, such as teaching staff, family child care providers, and home visitors (**D2NmCoEP**); family service workers or child counselors (**D2NmCoFS**); managers or coordinators (**D2NmCoMC**); and facilities or support staff (**D2NmCoOt**). For the same staff, we also constructed a variable using items D2Z07b_a through D2Z07b_j to indicate whether programs had added or increased at least one type of compensation in the prior year (**D2IncCEP**, **D2IncCFS**, **D2IncCMC**, **D2IncCOt**).

Program directors also reported on the well-being supports provided to staff in various positions, including educational personnel, family service workers or child counselors, managers or coordinators, and facilities or support staff. Using items D2Z18a through D2Z18m, we constructed a variable indicating the number of types of well-being supports provided for each type of staff: educational personnel (**D2NmSpEP**), family service workers or child counselors (**D2NmSpFS**), managers or coordinators (**D2NmSpMC**), and facilities or support staff (**D2NmSpOt**). For the same staff, we also constructed a

variable using items D2Z18b_a through D2Z18b_m to indicate whether programs had added or increased at least one type of well-being support in the prior year (**D2IncSEP, D2IncSFS, D2IncSMC, D2IncSOt**).

k. Center director mental health

We calculated the center director's depressive symptoms score (**C2DEPSCO**) from responses to 12 items (C2C01a through C2C01l) on the center director survey (drawn from the CES–D). The CES-D is a screening tool, not a diagnostic tool, but we corelated scores with clinical diagnosis (Radloff 1977). Items used a 4-point scale ranging in value from 1 ("rarely or never") to 4 ("most or all"). First, we recoded the 12 items so responses ranged from zero to 3 instead of 1 to 4. Second, we calculated **C2DEPSCO** from the sum of the recoded interview items. If any one of the items was missing, we calculated **C2DEPSCO** by multiplying the average of the other 11 items by 12. If any two items were missing, we calculated **C2DEPSCO** by multiplying the average of the other 10 items by 12. If three or more items were missing, we coded **C2DEPSCO** as missing. Continuous depressive symptoms scores range from zero to 36.

We used values from the continuous score (**C2DEPSCO**) to categorize the level of depressive symptoms (**C2DEPCAT**). We categorized cases with values of **C2DEPSCO** from zero to 4 as "no to few symptoms" (**C2DEPCAT** = 1); from 5 to 9 as "mild" (**C2DEPCAT** = 2); from 10 to 14 as "moderate" (**C2DEPCAT** = 3); and 15 or higher as "severe" (**C2DEPCAT** = 4). We coded any cases for which **C2DEPCAT** was missing as missing on **C2DEPCAT**.

We calculated the center director's anxiety level score (**C2AnxSco**) from responses to seven items (C2C03a through C2C03g) on the center director survey (drawn from the GAD–7, which is a valid and reliable tool to screen for anxiety (Spitzer et al. 2006). Items used a 4-point scale ranging in value from zero ("not at all") to 3 ("nearly every day"). We calculated **C2AnxSco** from the sum of these items. If any one of the items was missing, we calculated **C2AnxSco** by multiplying the average of the other six items by seven. If any two items were missing, we calculated **C2AnxSco** by multiplying the average of the other five items by seven. If three or more items were missing, we coded **C2AnxSco** as missing. Continuous anxiety symptoms scores range from zero to 21.

We used values from the continuous score (**C2AnxSco**) to categorize the level of anxiety symptoms (**C2AnxCat**) (Spitzer et al. 2006). We categorized cases with values of **C2AnxSco** from zero to 4 as "minimal" (**C2AnxCat** = 1); from 5 to 9 as "mild" (**C2AnxCat** = 2); from 10 to 14 as "moderate" (**C2AnxCat** = 3); and 15 or higher as "severe" (**C2AnxCat** = 4). We coded any cases for which **C2AnxCat** was missing as missing on **C2AnxCat**.

l. Program director mental health

We calculated the program director's depressive symptoms score (**D2DEPSCO**) from responses to 12 items (D2C01a through D2C01l) on the center director survey (drawn from the CES–D). As noted previously, the CES-D is a screening tool, not a diagnostic tool, but we correlated scores with clinical diagnosis (Radloff 1977). Items used a 4-point scale ranging in value from 1 ("rarely or never") to 4 ("most or all"). First, we recoded the 12 items so responses ranged from zero to 3 instead of 1 to 4. Second, we calculated **D2DEPSCO** from the sum of the recoded interview items. If any one of the items was missing, we calculated **D2DEPSCO** by multiplying the average of the other 11 items by 12. If any two items were missing, we calculated **D2DEPSCO** by multiplying the average of the other 10 items by 12. If three or more items were missing, we coded **D2DEPSCO** as missing. Continuous depressive symptoms scores range from zero to 36.

We used values from the continuous score (**D2DEPSCO**) to categorize the level of depressive symptoms (**D2DEPCAT**). We categorized cases with values of **D2DEPSCO** from zero to 4 as "no to few symptoms" (**D2DEPCAT** = 1); from 5 to 9 as "mild" (**D2DEPCAT** = 2); from 10 to 14 as "moderate" (**D2DEPCAT** = 3); and 15 or higher as "severe" (**D2DEPCAT** = 4). We coded any cases for which **D2DEPCAT** was missing as missing on **D2DEPCAT**.

We calculated the program director's anxiety level score (**D2AnxSco**) from responses to seven items (D2C03a through D2C03g) on the program director survey (drawn from the GAD–7). As mentioned previously, the GAD–7 is a valid and reliable tool to screen for anxiety (Spitzer et al. 2006). Items used a 4-point scale and ranged in value from zero ("not at all") to 3 ("nearly every day"). We calculated **D2AnxSco** from the sum of these items. If any one of the items was missing, we calculated **D2AnxSco** by multiplying the average of the other six items by seven. If any two items were missing, we calculated **D2AnxSco** by multiplying the average of the other five items by seven. If three or more items were missing, we coded **D2AnxSco** as missing. Continuous anxiety symptoms scores range from zero to 21.

We used values from the continuous score (**D2AnxSco**) to categorize the level of anxiety symptoms (**D2AnxCat**). We categorized cases with values of **D2AnxSco** from zero to 4 as "minimal" (**D2AnxCat** = 1); from 5 to 9 as "mild" (**D2AnxCat** = 2); from 10 to 14 as "moderate" (**D2AnxCat** = 3); and 15 or higher as "severe" (**D2AnxCat** = 4). We coded any cases for which **D2AnxCat** was missing as missing on **D2AnxCat**.

m. Program-level covariates

We include four program-level covariates that indicated a risk for nonresponse bias even after weighting adjustments (see chapter VI). Data users can choose to control for these covariates in their multivariate analyses. All four covariates were derived from data from the Head Start Program Information Report (PIR).

The first covariate is **METRO**, which is an indicator for whether the program's zip code is in a Metropolitan Statistical Area[63] or not. The second covariate is **A_15_CAT**, which is the cumulative program enrollment from the 2018-2019 PIR, divided into three groups (0=small, 1=medium, 2=large) using the unweighted 33rd and 67th percentiles of the distribution among the 281 programs in our eligible program sample (whether participating or not).

The third covariate is **PTEACHHVLEFT_CAT**, which refers to the percentage of teachers and home visitors who left the program in the prior year. This constructed variable used three variables from the 2020-2021 PIR: The total number of teachers and home visitors was calculated as the sum of B_3_1 (the number of lead classroom teachers) and B_8 (the number of home visitors). The number of teachers and home visitors who left was from B_17.[64] This constructed variable is calculated as the ratio of the number who left divided by the total number, times 100, then divided into three groups (0=low, 1=medium, 2=high) using the unweighted percentiles as described above.

The fourth covariate is **PTEACHHVREPL_CAT**, which refers to the percentage of teachers and home visitors that left the program in the prior year that were replaced. This constructed variable used two variables from the 2020-2021 PIR: The number of teachers and home visitors who left was from B_17

[63] Each metropolitan statistical area must have at least one urbanized area of 50,000 or more inhabitants.

[64] It was brought to our attention after using this construct that B_17 (the number of teachers and home visitors who left) also included assistant teachers and family childcare providers, who were not included in the denominator. We decided to include this version of the variable on the file, as that is what was used for the nonresponse bias analysis.

and the number replaced was from B_17_1.[65] This constructed variable is calculated as the ratio of the number who were replaced divided by the number who left, times 100, then divided into three groups (0=low, 1=medium, 2=high) using the unweighted percentiles as described above.

[65] Both the numerator and denominator of this variable also contained counts of assistant teachers and family childcare providers.

Exhibit VII.5. 2021–2022 study composite variables—Head Start program and center characteristics

Instrument	Variable name	Variable label	Instrument/survey items	Value labels and possible response ranges	2021–2022 study reliability (Cronbach's alpha)
Center characteristics					
Center director survey	C2TCHTRN	Lead teacher turnover in center	C2A01, C2A04	0–>100	n.a.
Center director survey	C2SPNTCH	Spanish spoken by teachers and families	C2A12i12, C2A12k12	Yes; No	n.a.
Center director survey	C2NMFLNG	Number of languages spoken by center families	C2A12iXX, C2A12iOth	0–10[a]	n.a.
Center director survey	C2PCTFLN	Percent of family languages spoken by teachers	C2A12iXX, C2A12iOth, C2A12kXX, C2A12koth	0–>100	n.a.
Center director survey	C2YRLGTH	Length of Head Start program year in months	C2A001bmm, C2A001bdd, C2A001byy, C2A001amm, C2A001add, C2A001ayy	0–12	n.a.
Center director survey	C2PGMSCD	HS children receive services 4 or 5 days per week	C2A02a1, C2A02a2	4 days per week only; 5 days per week only; both 4 and 5 days per week	n.a.
Program-level aggregates and characteristics					
Program director survey	D2PNTFND	Parent tuition/fees 1 of 2 largest non HS revenues	D2O02a–D2O02i; D2O03_1	Yes; No	n.a.
Program director survey	D2STFND	State govt tuition/fees 1 of 2 largest non HS revenues	D2O02a–D2O02i; D2O03_2	Yes; No	n.a.
Program director survey	D2LCLFND	Local govt 1 of 2 largest non HS revenues	D2O02a–D2O02i; D2O03_3	Yes; No	n.a.
Program director survey	D2FEDFND	Fed govt other than HS 1 of 2 largest non HS revenues	D2O02a–D2O02i; D2O03_4	Yes; No	n.a.
Program director survey	D2GNTFND	Grants or comm orgs 1 of 2 largest non HS revenues	D2O02a–D2O02i; D2O03_5	Yes; No	n.a.
Program director survey	D2GFTFND	Fund raising, gifts, etc. 1 of 2 largest non HS revenues	D2O02a–D2O02i; D2O03_6	Yes; No	n.a.

Table VII.9 *(continued)*

Instrument	Variable name	Variable label	Instrument/survey items	Value labels and possible response ranges	2021–2022 study reliability (Cronbach's alpha)
Program director survey	D2PKFND	State or local Pre-K funds 1 of 2 largest non HS revenues	D2O02a-D2O02i, D2O03_9	Yes; No	n.a.
Program director survey	D2CCSFND	Child care subsidy funds 1 of 2 largest non HS revenues	D2O02a-D2O02i, D2O03_10	Yes; No	n.a.
Program director survey	D2OTHFND	Other 1 of 2 largest non HS revenues	D2O02a–D2O02i, D2O03_7	Yes; No	n.a.
Program director survey	D2GVTFND	Govt other than HS 1 of 2 largest non HS revenues	D2O02a–D2O02i, D2O03_2, D2O03_3, D2O03_4	Yes; No	n.a.
Program director survey	D2REVSRC	Count of program revenue sources other than HS	D2O02a-D2O02i	0–6[a]	n.a.
Program director survey	D2MNTRCTR	Count of mentors in program	D2B24b, D2B24c, D2B24d, and D2B24e	0–53[a]	n.a.
Program director survey	D2NCMALL	Number of staff who serve as mentor/coach (all programs)	D2B24b, D2B24c, D2B24d, and D2B24e	0–53[a]	n.a.
Program director survey	D2NCMANY	Num of mentors/coaches (progs w/mntrs or cchs)	D2B03h_6, D2B24b, D2B24c, D2B24d, and D2B24e	1–53[a]	n.a.
Program director survey	D2DTALNK	Count of data types that can be linked electronically to assessment data	D2N05c	0–12	n.a.
Program director survey	D2ANYHSP	Program director's years of experience in any Head Start program	D2I02a	3 years or less, 4 to 9 years, 10 to 19 years, 20 or more years	n.a.
Program director survey	D2CURHSP	Program director's years of experience in current Head Start program	D2I02b	3 years or less, 4 to 9 years, 10 to 19 years, 20 or more years	n.a.

Table VII.9 *(continued)*

Instrument	Variable name	Variable label	Instrument/survey items	Value labels and possible response ranges	2021–2022 study reliability (Cronbach's alpha)
Staff compensation and well-being supports					
Program director survey	D2InMnWg	Program increased wages for one or more staff positions from the minimum wage or within two dollars of minimum wage	D2Z04b_a–D2Z04b_e	0–1	n.a.
Program director survey	D2NmCoEP	Number of types of compensation provided to educational personnel	D2Z07a_a–D2Z07a_i	0–10	n.a.
Program director survey	D2NmCoFS	Number of types of compensation provided to family service workers or child counselors	D2Z07a_a–D2Z07a_i	0–10	n.a.
Program director survey	D2NmCoMC	Number of types of compensation provided to managers or coordinators	D2Z07a_a–D2Z07a_i	0–10	n.a.
Program director survey	D2NmCoOt	Number of types of compensation provided to other staff, such as facilities or support staff	D2Z07a_a–D2Z07a_i	0–10	n.a.
Program director survey	D2IncCEP	Increased one or more types of compensation provided to educational personnel	D2Z07b_a–D2Z07b_j	Yes; No	n.a.
Program director survey	D2IncCFS	Increased one or more types of compensation provided to family service workers or child counselors	D2Z07b_a–D2Z07b_j	Yes; No	n.a.
Program director survey	D2IncCMC	Increased one or more types of compensation provided to managers or coordinators	D2Z07b_a–D2Z07b_j	Yes; No	n.a.

Table VII.9 *(continued)*

Instrument	Variable name	Variable label	Instrument/survey items	Value labels and possible response ranges	2021–2022 study reliability (Cronbach's alpha)
Program director survey	D2IncCOt	Increased one or more types of compensation provided to other staff, such as facilities or support staff	D2Z07b_a–D2Z07b_j	Yes; No	n.a.
Program director survey	D2NmSpEP	Number of types of supports for staff well-being provided to educational personnel	D2Z18a_a–D2Z18a_i	0–13	n.a.
Program director survey	D2NmSpFS	Number of types of supports for staff well-being provided to family service workers or child counselors	D2Z18a_a–D2Z18a_i	0–13	n.a.
Program director survey	D2NmSpMC	Number of types of supports for staff well-being provided to managers or coordinators	D2Z18a_a–D2Z18a_i	0–13	n.a.
Program director survey	D2NmSpOt	Number of types of supports for staff well-being provided to other staff, such as facilities or support staff	D2Z18a_a–D2Z18a_i	0–13	n.a.
Program director survey	D2IncSEP	Program increased one or more types of supports for staff well-being provided to educational personnel	D2Z18b_a–D2Z18b_j	Yes; No	n.a.
Program director survey	D2IncSFS	Program increased one or more types of supports for staff well-being provided to family service workers or child counselors	D2Z18b_a–D2Z18b_j	Yes; No	n.a.
Program director survey	D2IncSMC	Program increased one or more types of supports for staff well-being provided to managers or coordinators	D2Z18b_a–D2Z18b_j	Yes; No	n.a.

Table VII.9 *(continued)*

Instrument	Variable name	Variable label	Instrument/survey items	Value labels and possible response ranges	2021–2022 study reliability (Cronbach's alpha)
Program director survey	D2IncSOt	Program increased one or more types of supports for staff well-being provided to other staff, such as facilities or support staff	D2Z18b_a–D2Z18b_j	Yes; No	n.a.
Center director characteristics					
Center director survey	C2RACE	Race/ethnicity of center director	C2I26, C2I28	White, non-Hispanic; Black or African American, non-Hispanic; Hispanic, Latino/a/, or Chicano/a/o; American Indian or Alaska Native, non-Hispanic; Asian or Pacific Islander, non-Hispanic; Multi-racial/bi-racial, non-Hispanic; Other race, non-Hispanic	n.a.
Center director survey	C2ANYHSP	Center director's years of experience in any Head Start program	C2I02a	3 years or less, 4 through 9 years, 10 through 19 years, 20 or more years	n.a.
Center director survey	C2CURHSP	Center director's years of experience in current Head Start center	C2I02b	3 years or less, 4 through 9 years, 10 through 19 years, 20 or more years	n.a.
Center director survey	C2DEPSCO	Center director's depressive symptoms, CES-D Short Form	C2C01a-C2C01l	0–36	0.92
Center director survey	C2DEPCAT	Center director's depressive score, CES-D Short Form, Categories	C2DEPSCO	Not depressed; Mildly depressed; Moderately depressed; Severely depressed	n.a.
Center director survey	C2AnxSco	Center director anxiety level score, GAD-7	C2C03a-C2C03g	0–21	0.97
Center director survey	C2AnxCat	Center director anxiety level category, GAD-7	C2AnxSco	Minimal anxiety; Mild anxiety; Moderate anxiety; Severe anxiety	n.a.

Table VII.9 *(continued)*

Instrument	Variable name	Variable label	Instrument/survey items	Value labels and possible response ranges	2021–2022 study reliability (Cronbach's alpha)
Center director survey	C2CVJbSt	Center Director COVID-19 Job Stress	C2C04a-C2C04d	1–5	0.74
Program director characteristics					
Program director survey	D2RACE	Race/ethnicity of program director	D2I26, D2I28	White, non-Hispanic; Black or African American, non-Hispanic; Hispanic, Latino/a/x, or Chicano/a/o; American Indian or Alaska Native, non-Hispanic; Asian or Pacific Islander, non-Hispanic; Multi-racial/Bi-racial, non-Hispanic; Other Race, non-Hispanic	n.a.
Program director survey	D2DEPSCO	Program director's depressive symptoms, CES-D Short Form	D2C01a-D2C01l	0–36	0.91
Program director survey	D2DEPCAT	Program director's depressive score, CES-D Short Form, Categories	D2DEPSCO	Not depressed; Mildly depressed; Moderately depressed; Severely depressed	n.a.
Program director survey	D2AnxSco	Program director anxiety level score, GAD-7	D2C03a-D2C03g	0–21	0.98
Program director survey	D2AnxCat	Program director anxiety level category, GAD-7	D2AnxSco	Minimal anxiety; Mild anxiety; Moderate anxiety; Severe anxiety	n.a.
Program director survey	D2CVJbSt	Program Director COVID-19 Job Stress	D2C04a-D2C04d	1–5	0.75
Program-level covariates					
2018-2019 Head Start Program Information Report	METRO	Program in MSA	Program_ZIP_Code	1=MSA 2=non-MSA	n.a.
2018-2019 Head Start Program Information Report	A_15_CAT	Program cumulative enrollment	A_15	0=small 1=medium 2=large	n.a.

Table VII.9 *(continued)*

Instrument	Variable name	Variable label	Instrument/survey items	Value labels and possible response ranges	2021–2022 study reliability (Cronbach's alpha)
2020-2021 Head Start Program Information Report	PTEACHHVLEFT_CAT	Program percentage of teachers and home visitors who left	B_3_1, B_8, B_17	0=low 1=medium 2=high	n.a.
2020-2021 Head Start Program Information Report	PTEACHHVREPL_CAT	Program percentage of teachers and home visitors replaced	B_17, B_17_a	0=low 1=medium 2=high	n.a.

Note: The "n" used in a variable name is a placeholder for the wave number, with "1" referring to fall 2021 and "2" to spring 2022. If a particular variable is listed with the "n" in the text, it was constructed in more than one wave.

[a] Noted "values" reflect the actual response range obtained on the variable. For the data values associated with these labels, see Appendix E: Spring 2022 Center/Program-Level File Codebook.

CES-D = Center for Epidemiological Studies Depression Scale; GAD-7 = General Anxiety Scale-7; n.a. = not applicable.

This page has been left blank for double-sided copying.

References

Aarons, G., M. Hurlburt, and S. McCue Horwitz. "Advancing a Conceptual Model of Evidence-Based Practice Implementation in Public Service Sectors." *Administration and Policy in Mental Health*, vol. 38, no. 1, 2011, pp. 4–23.

Administration for Children and Families. "Head Start FACES 2000: A Whole-Child Perspective on Program Performance, Fourth Progress Report." Washington, DC: U.S. Department of Health and Human Services, 2003.

Administration on Children, Youth, and Families. "Head Start Program Performance Measures, Second Progress Report." Washington, DC: U.S. Department of Health and Human Services, 1998.

Aikens, N., A. Kopack Klein, E. Knas, J. Hartog, M. Manley, L. Malone, L. Tarullo, and S. Lukashanets. "Child and Family Outcomes During the Head Start Year: FACES 2014–2015 Data Tables and Study Design." OPRE Report 2017-100. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2017a.

Aikens, N., A. Kopack Klein, E. Knas, M. Reid, A. Mraz Esposito, M. Manley, L. Malone, et al. "Descriptive Data on Head Start Children and Families from FACES 2014: Fall 2014 Data Tables and Study Design." OPRE Report 2017-97. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2017b.

Aikens, N., E. Knas, L. Malone, L. Tarullo, and J.F. Harding. "A Spotlight on Dual Language Learners in Head Start: FACES 2014." OPRE Report 2017-99. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2017c.

Aikens, Nikki, Louisa Tarullo, Lara Hulsey, Christine Ross, Jerry West, and Yange Xue. "A Year in Head Start: Children, Families, and Programs." Report submitted to the Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. Washington, DC: Mathematica Policy Research, October 2010.

Amato, P. "The Impact of Family Formation Change on the Cognitive, Social, and Emotional Well-Being of the Next Generation." *The Future of Children*, vol. 15, no. 2, 2005, pp. 75–96.

Amato, Paul, and Frieda Fowler. "Parenting Practices, Child Adjustment, and Family Diversity." *Journal of Marriage and Family*, vol. 64, no. 3, February 2004, pp. 703–716.

Ammerman, R., F. Putnam, J. Kopke, T. Gannon, J. Short, J. Van Ginkel, M. Clark, et al. "Development and Implementation of a Quality Assurance Infrastructure in a Multistate Home Visitation Program in Ohio and Kentucky." *Journal of Prevention & Intervention in the Community*, vol. 34, no. 1/2, 2007, pp. 89–107.

Aquilino, W.S. "The Lifecourse of Children Born to Unmarried Mothers: Childhood Living Arrangements and Young Adult Outcomes." *Journal of Marriage and Family*, vol. 58, no. 2, 1996, pp. 293–310.

Atkins-Burnett, S., and N. Aikens. "Sweet Dreams: Linkages Between Sleep Quality and Quantity and Children's Developmental Outcomes in a Head Start Sample." Presented at the Biennial Meeting of the Society for Research in Child Development, Montreal, Canada, March 2011.

Baker, C., and S.E. Rimm-Kaufman. "How Homes Influence Schools: Early Parenting Predicts African American Children's Classroom Social-Emotional Functioning." *Psychology in the Schools,* vol. 51, no. 7, 2014, pp. 722–735.

Bane, M.J., and D.T. Ellwood. "The Dynamics of Dependence: The Routes to Self-Sufficiency." Cambridge, MA: Urban Systems Research and Engineering, 1983.

Bassok D., M. Fitzpatrick, S. Loeb, and A. Paglayan. "The Early Childhood Care and Education Workforce from 1990 Through 2010: Changing Dynamics and Persistent Concerns." *Education Finance and Policy,* vol. 8, no. 4, 2013, pp. 581–601.

Berry, Daniel J., Lisa J. Bridges, and Martha J. Zaslow. "Early Childhood Measures Profiles." Washington, DC: Child Trends, 2004.

Bialystok, E. "Consequences of Bilingualism for Cognitive Development." In *Handbook of Bilingualism: Psycholinguistic Approaches,* edited by J.F. Kroll and A.M.B. de Groot, 417–432. New York, NY: Oxford University Press, 2005.

Bickel, G., M. Nord, C. Price, W. Hamilton, and J. Cook. "Guide to Measuring Household Food Security, Revised 2000." Alexandria, VA: Food and Nutrition Services, U.S. Department of Agriculture, March 2000.

Bloom, P.J., S. Jackson, T.N. Talan, and R. Kelton. "Taking Charge of Change: A 20-Year Review of Empowering Early Childhood Administrators Through Leadership Training." Wheeling, IL: McCormick Center for Early Childhood Leadership, 2013.

Blue-Banning, M., J. Summers, H. Frankland, L. Nelson, and G. Beegle. "Dimensions of Family and Professional Partnerships: Constructive Guidelines for Collaboration." *Exceptional Children*, vol. 70, no. 2, 2004, pp. 167–184.

Blumberg, Stephen J., K. Bialostosky, W.L. Hamilton, and R.R. Briefel. "The Effectiveness of a Short Form of the Household Food Security Scale." *American Journal of Public Health,* vol. 89, no. 8, August 1999. Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1508674/pdf/amjph00008-0087.pdf. Accessed June 19, 2023.

Bodovski, K., and M. Young. "The Long Term Effects of Early Acquired Skills and Behaviors on Young Children's Achievement in Literacy and Mathematics." *Journal of Early Childhood Research,* vol. 9, no. 1, 2011, pp. 4–19.

Boyatzis, Richard, and Annie McKee. *Resonant Leadership*. Boston, MA: Harvard Business School Press, 2005.

Bradley, Robert H., Robert F. Corwyn, Margaret Burchinal, Harriette Pipes McAdoo, and Cynthia García Coll. "The Home Environments of Children in the United States Part II: Relations with Behavioral Development Through Age 13." *Child Development*, vol. 72, no.6, 2001, pp. 1868–1886.

Bronte-Tinkew, J., L. Guzman, S. Jekielek, K.A. Moore, S. Ryan, Z. Redd, J. Carrano, and G. Matthews. "Conceptualizing and Measuring 'Healthy Marriage' for Empirical Research and Evaluation Studies: A Review of the Literature and Annotated Bibliography (Task Three)." Prepared for Office of Planning and Research, Administration for Children and Families, U.S. Department of Health and Human Services. Bethesda, MD: Child Trends, 2004.

Bryk, A.S., and B. Schneider. "Trust in Schools: A Core Resource for School Reform." *Educational Leadership*, vol. 60, no. 6, 2003, pp. 40–45.

Burchinal, M., L. Vernon-Feagans, V. Vitiello, M. Greenberg, and The Family Life Project Key Investigators. "Thresholds in the Association Between Child Care Quality and Child Outcomes in Rural Preschool Children." *Early Childhood Research Quarterly*, vol. 29, no. 1, 2014, pp. 41–51.

Burchinal, M.R., A. Follmer, and D.M. Bryant. "The Relations of Maternal Social Support and Family Structure with Maternal Responsiveness and Child Outcomes." *Developmental Psychology*, vol. 32, no. 6, 1996, pp. 1073–1083.

Burchinal, M.R., E. Peisner-Feinberg, R. Pianta, and C. Howes. "Development of Academic Skills from Preschool Through Second Grade: Family and Classroom Predictors of Developmental Trajectories." *Journal of School Psychology,* vol. 40, no. 5, 2002, pp. 415–436.

Burchinal, M.R., Y. Xue, H. Tien, A. Auger, and A. Mashburn. "Testing for Threshold in Associations Between Child Care Quality and Child Outcomes." Presented at the Biennial Conference of the Society for Research in Child Development, Montreal, Canada, March 2011.

Burts, D.C., C.H. Hart, R. Charlesworth, and L. Kirk. "A Comparison of Frequencies of Stress Behaviors Observed in Kindergarten Children in Classrooms with Developmentally Appropriate Versus Developmentally Inappropriate Instructional Practices." *Early Childhood Research Quarterly*, vol. 5, no. 3, 1990, pp. 407–423.

Caspe, Margaret, and Elena Lopez. "Lessons from Family-Strengthening Interventions: Learning from Evidence-Based Practice." Cambridge, MA: Harvard Family Research Project, 2006.

Chromy, J.R. "Sequential Sample Selection Methods." *Proceedings of the Survey Research Methods Section of the American Statistical Association,* 401–406. Alexandria, VA: American Statistical Association, 1979.

Conger, R.D., K.J. Conger, G.H. Elder, F.O. Lorenz, R.L. Simons, and L.B. Whitbeck. "Family Economic Stress and Adjustment of Early Adolescent Girls." *Developmental Psychology*, vol. 29, no. 2, 1993, pp. 206–219.

Cooper, C.L. "Executive Stress: A Ten-Country Comparison." *Human Resource Management,* vol. 23, no. 4, 1984, pp. 395–407.

Copeland, B.W., and S. Kirsch. "Perceived Occupational Stress Among NCAA Division I, II, and III Athletic Directors." *Journal of Sport Management,* vol. 9, no. 1, 1995, pp. 70–77.

Crockenberg, S.B. "Infant Irritability, Mother Responsiveness, and Social Support Influences on the Security of Infant-Mother Attachment." *Child Development*, vol. 52, no. 3, 1981, pp. 857–865.

Currie, J. "Health Disparities and Gaps in School Readiness." *Future of Children*, vol. 15, no. 1, spring 2005, pp. 117–138.

Dawson, D.A. "Family Structure and Children's Health and Well-Being: Data from the 1998 National Health Interview Survey on Child Health." *Journal of Marriage and Family*, vol. 50, no. 3, 1991, pp. 619–648.

Dickinson, D.K., and P.O. Tabors. *Beginning Literacy with Language: Young Children Learning at Home and School*. Baltimore, MD: Paul H. Brookes Publishing, 2001.

Doran, Elizabeth, Natalie Reid, Sara Bernstein, Tutrang Nguyen, Myley Dang, Ann Li, Ashley Kopack Klein, et al. "A Portrait of Head Start Classrooms and Programs in Spring 2020: FACES 2019 Descriptive Data Tables and Study Design." OPRE Report #2022-15. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2022.

Douglass, Anne L. *Leading for Change in Early Care and Education: Cultivating Leadership from Within.* New York, NY: Teachers College Press, 2017.

Downer, Jason T., and Julia L. Mendez. "African American Father Involvement and Preschool Children's School Readiness." *Early Education & Development*, vol. 16, no. 3, 2005, pp. 317–340.

Downey, G., and J.C. Coyne. "Children of Depressed Parents: An Integrative Review." *Psychological Bulletin*, vol. 108, no. 1, 1990, pp. 50–76.

Duncan, G.J., and K. Magnuson. "Investing in Preschool Programs." *The Journal of Economic Perspectives: A Journal of the American Economic Association*, vol. 27, no. 2, 2013, pp. 109–132. https://doi.org/10.1257/jep.27.2.109.

Elliott, S.N., F.M. Gresham, T. Freeman, and G. McCloskey. "Teacher and Observer Ratings of Children's Social Skills: Validation of the Social Skills Rating System." *Journal of Psychoeducational Assessment*, vol. 6, no. 2, 1988, pp. 152–161.

Entwisle, D.R., K.L. Alexander, D. Cadigan, and P.M. Pallis. "The Emergent Academic Self-Image of First Graders: Its Response to Social Structure." *Child Development*, vol. 58, no. 5, 1987, pp. 1190–1206.

Fantuzzo, J., C. McWayne, M.A. Perry, and S. Childs. "Multiple Dimensions of Family Involvement and Their Relations to Behavioral and Learning Competencies for Urban, Low-Income Children." *School Psychology Review*, vol. 33, no. 4, 2004, pp. 467–480.

Fantuzzo, John, and Christy McWayne. "The Relationship Between Peer-Play Interaction in the Family Context and Dimensions of School Readiness for Low-Income Preschool Children." *Journal of Educational Psychology*, vol. 94, no. 1, March 2002, pp. 79–87.

Farewell, C.V., J. Quinlan, L. Gonzales, and J. Puma. "Changes in Demands and Resources Faced by the Early Childhood Education Workforce due to COVID-19." *Journal of Early Childhood Research,* vol. 21, no. 1, 2023, pp. 63–75. https://doi.org/10.1177/1476718X221136463.

Ferretti, L.K., and K.L. Bub. "The Influence of Family Routines on the Resilience of Low-Income Preschoolers." *Journal of Applied Developmental Psychology*, vol. 35, no. 3, 2014, pp. 168–180.

Field, T.M. "Infants of Depressed Mothers." In *Stress, Coping, and Depression*, edited by S.L. Johnson, A.M. Hayes, T.M. Field, N. Schneiderman, and P.M. McGabe, 3–22. Mahwah, NJ: Lawrence Erlbaum Associates, 2000.

Fiese, B.H., and M. Schwartz. "Reclaiming the Family Table: Mealtimes and Child Health and Well-Being." *Social Policy Report*, vol. 22, no. 4, 2008, pp. 1–20.

Foster, M.A., R. Lambert, M. Abbott-Shim, F. McCarty, and S. Franze. "A Model of Home Learning Environment and Social Risk Factors in Relation to Children's Emergent Literacy and Social Outcomes." *Early Childhood Research Quarterly*, vol. 20, no. 1, 2005, pp. 13–36.

Frankel, K.A., C.D. Croy, L.F. Kubicek, R.N. Emde, C.M. Mitchell, and P. Spicer. "Toddler Socioemotional Behavior in a Northern Plains Indian Tribe: Associations with Maternal Psychosocial Well-Being." *Infant Mental Health Journal*, vol. 35, no. 1, 2014, pp. 10–20.

Fry, P.S. "Perfectionism, Humor, and Optimism as Moderators of Health Outcomes and Determinants of Coping Styles of Women Executives." *Genetic, Social, and General Psychology Monographs,* vol. 121, no. 2, 1995, pp. 211–245.

Fuligni, A.S., E. Hoff, M. Zepeda, and P. Mangione. "Development of Infants and Toddlers Who Are Dual Language Learners." Working paper No. 2 from the Center for Early Care and Education Research–Dual Language Learners. Chapel Hill, NC: Frank Porter Graham Child Development Institute, University of North Carolina, March 2014.

Garfinkel, I., and S. McLanahan. *Single Mothers and Their Children: A New American Dilemma*. Washington, DC: The Urban Institute Press, 1986.

Gassman-Pines, A., and L. Gennetian. "COVID-19 Job and Income Loss Jeopardize Child Well-Being: Income Support Policies Can Help." SRCD Child Evidence Brief No. 9. Washington, DC: Society for Research in Child Development, December 2020.

Gassman-Pines, A., E.O. Ananat, J. Fitz-Henley II, and J. Leer. "Effect of Daily School and Care Disruptions During the COVID-19 Pandemic on Child Behavior Problems." *Developmental Psychology,* vol. 58, no. 8, 2022, pp. 1512–1527. https://doi.org/10.1037/dev0001373.

Gaylor, E., X. Wei, and M.M. Burnham. "Associations Between Nighttime Sleep Duration and Developmental Outcomes in a Nationally Representative Sample of Preschool-Age Children." *Sleep*, vol. 33, 2010, p. A17.

Gerber, Emily B., Marcy Whitebook, and Rhona S. Weinstein. "At the Heart of Child Care: Predictors of Teacher Sensitivity in Center-Based Child Care." *Early Childhood Research Quarterly*, vol. 22, no. 3, 2007, pp. 327–346.

Glasheen, C., G.A. Richardson, and A. Fabio. "A Systematic Review of the Effects of Postnatal Maternal Anxiety on Children." *Archives of Women's Mental Health,* vol. 13, no. 1, 2010, pp. 61–74. https://doi.org/10.1007/s00737-009-0109-y.

Glick, J.E., and B. Hohmann-Marrott. "Academic Performance of Young Children in Immigrant Families: The Significance of Race, Ethnicity, and National Origins." *International Migration Review*, vol. 41, no. 2, 2007, pp. 361–402.

Grant, A.A., L. Jeon, and C.K. Buettner. Relating early childhood teachers' working conditions and well-being to their turnover intentions. *Educational Psychology,* vol. 39, no. 3, 2019, pp. 294–312.

Greenhalgh, T., J. Russell, and D. Swinglehurst. "Narrative Methods in Quality Improvement Research." *Journal of Quality and Safety in Health Care*, vol. 14, no. 16, 2005, pp. 443–449.

Gresham, F.M., and S.N. Elliot. *Social Skills Rating System*. Circle Pines, MN: American Guidance Service, 1990.

Gullo, D. "A Structural Model of Early Indicators of School Readiness Among Children Of Poverty." *Journal of Children and Poverty,* vol. 24, no. 1, 2018, pp. 3–24.

Hall-Kenyon, K.M., R.V. Bullough, K.L. MacKay, and E.E. Marshall. "Preschool Teacher Well-Being: A Review of the Literature." *Early Childhood Education Journal,* vol. 42, no. 3, 2014, pp. 153–162.

Hart, B., and T.R. Risley. *Meaningful Differences in the Everyday Experiences of Young American Children.* Baltimore, MD: Paul H. Brookes, 1995.

Hascher, T., and J. Waber. "Teacher Well-Being: A Systematic Review of the Research Literature from the Year 2000–2019." *Educational Research Review,* vol. 34, 2021, article 100411.

Hayes, C.D., J.L. Palmer, and M.J. Zaslow. "Who Cares for America's Children? Child Care Policy for the 1990s." Washington, DC: American Family Press, 1990.

Hindman, A., and A. Bustamante. "Teacher Depression as a Dynamic Variable: Exploring the Nature and Predictors of Change over the Head Start Year." *Journal of Applied Developmental Psychology,* vol. 61, 2019, pp. 43–55.

Hiscock, Harriet, Louise Canterford, Obioha C. Ukoumunne, and Melissa Wake. "Adverse Associations of Sleep Problems in Australian Preschoolers: National Population Study." *Pediatrics*, vol. 119, no. 1, 2007, pp. 86–93.

Hofferth, Sandra L., and John F. Sandberg. "How American Children Spend Their Time." *Journal of Marriage and Family*, vol. 63, no. 2, 2001, pp. 295–308.

Horm, D., N. File, D. Bryant, M. Burchinal, H. Raikes, N. Forestieri, A. Encinger, and A. Cobo-Lewis. "Associations Between Continuity of Care in Infant-Toddler Classrooms and Child Outcomes." *Early Childhood Research Quarterly,* vol. 42, 2018, pp. 105–118.

Hsu, S.H.J., D.R. Chen, Y. Cheng, and T.C. Su. "Association of Psychosocial Work Hazards with Depression and Suboptimal Health in Executive Employees." *Journal of Occupational and Environmental Medicine,* vol. 58, no. 7, 2016, pp. 728–736.

Hulsey, L.K., N. Aikens, A. Kopack, J. West, E. Moiduddin, and L. Tarullo. "Head Start Children, Families, and Programs: Present and Past Data from FACES." OPRE Report 2011-33a. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2011.

Institute of Behavioral Research. "TCU Survey of Organizational Functioning (TCU SOF)." Fort Worth, TX: Institute of Behavioral Research, Texas Christian University, 2005. https://ibr.tcu.edu/forms/organizational-staff-assessments/

Jackson, A.P. "The Effects of Nonresident Father Involvement on Single Black Mothers and Their Young Children." *Social Work*, vol. 44, no. 2, 1999, pp. 156–166.

Janus, M., and E. Duku. "The School Entry Gap: Socioeconomic, Family, and Health Factors Associated with Children's School Readiness to Learn." *Early Education and Development*, vol. 18, no. 3, 2007, pp. 375–403.

Jeon, S., L. Jeon, S. Lang, and K. Newell. "Teacher Depressive Symptoms and Child Math Achievement in Head Start: The Roles of Family–Teacher Relationships and Approaches to Learning." *Child Development,* vol. 92, no. 6, 2021, pp. 2478–2495.

Kim, Y., E. Montoya, L.J.E. Austin, A. Powell, and W. Muruvi. "Early Care and Education Programs During COVID-19: Persistent Inequities and Emerging Challenges." Berkeley, CA: Center for the Study of Child Care Employment, University of California, Berkeley, 2022. https://cscce.berkeley.edu/publications/report/early-care-and-education-programs-during-covid-19.

Kish, Leslie. *Survey Sampling.* Wiley Classics Library Edition. New York: John Wiley & Sons, Inc., 1965, 1995.

Knight, D.K., J.E. Becan, and P.M. Flynn. "Organizational Consequences of Staff Turnover in Outpatient Substance Abuse Treatment Programs." *Journal of Substance Abuse Treatment,* vol. 42, no. 2, 2012, pp. 143–150. Available at https://www.sciencedirect.com/science/article/pii/S0740547211002121. Accessed June 19, 2023.

Kopack Klein, A., B. Lepidus Carlson, N. Aikens, A. Bloomenthal, J. West, L. Malone, E. Moiduddin, et al. "Head Start Family and Child Experiences Survey (2014) User's Manual." Report submitted to the U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation. Washington, DC: Mathematica Policy Research, 2018.

Kopack Klein, A., B. Lepidus Carlson, N. Aikens, A. Li, S. Bernstein, M. Dang, M. Scott, et al. "Head Start Family and Child Experiences Survey (FACES 2019): User's Manual." Report submitted to the U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation. Washington, DC: Mathematica, 2021a.

Kopack Klein, A., N. Aikens, A. Li, S. Bernstein, N. Reid, M. Dang, E. Blesson, et al. "Descriptive Data on Head Start Children and Families from FACES 2019: Fall 2019 Data Tables and Study Design." OPRE Report 2021-77. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2021b.

Korucu, I., E. Litkowski, and S. Schmitt. "Examining Associations Between the Home Literacy Environment, Executive Function, and School Readiness." *Early Education and Development,* vol. 31, no. 3, January 2020, pp. 455–473.

Kracht, C.L., P.T. Katzmarzyk, and A.E. Staiano. "Household Chaos, Family Routines, and Young Child Movement Behaviors in the U.S. During the COVID-19 Outbreak: A Cross-Sectional Study." *BMC Public Health,* vol 21, 2021, article 860.

Kruse, S. "Creating Communities of Reform: Continuous Improvement Planning Teams." *Journal of Educational Administration*, vol. 39, no. 4, 2001, pp. 359–383.

Krysik, J., and C.W. LeCroy. "Development and Initial Validation of an Outcome Measure for Home Visitation: The Healthy Families Parenting Inventory." *Infant Mental Health Journal,* vol. 33, no. 5, September/October 2012, pp. 496–505. https://doi.org/10.1002/imhj.21343.

Lavigne, John V., Richard Arend, Diane Rosenbaum, Andy Smith, Marc Weissbluth, Helen J. Binns, and Katherine Kaufer Christoffel. "Sleep and Behavior Problems Among Preschoolers." *Journal of Developmental & Behavioral Pediatrics*, vol. 20, no. 3, 1999, pp. 164–169.

LeCroy, C., K. Milligan, L. Malofeeva, and H. McPhail. "The Healthy Families Parenting Inventory (HFPI): Developing a Multi-Ise Tool for Guiding Services and Measuring Outcomes in Home Visitation Programs." Tucson, AZ: LeCroy & Milligan Associates, Inc., 2007. Available at https://www.acf.hhs.gov/sites/default/files/documents/cb/es2011_poster_19.pdf. Accessed June 19, 2023.

Lehman, W.E., J.M. Greener, and D.D. Simpson. "Assessing Organizational Readiness for Change." *Journal of Substance Abuse Treatment,* vol. 22, no. 4, 2002, pp. 197–209. https://doi.org/10.1016/S0740-5472(02)00233-7.

Liu, S., and P.A. Fisher. "Early Experience Unpredictability in Child Development as a Model for Understanding the Impact of the COVID-19 Pandemic: A Translational Neuroscience Perspective." *Developmental Cognitive Neuroscience,* vol. 54, April 2022, article 101091.

Lopez, M.E., S. Dorros, and H. Weiss. "Family-Centered Child Care." Cambridge, MA: Harvard Family Research Project, 1999.

Love, J.M., P.Z. Schochet, and A. Meckstrom. "Are They in Any Real Danger? What Research Does—and Doesn't—Tell Us about Child Care Quality and Children's Well-Being." Princeton, NJ: Mathematica Policy Research, 1996.

Löwe, B., O. Decker, S. Müller, E. Brähler, D. Schellberg, W. Herzog, and P.Y. Herzberg. "Validation and Standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the General Population." *Medical Care,* vol. 43, no. 3, 2008, pp. 266–274.

Lynch, J. "Infant Health, Race/Ethnicity, and Early Educational Outcomes Using the ECLS-B." *Sociological Inquiry*, vol. 81, no. 4, 2011, pp. 499–526.

Marcon, R. "Positive Relationships Between Parent School Involvement and Public School Inner-City Preschoolers' Development and Academic Performance." *School Psychology Review*, vol. 28, no. 3, 1999, pp. 395–412.

Marsh, J.A., J.F. Pane, and L.S. Hamilton. "Making Sense of Data-Driven Decision Making in Education: Evidence from Recent RAND Research." Santa Monica, CA: RAND Corporation, 2006.

Marzano, R., T. Waters, and B. McNulty. "School Leadership That Works: From Research to Results." Alexandria, VA: Association for Supervision and Curriculum Development, 2005.

Masarik, A.S., and R.D. Conger. "Stress and Child Development: A Review of the Family Stress Model." *Current Opinion in Psychology,* vol. 13, 2017, pp. 85–90.

McCormick Center for Early Childhood Leadership. "Head Start Administrative Practices, Director Qualifications, and Links to Classroom Quality." Chicago, IL: McCormick Center for Early Childhood Leadership Publications, November 9, 2010. Available at https://digitalcommons.nl.edu/mccormickcenter-pubs/20.

McLanahan, S., and G. Sandefur. *Growing Up with a Single Parent: What Hurts, What Helps*. Cambridge, MA: Harvard University Press, 1994.

McLean, Leigh, and Carol McDonald Connor. "Depressive Symptoms in Third-Grade Teachers: Relations to Classroom Quality and Student Achievement." *Child Development*, vol. 86, no. 3, May/June 2015, pp. 945–954.

McWayne, Christine, Virginia Hampton, John Fantuzzo, Heather Cohen, and Yumiko Sekino. "A Multivariate Examination of Parent Involvement and the Social and Academic Competencies of Urban Kindergarten Children." *Psychology in the Schools*, vol. 41, no. 3, February 2004, pp. 363–377.

Mensah, F.K., and K.E. Kiernan. "Parents' Mental Health and Children's Cognitive and Social Development: Families in England in the Millennium Cohort Study." *Social Psychiatry and Psychiatric Epidemiology,* vol. 45, 2010, pp. 1023–1035.

Moiduddin, E., N. Aikens, L. Tarullo, J. West, and Y. Xue. "Child Outcomes and Classroom Quality in FACES 2009." OPRE Report 2012-37a. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2012.

Moiduddin, Emily, Marta Induni, and Sara Duffy. "Developing a Shared Definition of Parent and Community Engagement: Supporting Analyses and a Proposed Definition." Final report submitted to First 5 LA. Washington, DC: Mathematica Policy Research, 2010.

Morrison, D.R., and A.J. Cherlin. "The Divorce Process and Young Children's Well-Being: A Prospective Analysis." *Journal of Marriage and Family*, vol. 57, no. 3, 1995, pp. 800–812.

Murphey, D., E. Cook, S. Beckwith, and J. Belford. "The Health of Parents and Their Children: A Two-Generation Inquiry." Bethesda, MD: Child Trends, 2018.

National Center on Parent, Family, and Community Engagement (NCPFCE). "Family Connections to Peers and Community." Available at http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/family/docs/rtp-family-connections.pdf. Accessed 2013.

National Research Council and Institute of Medicine. *Depression in Parents, Parenting and Children: Opportunities to Improve Identification, Treatment, and Prevention Efforts*. Washington, DC: National Academies Press, 2009. Available at https://www.ncbi.nlm.nih.gov/books/NBK215117/. Accessed June 19, 2023.

NICHD Early Child Care Research Network. "Early Child Care and Self-Control, Compliance, and Problem Behavior at 24 and 36 Months." *Child Development*, vol. 69, no. 4, 1998, pp. 1145–1170.

NICHD Early Child Care Research Network. "The NICHD Study of Early Child Care and Youth Development (SECCYD): Findings for Children up to 4½ Years." NIH Pub. No, 05-4318. Washington, DC: Eunice Kennedy Shriver National Institute of Health and Human Development, National Institutes of Health, Department of Health and Human Services, 2006.

NICHD Early Child Care Research Network. "The Relation of Child Care to Cognitive and Language Development." *Child Development*, vol. 71, no. 4, 2000, pp. 960–980.

Olinske, J.L., and C.M. Hellman. "Leadership in the Human Service Nonprofit Organization: The Influence of the Board of Directors on Executive Director Well-Being and Burnout." *Human Service Organizations: Management, Leadership & Governance,* vol. 41, no. 2, 2017, pp. 95–105.

Owen, M.T., J.F. Klauski, A. Mata-Otero, and M.O. Caughy. "Relationship-Focused Child Care Practices: Quality of Care and Child Outcomes for Children in Poverty." *Early Education and Development*, vol. 19, no. 2, 2008, pp. 302–329.

Patel, Sanjay R., and Frank B. Hu. "Short Sleep Duration and Weight Gain: A Systematic Review." *Obesity*, vol. 16, no. 3, 2008, pp. 643–653.

Patrick, S.W., L.E., Henkhaus, J.S. Zickafoose, K. Lovell, A. Halvorson, S. Loch, M. Letterie, and M.M. Davis. "Well-Being of Parents and Children During the COVID-19 Pandemic: A National Survey." *Pediatrics,* vol. 146, no. 4, 2020, article e2020016824.

Penuel, W.R., B.J. Fishman, R. Yamaguchi, and L.P. Gallagher. "What Makes Professional Development Effective? Strategies that Foster Curriculum Implementation." *American Educational Research Journal,* vol. 44, no. 4, 2007, pp. 921–958.

Peterson, J., and N. Zill. "Marital Disruption, Parent-Child Relationships, and Behavior Problems in Children." *Journal of Marriage and Family*, vol. 48, no. 2, 1986, pp. 295–307.

Pianta, Robert, K. LaParo, and B. Hamre. *The Classroom Assessment Scoring System Pre-K Manual*. Charlottesville, VA: University of Virginia, 2008.

Phillips, D.A., S. Anderson, A.R. Datta, and E. Kisker. "The Changing Portrait of Center-Based Preschool Teachers: 1990 and 2012." *Children and Youth Services Review,* vol 107, 2019, article 104558.

Radloff, L. "The CES-D Scale: A Self-Report Depression Scale for Research in the General Population." *Applied Psychological Measurement*, vol. 1, no. 3, 1977, pp. 385–401.

RAPID Survey Project. "Latest Data & Trends." Available at https://rapidsurveyproject.com/latest-data-and-trends. Accessed February 14, 2023 (updated monthly).

Ravid, Sarit, Iris Afek, Suheir Suraiya, Eli Shahar, and Giora Pillar. "Sleep Disturbances are Associated with Reduced School Achievements in First-Grade Pupils." *Developmental Neuropsychology*, vol. 34, no. 5, 2009, pp. 574–587.

Ribar, D. "Why Marriage Matters for Child Well-Being." *The Future of Children*, vol. 25, no. 2, 2015, pp. 11–27.

Rijlaarsdam, J., H. Tiemeier, A.P. Ringoot, M.Y. Ivanova, V.W. Jaddoe, F.C. Verhulst, and S.J. Roza. "Early Family Regularity Protects Against Later Disruptive Behavior." *European Child & Adolescent Psychiatry*, vol. 25, November 2015, pp. 1–9.

Rogers, R.E., E.Y. Li, and R. Ellis. "Perceptions of Organizational Stress Among Female Executives in the US Government: An Exploratory Study." *Public Personnel Management,* vol. 23, no. 4, 1994, pp. 593–609.

Rohacek, Monica, Gina C. Adams, and Ellen E. Kisker. "Understanding Quality in Context: Child Care Centers, Communities, Markets, and Public Policy." Washington, DC: The Urban Institute, 2010.

Ross, C.E., J. Mirowsky, and J. Huber. "Dividing Work, Sharing Work, and In-Between: Marriage Patterns and Depression." *American Sociological Review*, vol. 48, no. 6, December 1983, pp. 809–823.

Russell, B.S., M. Hutchison, R. Tambling, A.J. Tomkunas, and A.L. Horton. "Initial Challenges of Caregiving During COVID-19: Caregiver Burden, Mental Health, and the Parent–Child Relationship." *Child Psychiatry & Human Development,* vol. 51, 2020, pp. 671–682. https://doi.org/10.1007/s10578-020-01037-x.

Ryu, J.H., and J.S. Bartfeld. "Household Food Insecurity During Childhood and Subsequent Health Status: The Early Childhood Longitudinal Study—Kindergarten Cohort." *American Journal of Public Health*, vol. 102, no. 11, 2012, pp. e50–e55.

Sarama, J., D.H., Clements, C.B. Wolfe, and M.E. Spitler. "Professional Development in Early Mathematics: Effects of an Intervention Based on Learning Trajectories on Teachers' Practices." *Nordic Studies in Mathematics Education,* vol. 21, no. 4, 2016, pp. 29–55.

Sarche, M.C., C.D. Croy, C. Big Crow, C.M. Mitchell, and P. Spicer. "Maternal Correlates of 2-Year-Old American Indian Children's Social-Emotional Development in a Northern Plains Tribe." *Infant Mental Health Journal*, vol. 30, no. 4, 2009, pp. 321–340.

Schaack, D.D., V.N. Le, and J. Stedron. "When Fulfillment is not Enough: Early Childhood Teacher Occupational Burnout and Turnover Intentions from a Job Demands and Resources Perspective." *Early Education and Development,* vol. 31, no. 7, 2020, pp. 1011–1030.

Schlieber, M., and J. Han. "The Sleeping Patterns of Head Start Children and the Influence on Developmental Outcomes." *Child: Care, Health and Development,* vol. 44, no. 3, 2018, pp. 462–469.

Sekine, Michikazu, Takashi Yamagami, Kyoko Handa, Tomohiro Saito, Seiichiro Nanri, Katsuhiko Kawaminami, Noritaka Tokui, et al. "A Dose-Response Relationship Between Short Sleeping Hours and Childhood Obesity: Results of the Toyama Birth Cohort Study." *Child: Care, Health and Development*, vol. 28, no. 2, 2002, pp. 163–170.

Sharabi, A., and D. Marom-Golan. "Social Support, Education Levels, and Parents' Involvement: A Comparison Between Mothers and Fathers of Young Children with Autism Spectrum Disorder." *Topics in Early Childhood Special Education,* vol. 38, no. 1, 2018, pp. 54–64.

Sheldon, S.B. "Parents' Social Network and Beliefs as Predictors of Parent Involvement." *The Elementary School Journal*, vol. 102, no. 4, 2002, pp. 301–316.

Shonkoff, J.P., and D.A. Phillips. *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington, DC: National Academies Press, 2000.

Smith, Sheila, Veronica Benavides, and Faith Lamb-Parker. "New York City Pre-K Leadership Study." New York, NY: Bank Street College, April 2019.

Spielberg, L. "Successful Family Engagement in the Classroom: What Teachers Need to Know and Be Able to Do to Engage Families in Raising Student Achievement." Washington, DC: Flamboyan Foundation, 2011.

Spillane, J. "Data in Practice: Conceptualizing the Data-Based Decision-Making Phenomena." *American Journal of Education*, vol. 118, no. 2, February 2012, pp. 113–141.

Spitzer, R.L., K. Kroenke, J.B. Williams, and B. Löwe. "A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7." *Archives of Internal Medicine,* vol. 166, no. 10, 2006, pp. 1092–1097.

Swigonski, N.L., B. James, W/ Wynns, and K. Casavan. "Physical, Mental, and Financial Stress Impacts of COVID-19 on Early Childhood Educators." *Early Childhood Education Journal,* vol. 49, no. 5, 2021, pp. 799–806.

Talan, Teri N., Paula J. Bloom, and Robyn E. Kelton. "Building the Leadership Capacity of Early Childhood Directors: An Evaluation of a Leadership Development Model." *Early Childhood Research & Practice,* vol. 16, no. 1–2, 2014. Available at https://files.eric.ed.gov/fulltext/EJ1045231.pdf. Accessed June 19, 2023.

Taras, Howard, and William Potts-Datema. "Sleep and Student Performance at School." *Journal of School Health*, vol. 75, no. 7, 2005, pp. 248–254.

Tarullo, L. "The Children and Families of Head Start: National Data from FACES 2006." Presented at the U.S. Department of Health and Human Services, Secretary's Advisory Committee on Head Start Research and Evaluation, Washington, DC, January 25, 2011a.

Tarullo, L. "Learning Environments in Head Start: National Data from FACES." Presented at the U.S. Department of Health and Human Services, Secretary's Advisory Committee on Head Start Research and Evaluation, Washington, DC, April 12, 2011b.

Tarullo, L. "Head Start Families' Engagement with Programs: Data from FACES." Presented at the U.S. Department of Health and Human Services, Secretary's Advisory Committee on Head Start Research and Evaluation, Washington, DC, June 7, 2011c.

Tarullo, L. "Health and Developmental Status of Head Start Children: Data from FACES 2006." Presented at the U.S. Department of Health and Human Services, Secretary's Advisory Committee on Head Start Research and Evaluation, Washington, DC, September 22, 2011d.

Tarullo, L., E. Knas, A. Kopack Klein, N. Aikens, L. Malone, and J.F. Harding. "A National Portrait of Head Start Children and Families: FACES 2014." OPRE Report 2017-98. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2017.

Tarullo, Louisa, Nikki Aikens, Cheri Vogel, and Kimberly Boller. "Learning from National Research: Informing and Refining Practice in Head Start and Early Head Start Programs." Presented at the Head Start National Birth to Five Leadership Institute, National Harbor, MD, April 30, 2013.

Toldson, I., and B. Lemmons. "Social Demographics, the School Environment, and Parenting Practices Associated with Parents' Participation in Schools and Academic Success Among Black, Hispanic, and White Students." *Journal of Human Behavior in the Social Environment,* vol. 23, no. 2, 2013, pp. 237–255.

U.S. Census Bureau. "Survey of Income and Program Participation (SIPP) 2008 Panel: Topical Module Items Booklet Wave 9." Washington, DC: Census Bureau, 2011. Available at https://www2.census.gov/programs-surveys/sipp/questionnaires/2008/sipp-2008-panel-wave-09-topical-module-questionnaire.pdf. Accessed June 19, 2023.

U.S. Department of Education, National Center for Education Statistics. "Early Childhood Longitudinal Study, Kindergarten Class of 2022-23 (ECLS-K:2023), Preschool Field Test, OMB# 1850-0750 v.19, Attachment B1 Web Instruments." June 2020. https://omb.report/icr/201908-1850-007/doc/94454004.

U.S. Department of Education, National Center for Education Statistics. "Early Childhood Longitudinal Study—Kindergarten Class of 1998–99 (ECLS–K), Psychometric Report for Kindergarten through First Grade." NCES 2002–05. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, 2002.

U.S. Department of Education, National Center for Education Statistics. "National Household Education Survey of 1999: Data File User's Manual, Volume I, NCES 2000–076." Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education, October 2000. Available at https://nces.ed.gov/pubs2000/2000076.pdf. Accessed June 19, 2023.

van Buuren, S., and K. Groothuis-Oudshoorn. "mice: Multivariate Imputation by Chained Equations in R." *Journal of Statistical Software,* vol. 45, no. 3, 2011, pp. 1–67. https://doi.org/10.18637/jss.v045.i03.

Van IJzendoorn, M.H., L.W.C. Tavecchio, G. Stams, M. Verhoeven, and E. Reiling. "Attunement Between Parents and Professional Caregivers: A Comparison of Childrearing Attitudes in Different Child Care Settings." *Journal of Marriage and Family*, vol. 60, no. 3, 1998, pp. 771–781.

Vinopal, K. "Understanding Individual and Organizational Level Representation: The Case of Parental Involvement in Schools." *Journal of Public Administration Research and Theory,* vol. 28, no. 1, 2018, pp. 1–15.

Votruba-Drzal, E., R.L. Coley, and P.L. Chase-Lansdale. "Child Care and Low-Income Children's Development: Direct and Moderated Effects." *Child Development*, vol. 71, no. 1, 2004, pp. 296–312.

Wang, H., E. Demerouti, and P. Le Blanc. "Transformational Leadership, Adaptability, and Job Crafting: The Moderating Role of Organizational Identification." *Journal of Vocational Behavior,* vol. 100, 2017, pp. 185–195.

Weiland, C., M. McCormick, S. Mattera, M. Maier, and P. Morris. "Preschool Curricula and Professional Development Features for Getting to High-Quality Implementation at Scale: A Comparative Review Across Five Trials." *AERA Open,* vol. 4, no. 1, 2018, 2332858418757735.

Weiss, Heather, Margaret Caspe, and M. Elena Lopez. "Family Involvement in Early Childhood Education" Cambridge, MA: Harvard Family Research Project, 2006.

Wells, M.B. "Predicting Preschool Teacher Retention and Turnover in Newly Hired Head Start Teachers Across the First Half of the School Year." *Early Childhood Research Quarterly,* vol 30, 2015, pp. 152–159.

Whalen, S.P., H.L. Horsley, K.K. Parkinson, and D. Pacchiano. "A Development Evaluation Study of a Professional Development Initiative to Strengthen Organizational Conditions in Early Education Settings." *Journal of Applied Research on Children: Informing Policy for Children at Risk,* vol. 7, no. 2, 2016, article 9.

Whitaker, R.C., T. Dearth-Wesley, and R.A. Gooze. "Workplace Stress and the Quality of Teacher–Children Relationships in Head Start." *Early Childhood Research Quarterly,* vol. 30, 2015, pp. 57–69.

Whitebook, M., C. Howes, and D. Phillips. "Who Cares? Child Care Teachers and the Quality of Care in America." Oakland, CA: Child Care Employee Project, 1989.

Whitebook M., D. Phillips, and C. Howes. "Worthy Work, STILL Unlivable Wages: The Early Childhood Workforce 25 Years After the National Child Care Staffing Study." Berkeley, CA: Center for the Study of Child Care Employment, University of California, Berkeley, 2014.

Whitworth, B., and J. Chiu. "Professional Development and Teacher Change: The Missing Leadership Link." *Journal of Science Teacher Education,* vol. 26, 2015, pp. 121–137.

Worrall, L., and C.L. Cooper. "Executive Stress in Different Industrial Sectors, Structures and Sizes of Business." *Personnel Review,* vol. 24, no. 7, 1995, pp. 3–12.

Xue, Yange, Margaret Burchinal, Anamarie Auger, Hsiao-Chuan Tien, and Louisa Banks Tarullo. "Dosage Effects in Early Intervention and Child Care Quality: Evidence from Secondary Data Analysis." Presented at the Biennial Meeting of the Society for Research in Child Development, Montreal, Canada, March 31, 2011.

Yoshikawa, H., C. Weiland, J. Brooks-Gunn, M. Burchinal, L. Espinosa, W.T. Gormley, J. Ludwig, et al. "Investing in Our Future: The Evidence Base on Preschool." Washington, DC: Society for Research in Child Development, October 2013.

Zaslow, M., J. Bronte-Tinkew, R. Capps, A. Horowitz, K.A. Moore, and D. Weinstein. "Food Security during Infancy: Implications for Attachment and Mental Proficiency in Toddlerhood." *Maternal and Child Health Journal*, vol. 13, no. 1, 2009, pp. 66–80.

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

Zaslow, M., K. Tout, T. Halle, J.V. Whittaker, and B. Lavelle. "Toward the Identification of Features of Effective Professional Development for Early Childhood Educators: Literature Review." Washington, DC: U.S. Department of Education, Office of Planning, Evaluation and Policy Development, 2010b.

Zaslow, M.J. "Variations in Child Care Quality and Its Implications for Children." *Journal of Social Issues*, vol. 47, no. 2, 1991, pp. 125–138.

Zill, N., S. Alva, R.W. O'Brien, M. Vaden-Kiernan, M.A. D-Elio, R.H. McKey, S. Pai-Samant, et al. "The Changing Environmental Context of Head Start Children from Three Cohorts of the Head Start Family and Child Experiences Survey." Symposium presented at the biennial meeting of the Society for Research in Child Development, Atlanta, GA, April 2005.

This page has been left blank for double-sided copying.

This page has been left blank for double-sided copying.

Mathematica Inc.

Princeton, NJ • Ann Arbor, MI • Cambridge, MA
Chicago, IL • Oakland, CA • Seattle, WA
Woodlawn, MD • Washington, DC

mathematica.org

EDI Global, a Mathematica Company

Bukoba, Tanzania • High Wycombe, United Kingdom