

Head Start Family and Child Experiences Survey

FACES 2014

SELF-GUIDED TRAINING

Module 4

**Sampling Weights and Variance
Estimation**

Prepared by

Jerry West, Ph.D.

University of Maryland, College Park

**For Child Care and Early Education Research Connections,
National Center for Children in Poverty, Columbia University**

**With support from the Office of Planning, Research and Evaluation,
Administration for Children and Families,
U.S. Department of Health and Human Services**

The development of these training materials benefited greatly from comments and suggestions provided by the staff of the Office of Planning, Research and Evaluation. Thanks is also extended to students at the graduate program of the College of Education, University of Maryland at College Park for their feedback and to participants in the summer data training program at the Inter-university Consortium for Political and Social Research, University of Michigan, who pilot tested the training. Any errors remain the responsibility of the Research Connections project.

Module Objectives

- **Assist researchers in producing accurate and reliable answers to their research questions**
 - **Introduce researchers to sampling weights and why they need to use them when analyzing FACES 2014 data**
 - **Offer guidance on the use of weights and how to decide on the weight that is right for different types analyses**
 - **Introduce issues surrounding variance estimation and standard errors for complex sample designs**
 - **Describe methods for estimating variances and standard errors when using FACES 2014 data**
-

The overall goal of this module is to assist researchers in developing accurate and reliable answers to questions about Head Start and the children and families who are served by the program. The module includes two topics that are central to this overall goal. The first topical module focuses on sampling weights. It begins with a general introduction to sampling weights with the goal of answering researchers' questions about why they need to use sampling weights, and how they are created and used. This general introduction is followed by a description of the sampling weights included with the FACES 2014 data. It describes which of these weights to use when analyzing data at different levels and points in time, and when answering different types of research questions. Some of the most common questions users have about weights and their use are addressed.

The second topical module focuses on estimating variances and standard errors for complex sample designs, and is structured much like the module on sampling weights. First, researchers are introduced to the issues surrounding the estimation of variances and standard errors when the data come from a complex sample design, not a simple random sample (SRS). It introduces design-based methods for calculating standard errors and contrasts these with methods that assume the data come from a SRS. Second, it describes how design-based methods can be used to estimate standard errors for FACES 2014. Experienced researchers may want to opt out of those sections which introduce sampling weights and variance estimation in general. However, all researchers would benefit from reviewing those parts that deal directly with FACES 2014 sampling weights and variance estimation.

Topic 4.1

Sampling Weights

Sampling Weights Generate Many Questions from Data Users

- **In this section we will answer the following questions:**
 - **What is a sampling weight and why are sampling weights important?**
 - **Why use weights in your analyses?**
 - **Why are there so many different weights on the data files?**
 - **Can I use weights with my software and how?**
 - **Answers to some of the questions researchers ask most often about the use of weights are provided?**
-

No topic generates more questions from secondary data users than sampling weights. The goal of this section is to provide researchers with a basic understanding of sampling weights and why they should use them when analyzing data from FACES 2014 and earlier rounds, as well as when analyzing data from other large-scale studies with complex sample designs. The material is organized around the questions and topics that are listed on the slide.

Why Use Weights?

- **Weights account for not collecting data from the entire population and for using a complex sample design**
 - **Weights**
 - **Adjust for differential selection probabilities**
 - **Adjust for differential nonresponse and attrition**
 - **Reduce potential bias associated with nonresponse**
 - **Are used when estimating characteristics of the population**
-

Sampling weights adjust for different features of complex sample designs like the one used in FACES and for nonresponse and sample attrition that is inherent in any national longitudinal study. Specifically, weights adjust for the fact that not all units had an equal chance of selection into the sample. They also adjust for differences in nonresponse among certain groups of the population and for sample attrition. They can help to reduce the potential for bias when there is differential nonresponse and/or attrition. In FACES, the basic sampling weight (inverse probability of selection) is adjusted for nonresponse at each stage of sampling and for instrument nonresponse (unit nonresponse).

Sampling weights adjust for unit nonresponse (an entire case or instrument is missing). They do not adjust for item nonresponse (individual items in an instrument have missing data). Different imputation methods are used to adjust for item-level missing data.

Weights are used when estimating characteristics of the population, which is the primary goal of FACES. FACES is not interested in the characteristics of the 176 programs, 667 classrooms, and 2,462 children in the study sample per se, but in the characteristics of the population of Head Start programs, classrooms and children. The data collected from sample cases when used with the appropriate sampling weights produce reliable estimates for the population of programs, centers, classrooms, teachers and other staff, and children.

Why are Weights Important?

- Using unweighted data may give you different and perhaps inaccurate results:

Children's age	4	3	3	5	4
Weight	1	4	5	2	3

- Unweighted mean: $(4+3+3+5+4)/5 = 3.8$
 - Weighted mean:
 $\{(4*1)+(3*4)+(3*5)+(5*2)+(4*3)\}/15=3.5$
-

Let's continue our example, but now let's include some data – Let's say we want to estimate the average age of children in a class with 15 children. We could select a sample of the children and use their ages to estimate the average age of all 15 children in the class. If a simple random sample of the children in the classroom is selected, we could get an estimate of the average age of the children by calculating the simple mean of their ages. However, if a different sampling approach is used, one in which children might represent different numbers of children in the class or only themselves, calculating the average age without first adjusting for the differential probabilities of selection would result in an inaccurate estimate of the average age of 15 children in the class. As we will see later, using the weights also adjusts for other features of the survey such as nonresponse and sample attrition and can help to offset the potential bias associated with these.

This simple example shows the results of using the data from the same five children with and without weighting to estimate the age of the children in the class. The unweighted mean does not take into account the sample design where children were sampled with different probabilities and thus represent different numbers of children in the class. Here, each sampled child represents 1 to 5 children in the class. The weighted mean adjusts for these different probabilities of selection, and in our example using the weighted mean leads to a different and lower estimate of the average age of the class.

What is a Sampling Weight?

- **Characteristics of unweighted and weighted data**
 - When we use unweighted data, each observation is counted equally
 - Unweighted data represents only the observations of those in the sample with data
 - When we use weighted data, the observations are counted relative to sample members' representation in the population
-

As discussed in Module 1 (Topic 1.3 – FACES 2014 Sample), FACES is not a simple random sample of Head Start programs, classrooms and children. FACES uses a complex design that includes multi-levels of sampling and stratification to reduce the cost of the study and to ensure that the samples are representative of the population from which they were drawn. Sampling weights are used to compute estimates of the population (for example, average assessment scores, percentages of classrooms with a teacher who has a Bachelor's degree, and average enrollments in Head Start programs) that reflect the sample design. So, what is a sampling weight and what is the difference between unweighted and weighted data?

When unweighted data are used, each observation is counted equally and the data represents only the observations of those programs and individuals in the sample. For example, the mean age of children attending Head Start in fall 2014 who are participating in FACES 2014 can be calculated by summing the ages of all of the children with data on the FACES child file and dividing by the number of children with age data. When weighted data are used, the observations are counted relative to sample members' representation in the population from which the sample was drawn. Using the same data as before together with one of the FACES 2014 child-level weights we can estimate the mean age of the total population of children attending Head Start in fall 2014. We sum the weighted age of all children on the data file and divide by the sum of all the weights for those children.

What is a Sampling Weight? Let's Look at an Example

- **If we want to know the mean or average age for all children in one Head Start classroom, we just sum everyone's age and divide by the number of children in the classroom**
 - **Alternatively, we could take a sample of the children in the class and use their ages to estimate the average age of all the children in the class**
 - **We might use an approach were some children are sampled at a higher rate than others**
 - **Calculating an unweighted and weighted mean using the data from this sample would most likely produce a different result**
-

Let's look at a simple example in order to examine further the differences between unweighted and weighted data.

If we want to know the mean age for all the children in one Head Start classroom, we could sum the ages of all the children in the classroom and divide by the number of children in the classroom. Alternatively, we could select a sample of the children and use their ages to estimate the average age of all the children in the class (that is, the population average). Depending on how the sample is selected, some children might represent multiple children in the class and different numbers of children, and others might represent only themselves. Thus, calculating an unweighted (each child counts once) and weighted (different children represent different numbers of children in the class) averages would most likely produce a different result. NOTE: If we selected a simple random sample of the children in the class with each child having the same probability of being selected, the unweighted and weighted estimates would be the same.

FACES and Sampling Weights

- **FACES is designed to produce national estimates**
 - **FACES is a sample, i.e., the entire population was not surveyed.**
 - **FACES is not a simple random sample (SRS). Not all Head Start programs, classrooms, and children had an equal probability of selection.**
 - **Not all programs, teachers, parents and children participated.**
 - **Children who leave Head Start during the program year can become ineligible.**
-

There are several features of FACES and most large-scale national studies, that are important to keep in mind when considering to use weighted or unweighted data in your analysis.

- FACES is designed to produce national estimates for Head Start programs and centers, classrooms and teachers, and children and their families.
- FACES is a sample survey. It is not a census of Head Start programs, centers, classrooms and children. Only a small portion of each of these populations was selected and participated in the study.
- As mentioned several times before, FACES is not a simple random sample, and programs, classrooms and children did not have an equal probability of selection.
- Not all programs, staff, parents and children who were sampled or asked to complete a study instrument did so. There was unit nonresponse, which varied by instrument.
- Children who left their Head Start program after the fall data collection were not eligible for the spring follow up. Therefore, the sampling weights for the spring sample have been adjusted to reflect this.

FACES 2014 Weights Vary According to Several Factors

- **Level of analysis: program, center, classroom/teacher, program staff, child**
 - **Waves of data collection: Fall or spring (cross-sectional) or both fall and spring (program year)**
 - **Source(s) of data:**
 - **Core Study instruments – program/center director surveys, teacher surveys and teacher child reports, classroom observations, parent surveys, and child assessments**
 - **Family Engagement instruments – parent interview and family service staff interview**
-

The sampling weights developed for Core and Plus study data are designed for use in analyzing data at different levels and when using data from fall 2014 or spring 2015 alone (cross-sectional) or together (program year). There are also weights that support analyses of data from one or more of the Core study instruments and Family engagement interviews.

Why are there so many different weights on the FACES data files?

- **There are 21 different weights on the FACES 2014 data files**
 - **17 Core study weights**
 - 14 cross-sectional
 - 3 program year
 - **4 Family Engagement Plus study weights**
 - 3 cross-sectional
 - 1 program year
-

There are a total of 21 weights on the FACES 2014 data files, 17 Core study weights and 4 Family Engagement Plus study weights. A summary of the weights can be found in the User's Manual (See Chapter VI, Sections B and C, pgs. 180-192). There are three sets of weights listed in the document: fall 2014 cross-sectional weights, spring 2015 cross-sectional weights and program year weights. The cross-sectional weights are used when analyzing data from a single wave of the study (fall 2014 or spring 2015) and the program year weights are used to analyze data across both waves. Core Study weights are included in each set and the Family Engagement study weights are only included in the spring 2015 and program year sets.

There are only three fall cross-sectional weights that support analyses of the fall Core child assessment, parent survey and Teacher Child Report (TCR) data. There are many more spring cross-sectional weights that support analyses using these same three Core instruments plus several additional instruments that were administered only in spring 2015 (for example, Head Start teacher and director surveys, classroom observations). The number of program year weights is again smaller, supporting longitudinal analyses of the Core child assessment, parent and TCR data.

The Family Engagement study spring cross-sectional weights support analysis of the family service staff interview data and analyses of the parent interview data together with Core parent survey data alone or in combination with data from the spring child assessment and TCRs. The single Family Engagement program year weight supports analyses using the data from both the Core parent survey and Family Engagement parent survey along with fall and spring child assessment or TCR data.

Descriptions of how these weights were developed can be found in the User's Manual. (See Chapter VI, Section B, starting on pg. 180).

Which Weight Should I Use?

- **Weight should match your research question**
 - **What is the level of analysis?**
 - Program, center, classroom, teacher, staff, child/parent
 - **What data source(s) will be used?**
 - See tables on the next two slides
 - **Does answering your question require one or two waves of data?**
 - Fall 2014, spring 2015 or both
 - **There is not always a perfect weight!**
-

Deciding which of the 21 sampling weights to use can be a challenge, but there are a few simple rules that you can follow to make this task more manageable. First, and most important, the weight you choose should support the analysis that is required to answer your research question. More specifically, what is the target population and the level of analysis required to answer your research question(s)? What data sources will you be working with or where will you find data on the behaviors, outcomes, and characteristics of interest? Does answering your question require that you use one or two waves of data (is your focus on a single time point or on change between the fall and spring of the program year)? See User's Manual (Chapter VI, Sections B and C, pgs. 180-192) to go back to the document you reviewed earlier. It describes the decision making process and is a handy reference to use when you are working with the FACES data.

Answering these questions will help you decide on which of the 21 weights is best for your purposes. You will not always find a weight that perfectly fits your needs. In many cases, you will need to decide between two or more weights, basing your decision on several factors such as how much missing data are acceptable and how important is it to have full information even if the number of cases must be reduced to do so. Comparing the number of cases that have a positive value for each of the weights you are considering will identify the impact of each weight on the size of your analytic sample. Examining the unweighted distributions for key variables for cases with a positive value on each of the different weights will quickly tell you how much missing data you will have if you use one weight versus another.

Summary of Data Sources by Wave: Core Studies

Data Source	Fall 2014	Spring 2015
Direct Child Assessment	X	X
Teacher Child Report	X	X
Parent Survey	X	X
Teacher Survey		X
Classroom Observation		X
Program Director Survey		X
Center Director Survey		X

As a reminder, data for the Core studies were collected in the fall and spring of the 2014-2015 program year (in the 60 programs with child-level data collection) or only in spring 2015 (in the 116 programs with no child-level data collection). Listed in the table are all of the Core study instruments along with information on when each was used.

Summary of Data Sources by Wave: Family Engagement Plus Study

Data Source	Fall 2014	Spring 2015
Parent Interview		X
Family Service Staff Interview		X
Supplement Core Parent Survey Items		X
Supplemental Core Teacher Survey Items		X

There are fewer data sources for the Family Engagement Plus study and all data collection occurred in spring 2015.

Example 1 – Choosing the Best Weight

- **Do children’s language, literacy, and early math skills in the fall of the program year differ by gender and race/ethnicity?**
 - How many waves of data are needed?
 - What is/are the source(s) of these data?
 - What is the level of analysis?

 - **Best weight: *P1_RA1WT***
 - To answer the question you will use children’s scores on the direct child assessment that was administered in fall 2014
 - *P1_RA1WT* is one of two cross-sectional weights that are used when analyzing child assessment data from the fall
 - The other weight is *PRA1WT*; however, it also requires Teacher Child Report data, which are not needed here
-

Let’s take a look at several examples of research questions that can be answered with FACES 2014 and decide which weight is the best to use when answering each question.

The best weight to use for example 1 is *P1_RA1WT*. To answer the question you will be using data from a single wave (fall 2014). Furthermore, data on children’s language, literacy, and early math skills come from the direct child assessments. You are not using data from the Teacher Child Report, so there is no need to lose cases by selecting the weight that requires both the child assessment data and TCR data (*PRA1WT*). Also, the unit analysis here is the child, so you are opting for a child-level weight. Both cross-sectional child-level weights as well as all program year weights are conditioned on having a parent survey.

Example 2 – Choosing the Best Weight

- Do the gains children make in their language, literacy, and math skills across the Head Start year differ by their skill level in the fall of the year?
 - How many waves of data are needed?
 - What is/are the source(s) for these data?
 - What is the level of analysis?
 - **Best weight: *PRA12WT***
 - To answer the question you will use data from both the fall and spring direct assessments
 - *PRA12WT* is one of two child-level program year weights that are used when analyzing child assessment data from both the fall and spring
 - The other weight is *PRA12OCW*; however, it also requires spring teacher survey and classroom observation data that are not needed here
-

The best weight for example 2 is *PRA12WT*. The unit analysis here is the child, so you are opting for a child-level weight. You are not using data from the spring teacher survey or classroom observations. Therefore, *PRA12OCW* would not be the appropriate weight to use.

Example 3 – Choosing the Best Weight

- **Are the gains children make in their language, literacy, and math skills during their time in Head Start influenced by the quality of their Head Start classrooms and teachers?**
 - How many waves of data are needed?
 - What is/are the source(s) of these data?
 - What is the level of analysis?

 - **Best weight: *PRA12OCW***
 - To answer this question you will use data from the teacher survey and classroom observations together with data from the fall and spring direct child assessments
 - *PRA12OCW* is the only child-level program year weight that supports analyses using this combination of instruments
-

The best weight for example 3 is PRA12OCW. To answer the question you will be using data from both the fall and spring child assessment, and the spring classroom observations and teacher survey. The unit analysis here is the child, so you are opting for a program year, child-level weight.

Example 4 – Choosing the Best Weight

- Does classroom quality vary by classroom and teacher characteristics such as class size and teaching experience?
 - How many waves of data are needed?
 - What is/are the source(s) of these data?
 - What is the level of analysis?

 - **Best weight:** *TO2CLSWT*
 - To answer the question you will use data from the spring classroom observations and teacher survey.
 - *TO2CLSWT is the only spring 2015 cross-sectional weight that supports analyses that use data from both instruments*
-

The best weight for example 4 is TO2CLSWT. To answer the question you will be using data from the spring classroom observations and teacher survey. All of these instruments were administered only in spring 2015, so you will be using a cross-sectional weight. The unit analysis here is the classroom, so you are opting for a classroom-level weight. In this analysis, teacher characteristics such as education, work experience and job satisfaction will be treated as classroom variables.

If your research question required that you only use data from the classroom observations or the teacher survey, you would use O2CLSWT or T2CLSWT, respectively.

Example 5 – Choosing the Best Weight

- **In what ways are Head Start programs able to link data on family needs and services to child assessment information?**
 - How many waves of data are needed?
 - What is/are the source(s) of these data?
 - What is the level of analysis?

 - **Best weight: *D2WT***
 - *D2WT* is the only weight that is provided for use in analyzing data from the program director survey
-

The best weight to use here is D2WT, which is the only weight that is available for use when working with the program director survey data that are required to answer this question. The program director survey is administered only once, in the spring the program year (spring 2015). The unit analysis here is the program (or program director), so you are opting for a program-level weight.

Example 6 – Choosing the Best Weight

- **How do Head Start parents and staff characterize their relationships and interactions with one another?**
 - How many waves of data are needed?
 - What is/are the source(s) of these data?
 - What is the level of analysis?
 - **Best weight: PE2WT, F2WT, T2TCHWT**
 - To answer this question you could use data from several different sources, including the two Family Engagement interviews and the supplemental items that are asked in the Core parent and teacher surveys
 - The three weights will support three different levels of analysis (child-, staff-, and teacher-level)
-

The best weight to use here is a bit more complicated. To answer the question you could use data from several different sources, including the two Family Engagement interviews (parent and family service staff) and/or data from the supplemental items in the Core parent or teacher surveys. If you decide to focus on the responses to questions in the two interviews, you would use F2WT for the family service staff interview data and PE2WT for the parent interview data. PE2WT is also designed to be used when analyzing data from the Family Engagement parent interview together with data from the Core parent survey. You will need to use T2TCHWT if you want to use data from the supplemental items in the Core teacher survey. The data from each of these sources were collected only once, in the spring of the program year (spring 2015).

The three weights will support three different levels of analysis: child-level (PE2WT), staff-level (F2WT for family service staff), and teacher-level (T2TCHWT).

Choosing the best weight for your particular analysis can be challenging at times. You can make this task less difficult and stressful through practice. *Reference slides at the end of this module for additional practice.*

How to Use Weights with Commonly Used Software Packages

- **Use**
 - “WEIGHT” statement in SAS, or
 - “WEIGHT BY xxxxx” statement in SPSS
 - [weight=xxxxx] in Stata
 - **FACES weights sum up to population totals**
 - Can be problematic for some statistical packages
 - Leads to incorrect standard errors and erroneous tests of significance
 - **Normalizing weights**
 - Normalize weights for certain software packages (SPSS)
 - Normalized weights will sum to sample sizes, not population totals
-

FACES weights can be used with most commonly used statistical software packages. Shown on the slide are the weight statements that are found in SAS, SPSS, and Stata. In Stata, the typical statement is [weight] but there are 4 weight types (f=frequency weight; a=analysis weight; i=importance weight; p=probability weight) and the default varies by procedure. For frequencies the default is the “f” weight but it does not use non-integers values, so you will need to specify the “iweight” to get the weight applied accurately if it includes decimal points, which FACES weights do. For regressions, and correlation the default is “a” weight and it works fine with non-integer values. You should consult the software documentation for more information on how to incorporate weights in your analyses.

FACES weights sum to population totals, which can be quite large. Some statistical packages will use this weighted sum for the number of observations in the calculation of variances and standard errors. This can lead to incorrect and very small variances and standard errors and adversely affect tests of statistical significance. If the package you are using for your analyses uses the weighted n instead of the unweighted n or sample size, you will want to normalize your weights. The weight will now sum to the sample size and the unweighted n will be used when calculating variances and standard errors.

How Do I Normalize the Weight?

- Find the **FACES weight** that you want to use in your analysis
 - Identify the number of cases with a positive value for this weight and the sum of the values of this weight
 - Calculate a new weight using the following formula
 - $\text{NewWeight} = \text{FACES weight} * (\text{n}/\text{sum of FACES weight})$
 - The sum of **NewWeight** will be the number of cases (sample size)
-

It is easy to normalize the weight. When you do this, you will be creating a new weight that sums to the sample size rather than to the population total. The new weight is equal to the FACES weight that you are using for your analysis multiplied by the ratio of the number of cases with this weight to the sum of the weight. You can use this weight when estimating means, percentages, and when doing correlational analyses, and the results for these estimates will be the same as when using the original weight. Standard errors for these estimates will now be calculated using the sum of the normalized weight (sample size).

A word of caution, the normalized weight will include non-integer values and some may be quite small. You should be sure that the procedure you are using for estimation purposes does not drop or round very small non-integer values.

Example - Normalizing Weights

- Weight to be normalized: PRA1WT
 - Sum of PRA1WT weights = 769,686
 - Total number of cases with a positive value for PRA1WT: 1,786
 - Normalized weight =
 - $NWPRA1WT = PRA1WT * (1,786 / 769,686)$
 - Sum of normalized weight = 1,786
-

Here is an example, using one of the sampling weights found on the FACES 2014 child file. PRA1WT is one of the three fall 2014 cross-sectional weights that is used to analyze data from the fall parent survey and data from the direct child assessment and/or Teacher Child Report. The values for this weight sum to the approximate number of children enrolled in Head Start in fall 2014. There are 1,786 cases in the data file with a positive value for this weight. Using these two values and the formula described in the previous slide a new weight (normalized weight) is created that now sums to 1,786 or the number of cases with a positive, non-zero PRA1WT weight.

Frequently Asked Questions

- **When selecting a weight, do I have to subset my dataset?**
 - **What happens to cases where there is no positive weight?**
 - **What weights do I use if analyzing a subsample of cases?**
 - **What if I'm running a regression - what weights do I use?**
 - **What weight do I use if I'm using a multi-level model?**
 - **Answers to these questions can be found in an attachment to this module, Sampling Weights: Answers to Frequently Asked Questions**
-

Listed on this slide are some of the questions that researchers often ask about the use of sampling weights. Answers to these and other questions can be found in Sampling Weights: Answers to Frequently Asked Questions can be found in Slides 40-45. Many of these questions are also addressed in the FACES 2014 User's Manual.

Summary

- **Weights should be used when analyzing FACES data**
 - **The appropriate weight should be selected based on:**
 - Level of analysis
 - Wave(s) of data
 - Source(s) of data
 - **Weights adjust for unit nonresponse, but not for item nonresponse**
 - **There may not be a “perfect” weight for some analyses. The decision about which weight to use can be informed by some descriptive analyses**
-

You should always use weights when analyzing any FACES data. This helps to ensure that your findings generalize to the populations represented by the different FACES samples (Head Start programs and centers, classrooms and teachers, family service staff, and children) and is one way that you can reduce the potential for nonresponse bias in your findings. You should use the weight that is appropriate for your research question, taking into account the level of analysis, waves of data, and the sources of the data you will be using.

FACES weights adjust for unit nonresponse (an entire instrument missing), but not for item nonresponse (items data missing for an otherwise completed instrument). For the most part, there are very little missing item data in the FACES. However, you should evaluate the amount of missing data on the variables you are using and decide how you want to address it in your analysis.

Various descriptive analyses, such as comparing the unweighted sample sizes for different weights and examining the unweighted distributions for key variables to identify the level of missing data, are useful tools when deciding between two or more weights. You may want to choose a weight that will maximize the number cases and the number of cases with complete data that are available for your analysis.

Topic 4.2

Variance Estimation

Overview

- **Why are variances and standard errors important?**
 - **Why not use standard errors that assume a simple random sample (SRS)?**
 - **How to use “design-based” methods for estimating standard errors**
 - **How to use approximation methods for estimating standard errors**
-

Data from FACES 2014 are used to produce population estimates for characteristics of Head Start, including its programs, centers, classrooms, teaching and non-teaching staff, and children. At each level of sampling (for example, program, classroom and child), the sample selected is only one of many that could have been chosen. Recognizing that estimates derived from a single sample may not represent the true value in the population from which the sample was drawn, researchers often compute the variance and standard error for the population estimates. These are used when placing confidence intervals around point estimates and when testing the statistical significance of group differences.

This section reviews reasons for estimating population variances and standard errors. It examines differences in standard errors for estimates derived from simple random sample versus complex sample designs. It describes design-based and approximation methods for estimating standard errors for complex sample designs like the one used in FACES.

Why are variances and standard errors important?

- **Data from FACES are used to estimate population characteristics**
 - Standard error is a measure of the variability or precision of those estimates
 - **Data from FACES are often used to test hypotheses**
 - Differences between two or more groups of children, classrooms or programs
 - Relationships between characteristics of children and/or their environments
 - Relationships between classroom quality and classroom and teacher characteristics
 - **Standard errors are used when testing whether the differences or the relationships are statistically significant**
-

FACES is a sample survey and the data collected from sample participants are used to estimate population characteristics (for example, average receptive and expressive vocabulary skills of Head Start children, percentage of Head Start teachers who have received training on their curriculum). Standard error is a measure of the variability of these estimates. It indicates how close an estimate derived from one sample is to the actual value in the population. The smaller the standard error the more precise the estimate.

Researchers use FACES data to test a wide range of hypotheses about the characteristics of two or more groups of children, classrooms or programs, and the relationships between two or more variables. Standard errors are used when testing research hypotheses and when making inference to a population from the sample results.

Estimating Variances and Standard Errors

- **Standard procedures in commonly-used statistical software packages often assume data are from a simple random sample**
 - **These procedures are not appropriate for complex, multi-stage, clustered sample designs with unequal probabilities of selection such as FACES**
 - **Software packages designed for simple random samples tend to underestimate the standard errors for complex sample designs**
-

Because standard errors play such an important role in research, it is important that they are accurate and that they capture the key features of the sample design. Many of the standard procedures found in commonly-used statistical packages assume that the data are from a simple random sample where the cases are assumed to be independent from one another and every case had an equal probability of being selected.

Classrooms and children in the FACES sample are not independent given the clustered nature of its design (for example, classrooms are clustered in program/centers and children are clustered in classrooms) and not all sampling units had the same probability of selection. Thus, standard procedures are not appropriate for estimating standard errors generated from the complex sample design used in FACES. Software packages designed for simple random samples would tend to underestimate the standard errors for estimates derived from FACES data.

Using inaccurate standard errors when testing research hypotheses can lead to identifying statistically significant results where none are present (Type I Error) and vice versa (Type II Error). Using the correct standard error will help to reduce the chances of both types of error.

Methods for Estimating Variances and Standard Errors for Complex Sample Designs

- **Design-based variance estimation methods**
 - Replication methods
 - Taylor Series methods
 - **Approximation method**
-

When estimating variances and standard errors for estimates derived from FACES, researchers should use a method that accounts for the FACES sample design. Over the years, FACES has used two design-based variance estimation methods. Prior to FACES 2006, a replication method was used and the data files included a set of replicate weights that researchers would use when estimating standard errors. Since then, FACES has used a Taylor Series methodology and data files include the design variables that are needed to calculate these weights.

It is strongly recommended that you use the Taylor Series method when estimating variances and standard errors for FACES 2014 (and for FACES 2006 and 2009 as well). However, if you do not have access to software that supports this method, you may use a method that approximates the impact of the design on these estimates. The method is simple to apply and the information you need to do so is included in Chapter VI of the User's Manual.

Design-Based versus Simple Random Sample Standard Errors - FACES

Characteristic	Percent	Standard Errors	
		Standard Method	Taylor Series
All Children			
Overweight	17.1	0.68	0.80
Obese	16.6	0.67	0.79
White Children			
Overweight	17.7	1.57	1.74
Obese	13.4	1.40	1.56
Hispanic Children			
Overweight	17.4	1.12	1.28
Obese	22.9	1.24	1.46

The table on this slide shows the standard errors for the percentage of children who are overweight and obese using the standard approach found in SAS (an approach that assumes simple random sampling) and the standard errors when the Taylor Series method is used. The table uses data from FACES 2009, but is included to illustrate the point.

As you can see, for children as a whole and for children in different racial/ethnic groups, the standard errors that are calculated using the standard method are smaller than those calculated with the Taylor Series method.

Design-Based Methods and FACES 2014

- **Taylor Series**
 - Use STRAT (sampling strata) and PSU (primary sample units) IDs from data file
 - Software
 - SUDAAN
 - STATA
 - SAS (using PROC SURVEY procedures)
 - IBM SPSS
 - AM
-

Section E (variance estimation) in Chapter VI of the FACES 2014 User's Manual includes the specifications and code that you would use to calculate Taylor Series standard errors in SUDAAN and SAS. For each, it shows the weight, stratum and PSU variables that you would use for analyses conducted at the program, center, classroom, teacher and child levels. For analyses at the program level you need to specify the sampling weight (D2WT) and the first-stage sampling strata (STRAT). For analyses at the center, classroom or teacher levels, you need to specify the weight that is appropriate for your research question and both the strata and PSU variables (STRAT_C and PSU_C). For child-level analyses you need to use the appropriate weight and STRAT and PSU.

Special procedures in other software packages can also be used to generate accurate estimates for FACES 2014. Some of these are listed on this slide.

Approximation Method

Two steps:

1) Normalize weights

- May need to adjust weights so standard error is based on actual sample size rather than population size

2) Use design effect (DEFF) to account for complex sampling design

- Adjust variance by DEFF and standard error by square root design effect or DEFT
 - DEFF is the ratio of the variance of dependent variable produced by specialized software that accounts for the complex design to the variance of the same dependent variable produced using procedures that assume simple random sampling
-

Whenever possible you should use the Taylor Series method to calculate standard errors for weighted estimates derived from FACES 2014 data. However, if that is not possible, you can use an approximation method. This approach is simple and straight forward, and involves adjusting the standard error generated by standard software (simple random sample standard error) by the design effect (DEFF), which is the ratio of the variance of a variable produced by specialized software that accounts for the complex design (design-based method) to the variance of the same variable using procedures that assume simple random sampling.

If the software package you are using uses the weighted n (sum of the weights) rather than the unweighted n (sample size) when estimating variances and standard errors, you will need to normalize the weights and use this new weight to compute the standard errors before adjusting for DEFF. The normalized weight is created by multiplying the weight found on the FACES data file that you are using for your analysis by the ratio of the sample size (number of cases with a positive value on the FACES weight) to the sum of the values for the same FACES weight. The new weight (normalized weight) will sum to the sample size.

Caution! Certain procedures in available software (for example, proc univariate in SAS) will provide an erroneous number of cases if there are cases with a normalized weight equal that rounds to zero. Always check to confirm that the number of cases in your output is equal to the sum of the new weight.

When estimating the variance of an estimate you will adjust the variance by DEFF. When estimating the standard error you will adjust the standard error by the square root of the design effect or DEFT. Average overall FACES 2014 design effects (means and medians) for all levels of analysis and for select subgroups are provided in the FACES 2014 User's Manual (See Chapter VI, Section F, pgs. 201-207).

FACES 2014 Design Effects: Spring 2015

Type of analysis (weight)	Mean design effect	Median design effect
Program (D2WT)	1.77	1.90
Center (C2WT)	1.81	1.77
Classroom (O2CLSWT)	2.24	2.11
Teacher (T2TCHWT)	1.98	1.89
Child/Parent (P1_RA1WT)	2.32	2.28
Child/Parent (P21RA2WT)	2.42	2.14

This table shows the mean and median design effects that you would use when analyzing the spring 2015 Core study data. Other tables in the User's Manual include DEFF for different subgroups of children. There is no one correct DEFF to use when analyzing these data, but you should choose one that best aligns with the variables and subgroups in your analysis.

3 Ways of Using the DEFF

- **Multiply the simple random sample (SRS) variance by the DEFF or the SRS standard error by the square root of the DEFF (i.e., DEFT)**
 - **Adjust the t-statistic by dividing it by the square root of the design effect (DEFT) or adjust the F-statistic by dividing it by the DEFF**
 - **Adjust the sampling weight such that an adjusted standard error is produced**
-

There are three ways that you can use DEFF or DEFT to adjust the simple random sample (SRS) variances and standard errors. You can multiply the SRS variance and standard error by DEFF and DEFT, respectively. The adjusted variance and standard errors would then be used when calculating the test statistic. Alternatively, you could adjust the t-statistic by dividing by DEFT or the F-statistic by dividing by DEFF. Lastly, you could adjust the sampling weight by dividing the sampling weight (or the normalized weight) by DEFF and then use this adjusted weight when calculating variances and standard errors.

Regression Analysis with Complex Sample Designs

- **Use a software package that supports the use of the Taylor Series method such as AM, WESVAR, SUDAAN, or SAS (SURVEYREG procedure).**
 - **For SAS (PROC REG procedure) adjust the standard errors by DEFT or use a DEFF-adjusted weight.**
 - **For SPSS, normalize weights and adjust standard errors by DEFT or use a DEFF-adjusted, normalized weight.**
-

Many software packages allow the use of sampling weights when doing regression analyses. However, when conducting regression analyses using FACES 2014 data, you want to use a software package that uses a Taylor Series method when calculating standard errors of the regression coefficients. Listed on this slide are several packages or routines within packages that do this. Again, if you do not have access to such a package, you could adjust the standard errors for packages that assume simple random sampling or adjust your sampling weight by DEFF and use this weight in our analyses. Once again, you may need to normalize your weights as well.

Summary

- **FACES 2014 uses a complex sample design and not a simple random sample**
 - **The standard errors should be adjusted for the complex design either using software or the design effects**
 - **The standard error calculation should be based on sample total rather than population total**
-

The sample design used to select the FACES sample is complex and is not a simple random sample. You should adjust for the complexities of the sample design when estimating standard errors by using the Taylor Series method or adjusting simple random sampling standard errors by the design effect (or the square root of the design effect or DEFT). You are strongly encouraged to use the Taylor Series method. Failing to adjust the standard errors by using one of these approaches can lead to inaccurate findings.

Some software packages use the weighted n (population count) rather than the unweighted n (sample size) when calculating standard errors (and variances). You need to be on the alert for this as it can lead to extremely small standard errors and incorrect test statistics.

Module 4 Review

- **To review some of what we have learned about sampling weights and their use and about how to compute variances and standard errors for estimates generated from FACES 2014 data, please answer the questions in the Module 4 Review Quiz.** *Reference slides at the end of this module.*
 - **Once you have answered all the questions, check your answers using Module 4 Review Quiz Answers.** *Reference slides at the end of this module.*
-

Resources

- **Dataset documentation and access via *Research Connections***
 - FACES Series
 - FACES 2014
 - **FACES 2014 User's Manual**
 - Chapter II – Sample and Sampling Design
 - Chapter VI – Data File Content, Structure and Use
 - Appendix A – Elements of the FACES Design and Key Measures used (and Child Outcomes Captured):
FACES 1997 – FACES 2014
 - **FACES 1997 – 2009 user's manuals**
-

Research Connections (www.researchconnections.org) provides users with extensive documentation on both the FACES dataset series: <https://www.researchconnections.org/childcare/studies/36643>. This documentation includes the User's Manual, questionnaires, and codebooks. These pages also provide information on accessing the datasets.

The topics covered in this module are discussed in the FACES 2014 User's Manual and in earlier FACES manuals. A detailed description of the complex design used to select the samples for the Core studies and Family Engagement Plus study is found in Chapter II of the Manual. Discussion of sampling weights and variance estimation is found in the following sections of the Manual:

- Sampling weights (See Chapter 2, Section B).
- Choosing the best weight (See Chapter 2, Section C).
- Variance estimation (See Chapter 2, Section E).

Appendix A contains a summary of key features of the FACES sample design and how it has changed and stayed the same over the years.

More details on the sample designs used in earlier rounds of FACES can be found in the FACES 1997 – 2009 user's manuals.

FACES 1997-2009 data files also contain a set of sampling weights, which you can learn more about by reading the user's manuals for those rounds of FACES. A different method was used to calculate standard errors in these earlier rounds of FACES (replication method). You can learn more about this method and its use by reading the manuals.

Frequently Asked Questions Sampling Weights

FAQ

1. When selecting a weight, do I have to subset my dataset?

ANSWER: There is no need to subset your dataset when using one of the weights on the FACES data files. Cases that are not assigned a valid weight (a non-zero, positive weight) have a missing value for the weight or a zero and would drop out of your analyses or contribute nothing to your estimates. That said, you could be cautious and select only those cases with a non-zero, positive weight.

2. What happens to cases where there is no positive weight?

ANSWER: Cases with a zero value or missing value for the weight variable do not contribute to the weighted estimates (for example, the case makes no contribution to either the numerator or denominator of a weighted average or weighted percent distribution). As noted above, a cautious approach is to select only those cases with a non-zero, positive weight when creating your analysis file.

FAQ

3. What weights do I use if analyzing a subsample of cases?

ANSWER: FACES weights were not designed for estimates of any particular subpopulation. Therefore, there is no guarantee that the weighted sum for the sample subpopulation will match that subpopulation in the FACES reference population (Head Start children enrolled in fall 2014), but the weights should still provide unbiased estimates of means, percentages, etc. for the subpopulation. You should select the weight that best matches your research question, following the same rules that were outlined in Module 4, Topic 4.1.

One thing to keep in mind when analyzing a subpopulation (such as children enrolled in their first year of Head Start or children who stayed in the same classroom between fall and spring): it is best to read in the full sample file and specify in your statistical software the subpopulation to be analyzed (see bottom of page 164 of the FACES User's Manual), rather than just keeping the cases you are interested in. That way, the complex sample design will be fully accounted for when calculating variances and standard errors.

4. What if I'm running a regression - what weights do I use?

ANSWER: When running a regression analysis or any correlational analysis, you should select the weight that best aligns with your research question, following the same decision rules that were outlined in Module 4, Topic 4.1. You should choose your weight taking into account the rounds of data that you are using, the sources of the data, level of analysis and so forth.

FAQ

5. What weight do I use if I'm using a multi-level model?

ANSWER: Researchers will use multilevel modeling to examine the associations between Head Start classroom or program characteristics and child outcomes. Because FACES child weights take into account earlier stages of sampling (programs, centers, classrooms), you would only use a single child weight at level 1. That is, you do not need to use a child weight for level 1 and a different weight (for example, a classroom weight) at level 2. If you are doing a center- or classroom-level analysis, you would only need to weight the data with the appropriate center or classroom weight, respectively.

6. Does the use of imputed data affect how the weights are used?

ANSWER: Several factors must be considered when determining the weight to be applied when using imputed data. Were the missing data within an instrument (item-level missing) imputed and/or were all data missing because of the absence of a survey response (missing as a result of unit nonresponse)? The weights take care of "unit nonresponse," meaning that we are missing an entire instrument, but do not take care of "item nonresponse," meaning specific items being missing from an otherwise complete instrument. Some items are missing due to deliberate skip patterns, while others are missing due to refusals or "don't know" responses, so you have to disaggregate the two. Unless you statistically impute values for the refusals and don't know responses, cases missing one or more variables in your analysis will simply drop out of the estimate, potentially introducing bias. This is probably not a major concern for small amounts of missing data, but could be for larger amounts. If you have larger amounts of missing responses for a categorical response, you might want to create a missing value category with a unique designated (non-missing) value, and run your analysis that way.

FAQ

If the imputation was for an entire instrument (that originally was unit-missing), the use of a weight that includes that source instrument will likely result in the case still being dropped as the value of the weight for cases with unit missing is set to system missing. Here, you would want to use a weight that adjusts for the probability of selecting the case and perhaps for attrition in a longitudinal design, but that is not conditional on having a complete instrument for the data that are being imputed. If the imputation was done for individual items within an instrument (item-missingness), then that case would have a positive value for the weight. The use of weights in that scenario has no impact on your imputed data; it will weigh the case relative to its strength in the population based on its probability of selection.

7. For propensity score analysis, should weights be used at the point of creating the propensity score as the propensity score is sample specific?

ANSWER: At the time of creating a matched sample (either estimating the propensity score or using the propensity score to produce a matched sample), weights are not needed. The focus is to select the best match for each treatment unit for internal validity and matching as many treatment units as possible to have external validity. However, once you start analyzing the matched sample you would want to apply weights for generalizability to a larger population

FAQ

8. How are weights used in analysis using M-Plus or R?

ANSWER: M-Plus does allow the use of weights and the WEIGHTS option normalizes the weights automatically. For additional information for using weights related to variance estimation, please refer to <http://www.statmodel.com/cgi-bin/discus/discus.cgi> (in particular the section on “multilevel data/complex samples”)

R can use weights if applied manually, but whether a given function in the package uses weights depends on the specific package you are using. Similarly, whether the particular command requires the analyst to normalize the weights (as compared to R doing it automatically) depends on the particular package. For variance calculations, design-based methods (replication, Taylor series) are supported by R, but again that depends on the package. The R survey package supports the Taylor Series method. Ultimately, the analyst should review the specific package documentation to determine if and how the functions within it handle weights.

Exercise
(with answers)

**Which weight is the most
appropriate?**

Exercise Q&A

Instructions: Based on the research questions below, determine at what level the analysis will be conducted (for example, child, program, classroom), whether one (fall 2014 or spring 2015) or two (fall 2014 and spring 2015) waves of data are needed to support cross-sectional or program-year analysis, sample units that will be used (for example, program, classroom, teacher, child) and from what source(s) the data come (for example, child assessment, Teacher Child Report, family service staff). Using these pieces of information, determine which weight is the most appropriate.

1. Research question: When children enter Head Start, do their language, literacy, and early math skills differ by gender and race/ethnicity?

Level of analysis: Child

Wave(s) of data: Fall 2014

Sample units: Children who were attending their first year of Head Start in fall 2014 (NEWT0HS=1)

Source(s) of data: Child assessment, Parent survey

Suggested weight: Fall 2014 cross-sectional weight P1_RA1WT. Because we don't need TCR data, we can be less restrictive and use the weight that includes either TCR or child assessment data.

2. Research question: In the fall of the program year, do Head Start children's social skills differ by home resources such as household income and number of books in the home?

Level of analysis: Child

Wave(s) of data: Fall 2014

Sample units: Children who were attending their first or second year of Head Start (NEWT0HS=0,1)

Source(s) of data: Parent Survey, Teacher Child Reports

Suggested weight: Fall 2014 cross-sectional weight P1_RA1WT. Because we don't need child assessment data, we can be less restrictive and use the weight that includes either TCR or child assessment data.

Exercise Q&A

3. Research question: Do the gains children make in their language, literacy, and math skills during their first or second year of Head Start differ by their skill level at program entry?

Level of analysis: Child

Wave(s) of data: Fall 2014, Spring 2015

Sample unit: Children who were attending their first or second year of Head Start (NEWTOHS=0,1)

Source(s) of data: Child Assessment

Suggested weight: Program year weight PRA12WT. We are looking at gains across the first year or second year of Head Start, so we need a program year child-level weight. Because we only need child assessment data we can use a less restrictive weight that does not require teacher survey and observation data.

|

4. Research question: What is the relationship between children's fall-spring gains in literacy and the quality of their Head Start classroom?

Level of analysis: Child

Wave(s) of data: Fall 2014, Spring 2015

Sample Units: Children who were attending their first or second year of Head Start (NEWTOHS=0,1) and whose classrooms were observed in Spring 2015

Source(s) of data: Child assessment, classroom observation

Suggested weight: Program year child-level weight PRA12OCW.

5. Research questions: What percent of Head Start teachers have a Bachelor's degree? More than 3 years of experience teaching in a Head Start program?

Level of analysis: Teacher

Wave(s) of data: Spring 2015

Sample units: Head Start teachers

Source(s) of data: Teacher survey

Suggested weight: Spring 2015 teacher cross-sectional: T2TCHWT. We use this weight rather than T2CLSWT because the question is asking about teachers, not about the qualifications of teachers in Head Start classrooms.

Exercise Q&A

6. How does Spanish-speaking dual language learners' growth in English language skills compare to their growth in Spanish language skills during their second year of Head Start?

Level of analysis: Child

Wave(s) of data: Fall 2014 and Spring 2015

Sample units: Children who were attending their second year of Head Start (NEWTOHS=0)

Source(s) of data: Child assessment, Parent survey

Suggested weight: Program year weight PRA12WT.

7. What is the structural quality (e.g., child-to-teacher ratio) and process quality (e.g., instructional support) of Head Start classrooms?

Level of analysis: Class

Wave(s) of data: Spring 2015

Sample units: Head Start classrooms and their teachers

Source(s) of data: Classroom observation, teacher survey

Suggested weight: Spring cross-sectional weight TO2CLSWT. If we do not need teacher survey data, we can use the weight O2CLSWT.

8. What is the frequency of literacy and mathematics activities in Head Start classrooms?

Level of analysis: Class

Wave(s) of data: Spring 2015

Source(s) of data: Teacher survey

Sampling units: Head Start classrooms and their teachers

Suggested weight: Spring 2015 cross-sectional weight T2CLSWT. This weight is chosen because the data on classroom activities are collected in the teacher survey and because we are not using any classroom observation data.

Exercise Q&A

9. What percentage of Head Start programs have a director with a Master's Degree or above?

Level of analysis: Program

Wave(s) of data: Spring 2015

Source(s) of data: Program director survey

Sampling units: Head Start programs

Suggested weight: Spring 2015 cross-sectional weight D2WT. This is the only weight when using data from the Core program director survey

10. What training have Head Start family service staff had to prepare them to identify the needs of the children and families that they serve?

Level of analysis: Family Service Staff

Wave(s) of data: Spring 2015

Source(s) of data: Family service staff Interview

Sampling units: Family Service Staff

Suggested weight: Spring 2015 cross-sectional weight F2WT. This is the only weight when using data from the family service staff interview alone.

11. How are families engaged in Head Start and in their children's learning and development at home and in the community?

Level of analysis: Child/parent

Wave(s) of data: Fall 2014 and/or Spring 2015

Source(s) of data: Parent survey, Family Engagement parent interview

Sampling units: Parents selected for the Family Engagement interview

Suggested weight: Program year weight PERA12WT. This is the most inclusive weight when the analysis includes data collected from the Core parent survey and the Family Engagement parent interview.

MODULE 4
Review Quiz & Answers

Module 4 Review Quiz

1. As a rule of thumb, it is always best to use the FACES sampling weight that gives you the maximum number of cases for your analysis.
 - a. True
 - b. False
 2. The sampling weights included with the FACES data adjust for the following (Check all that apply):
 - a. Unequal probabilities of selection
 - b. Parents unwillingness to consent to their child's participation in the study
 - c. Children leaving the study or leaving Head Start after the fall sample is selected
 - d. Teachers failing to answer questions in the teacher survey about their age and salary
 3. The weights included with the FACES data do not adjust for all possible combinations of missing data across the study instruments. You may need to choose a weight that is less than a perfect match for your research question.
 - a. True
 - b. False
-

Module 4 Review Quiz, Cont'd

4. Some software packages use the weighted n rather than the unweighted when calculating standard errors. What effects might this have on your findings? (Choose the correct answer)
 - a. Standard errors would be much smaller than expected
 - b. Standard errors would be small, but estimates of population means, percentages and correlations would be unaffected
 - c. Both the standard errors and all other statistics (for example, population means, percentages, correlations) would be negatively impacted
 - d. This has no real impact on the findings

 5. It is always a good idea to adjust standard errors for design effects when analyzing FACES data even if the standard errors are estimated using a Taylor Series method.
 - a. True
 - b. False

 6. When analyzing FACES data, it is best to use a design-based method to estimate the standard errors of means and percentages, but that is not necessary when running a regression analysis.
 - a. True
 - b. False
-

Module 4 Review Quiz Answers

- 1** *ANSWER: False. You should choose use the weight that is best suited for your research question, taking into account the unit of analysis, data sources and whether you will be using one or two rounds of data. When deciding between two weights that seem appropriate for answering your research question, you may want to consider how much missing data you are willing to accept, but this should not be your major reason for choosing the best weight.*

 - 2** *ANSWER: The correct answers are a, b, and c. FACES weights adjust for different probabilities of being selected to participate in the study and for nonresponse among those selected. Parents' unwillingness to consent for their child to be in the study is one form of nonresponse (this affects the child consent or participation rate). The spring 2015 cross-sectional child weights and program year weights both adjust for sample attrition (children leaving the study) and for children leaving Head Start. Children who leave Head Start are no longer considered to be part of the population that is being studied.*

 - 3** *ANSWER: True. Not all combinations of instruments and nonresponse are adjusted for in the weights that are included on the data files. There may not be a perfect weight for your specific analysis. When this is the case, you need to decide how much missing data you are willing to accept for specific instruments and variables in your analysis. It is often a good idea to compare the results from analyses using different weights when you have two to choose from and to examine the amount of missing data on your key variables.*
-

Module 4 Review Quiz Answers Cont'd

- 4** *Answer: The correct answer is b. Standard errors would be underestimated and test statistics such as t-values would be inflated, but this has no effect on the estimates for other statistics (means, percentages, correlations, regression coefficients). The latter would be estimated correctly.*
- 5** *Answer: False. The Taylor Series method adjusts for the complex sample design used by FACES. There is no reason to make any future adjustments to these standard errors.*
- 6** *Answer: False. Design-based methods should be used to estimate the standard errors of regression coefficients. If your software does not support Taylor series, the standard errors should be adjusted by the design effort using one of the three approximation methods discussed in Module 4.*
-