

***IDEA* Part C Data Collection**

Handling Missing Data When Reporting Race/Ethnicity

(Revised January 2009)

This product is intended as a technical assistance document for data stakeholders who complete data submissions that include race/ethnicity information; it is consistent with *IDEA*, Section 618. The primary use of this product is for implementation of policy related to data collection and reporting. OSEP reviewed the document, and it is updated and distributed by the Data Accountability Center (DAC). Please direct your questions about this information to DAC at IDEAdata@westat.com.

States should use the following guidelines only if they have NOT implemented the *Final Guidance on Maintaining, Collecting and Reporting Racial and Ethnic Data* to the U.S. Department of Education, published in the *Federal Register* October 2007 (Vol. 72, No. 202, available at <http://www.gpoaccess.gov/nara/index.html>). As described in the new guidance, the procedures for collecting, aggregating, and reporting race/ethnicity using seven (7) categories must be implemented by no later than the report of the 2010-11 *IDEA* data. Although not required, states may, as early as for reports referencing reporting period 2008-09, collect and aggregate their *IDEA* data as specified in the new guidance. When implementing the new guidance, states must do so at the beginning of the relevant reporting period and must apply the guidance consistently to all *IDEA* data reported within that period.

States are required to report race/ethnicity data for all infants and toddlers reported for the child count (Table 1), settings (Table 2), and exiting (Table 3) data collections. There may be occasions, however, when the racial/ethnic background of some children is unknown. If a state has NOT implemented the new seven (7) race/ethnicity categories, it should use the following guidance for missing race/ethnicity data.

One method for addressing missing race/ethnicity data is to apply the racial/ethnic distribution of those children whose race/ethnicity is known to the children whose race/ethnicity is unknown. This method is described in more detail below.

Please note that OSEP is not asking states to assign a race/ethnicity to individual students. Rather, states are asked to estimate missing racial/ethnic data for aggregate data reporting only. When using this estimation method, states should provide a comment to describe the number of students in each reporting category for whom race/ethnicity was estimated.

The second method for addressing missing race/ethnicity is to omit children/students for whom race/ethnicity is unknown or who have multiple race/ethnicity status from the race/ethnicity cross-tabulations. Note that while a state may opt to omit children/students from the race/ethnicity cross-tabulations, they must be included in the row and column totals. This practice will produce red cell errors and requires a Data Note to explain the omission (including a breakdown of the number of omitted children/students for whom race/ethnicity is unknown vs. those who are multiracial).

Levels of Data Reporting

States should estimate race/ethnicity where the missing data occur in terms of reporting level and reporting category.

- (1) Reporting level. Estimate missing racial/ethnic data at the local program level rather than at the state level whenever possible. When estimating missing program-level race/ethnicity

data, use the program-level racial/ethnic distribution of the reporting category. If local program-level data are unavailable and a different reporting level is used, then estimate race/ethnicity using the racial/ethnic distribution of that reporting level.

- (2) Reporting category. Estimate race/ethnicity using the racial/ethnic distribution of the reporting category with the missing data. For example, if missing data occur in one of the program setting categories, then use the racial/ethnic distribution of that specific program setting (e.g., home). For exiting, use the racial/ethnic distribution of the specific basis for exiting (e.g., Part B eligible).

Multiracial Infants and Toddlers

These estimation procedures are not ideal for multiracial infants and toddlers. The race/ethnicity of these children is known, so, when possible, report them in one of the five racial/ethnic categories based on their racial/ethnic background. Only estimate race/ethnicity for multiracial children if individual race/ethnicity data are not available (e.g., an aggregate collection that uses a single multiracial category).

Degree of Missing Data

Before estimating race/ethnicity, states should investigate the degree of missing racial/ethnic data. This estimation method assumes that the racial/ethnic distribution of the children whose race/ethnicity is unknown is the same as the racial/ethnic distribution of the children whose race/ethnicity is known. This assumption may or may not be true. The larger the percentage of children with missing race/ethnicity data, the less likely it is this assumption is appropriate. As a result, inaccurate racial/ethnic estimates may occur. Whenever there is a large percentage of missing data, states should investigate why the data are missing and take appropriate action to correct the reporting problem.

Example #1: Child Count Data

States report child count data by age and by race/ethnicity. The number of children reported for the five racial/ethnic groups should sum to the total number of children ages birth through 2 receiving early intervention services.

The following table shows fictitious race/ethnicity data for infants and toddlers receiving early intervention services. In this example, the sum across racial/ethnic groups does not equal the total reported by age. As shown in the last column, the racial/ethnic data are missing for several children.

	American Indian/Alaska Native	Asian/Pacific Islander	Black (not Hispanic)	Hispanic	White (not Hispanic)	Racial/Ethnic Sum	Reported Child Count Total	Unknown Race/Ethnicity
Child Count	158	474	1,501	237	1,580	3,950	4,000	50

To estimate race/ethnicity for these children, calculate the racial/ethnic distribution of the child count based on the children whose race/ethnicity is known.

For example, there are 3,950 children whose race/ethnicity is known. Of these children, 158 are American Indian/Alaska Native, 474 are Asian/Pacific Islander, 1,501 are Black, 237 are Hispanic, and 1,580 are White. Translating the counts into percentages, 4% of the 3,950 children are American Indian/Alaska Native (158/3,950), 12% are Asian/Pacific Islander (474/3,950), 38% are Black (1,501/3,950), 6% are

Hispanic (237/3,950), and 40% are White (1,580/3,950). The known racial/ethnic distribution for the child count is presented below.

American Indian/Alaska Native	Asian/Pacific Islander	Black (not Hispanic)	Hispanic	White (not Hispanic)
4%	12%	38%	6%	40%

Now, apply the racial/ethnic distribution to the children whose race/ethnicity is unknown. To do this, multiply the number of children whose race/ethnicity is unknown by each of the racial/ethnic group percentages to determine how many of those children to assign to each race/ethnicity group. Below is the example.

Child Count		
American Indian/Alaska Native	$50 * 0.04$	2 children
Asian/Pacific Islander	$50 * 0.12$	6 children
Black (not Hispanic)	$50 * 0.38$	19 children
Hispanic	$50 * 0.06$	3 children
White (Not Hispanic)	$50 * 0.40$	20 children
<i>Total = 50 children</i>		

Sometimes, due to rounding, the resulting number of children is one too few or one too many. This minor discrepancy can be easily adjusted. For example, if you have one too few children, identify the prounded value with the largest decimal below 0.5 and round that count up. If you have one too many children, identify the prounded value with the smallest decimal 0.5 or above and do not round that count up.

At the end of this step, you know how many of the unidentified children to assign to each racial/ethnic group. Add these children to the child count as reported by race/ethnicity data (see below).

	American Indian/Alaska Native	Asian/Pacific Islander	Black (not Hispanic)	Hispanic	White (not Hispanic)	Racial/Ethnic Sum	Reported Child Count Total	Unknown Race/Ethnicity
Child Count	$158+2=160$	$474+6=480$	$1,501+19=1,520$	$237+3=240$	$1,580+20=1,600$	4,000	4,000	0

Example #2: Non-Child Count Data

This estimation method can also be applied to non-child count data, including program setting (Table 2) and exiting (Table 3) data. For each data collection, the number of children reported for the five racial/ethnic groups should sum to the total number of children reported for that table. Any discrepancies indicate missing racial/ethnic data.

As discussed earlier, you should estimate missing race/ethnicity data using the known racial/ethnic distribution specific to the reporting category and reporting level where the missing data occur. Use the

distribution of the program setting category for program setting data and the distribution of the exiting basis for exiting data. Apply the estimating procedures at the local program level whenever possible.

The following table shows fictitious race/ethnicity data for the children reported in two of the program setting categories. In each example, the sum across racial/ethnic groups does not equal the total reported for the program setting category. As shown in the last column, the racial/ethnic data are missing for several children.

Program Setting	American Indian/Alaska Native	Asian/Pacific Islander	Black (not Hispanic)	Hispanic	White (not Hispanic)	Racial/Ethnic Sum	Reported Total for Category	Unknown Race/Ethnicity
Home	27	153	342	171	207	900	925	25
Community-based Setting	21	7	49	175	98	350	370	20

To estimate race/ethnicity for these children, calculate the racial/ethnic distribution of the program setting category based on the children in the category whose race/ethnicity is known.

For example, for the Home category, there are 900 children whose race/ethnicity is known. Of these children, 27 are American Indian/Alaska Native, 153 are Asian/Pacific Islander, 342 are Black, 171 are Hispanic, and 207 are White. Translating the counts into percentages, 3% of the 900 children are American Indian/Alaska Native (27/900), 17% are Asian/Pacific Islander (153/900), 38% are Black (342/900), 19% are Hispanic (171/900), and 23% are White (207/900). The known racial/ethnic distributions for the two program setting categories are presented below.

Program Setting	American Indian/Alaska Native	Asian/Pacific Islander	Black (not Hispanic)	Hispanic	White (not Hispanic)
Home	3%	17%	38%	19%	23%
Community-based Setting	6%	2%	14%	50%	28%

Now, apply the racial/ethnic distribution to the children whose race/ethnicity is unknown. To do this, multiply the number of children whose race/ethnicity is unknown by each of the racial/ethnic group percentages to determine how many of those children to assign to each race/ethnicity group. Below are the examples.

Home		
American Indian/Alaska Native	25 * 0.03	1 child
Asian/Pacific Islander	25 * 0.17	4 children
Black (not Hispanic)	25 * 0.38	9 children
Hispanic	25 * 0.19	5 children
White (not Hispanic)	25 * 0.23	6 children
<i>Total = 25 children</i>		

Community-based Setting		
American Indian/Alaska Native	20 * 0.06	1 child
Asian/Pacific Islander	20 * 0.02	0 children
Black (not Hispanic)	20 * 0.14	3 children
Hispanic	20 * 0.50	10 children
White (not Hispanic)	20 * 0.28	6 children
<i>Total = 20 children</i>		

At the end of this step, you know how many of the unidentified children to assign to each racial/ethnic group. Add these children to the program settings as reported by race/ethnicity data (see below).

Program Setting	American Indian/Alaska Native	Asian/Pacific Islander	Black (not Hispanic)	Hispanic	White (not Hispanic)	Racial/Ethnic Sum	Reported Total for Category	Unknown Race/Ethnicity
Home	27+1 = 28	153+4 = 157	342+9 = 351	171+5 = 176	207+6 = 213	925	925	0
Community-based Setting	21+1 = 22	7+0 = 7	49+3 = 52	175+10 = 185	98+6 = 104	370	370	0

The same estimation method can also be applied to missing race/ethnicity data in the exiting data collections.