**Representativeness of Family Survey Data**

**Questions and Answers**

1. **What does it mean that family survey data must be representative?**

The goal in being representative is that the survey results that you receive accurately represent all the families you wanted to survey. To determine representativeness, a state needs to analyze whether families who completed the survey have similar characteristics as the families who were sent the survey.

Take this example: You send your survey out to 200 people, and you received all 200 surveys back. In this case you know that that these data exactly represent the group you wanted to hear from. If, however, you receive 150 surveys back, you might wonder about the 50 people that did *not* return it: would they have answered your questions differently? Do they feel similarly to the 150 that did respond? Do they represent a subgroup that you now are missing? This is the essence of representativeness.

1. **What data should you use to compare your returned surveys to?**

You should compare your returned surveys with your target population. Your target population is the group that you want to survey. If your target population is “all families currently in the program,” you will need demographic information on all families currently in the program. If your target population is “all families who are exiting with at least six months of services” then you will need demographic information on the exiting families with six months of service.

To illustrate the importance of comparison to the correct target population, Table 2 shows how these groups differ according to “length of time in program”. This state surveys families who are exiting the program (this is their target population). When comparing these two groups, we see that higher percentages of families who are exiting have been in early intervention for longer amounts of time. To assure an accurate comparison, this state would want to compare their returned survey characteristics to the target population (exiting families), not to all enrolled families in order to analyze whether their survey is representative according to length of time in service.

**Table 1: Comparing target and total populations by**

**length of time in program**

|  |  |  |
| --- | --- | --- |
| **Length of Time in Program** | **Percent of exiting families** | **Percent of all enrolled families** |
| Less than 6 months | 5% | 15% |
| 6 to 12 months | 10% | 25% |
| 12 to 24 months | 45% | 45% |
| More than 24 months | 40% | 15% |
| **Total** | **100**% | **100**% |

1. **How do you get the demographic information to use for comparisons?**

This depends on the data system’s structure and capacity, as well as the subgroup of families surveyed. Two main methods are (a) linking demographic data to a family’s individual survey, or (b) asking families to fill in demographic information on the surveys. Linking demographic information about a family can be done using a unique identifier or code. The code on the returned survey would then be linked to provide information such as the family’s program, race/ethnicity, length of time in service, or any factors that you want to document. This method allows you to directly compare demographics of those who responded with all families to whom you sent a survey.

For programs without this capacity, an additional source of data is needed for comparison. States may use program demographic data as comparison, even if it is not connected with an individual families’ survey response. Another option is to use external data such as the 618 data tables. For this option you will need to use the data table(s) that best matches your family survey population. For example, if your state surveys all families of exiting children and families fill in race/ethnicity on their survey, you might compare returned survey demographics with the 618 “Percent exiting by race/ethnicity” data table. Although the numbers will not be exact due to different data collection time frames, you might reasonably compare to these data.

1. **What demographic information should a state use to analyze representativeness?**

All states are required to analyze and report family survey data by local program (LEA/Part C program). States can decide what demographic information they will examine in addition to local program. Variables to consider are race/ethnicity, length of time in service, child’s age, disability category, family income, and/or family’s primary language.

Analyzing your data using multiple variables will increase confidence in your results. For example, you may find that responses are proportionate among race/ethnicity groups, but not among family income categories: you may have under-representation of low-income families across all race/ethnicity categories. An example using multiple variables is shown in Table 5.

1. **How would a state analyze their data for representativeness?**

There are several ways to think about analyzing how representative your responses are. Two main ways are looking at response *rates* by subgroup and response *proportions* by subgroup. Examples of each are provided, as well as an example across multiple variables.

**Examine Response Rates.** One way to determine if data are representative is by looking in detail at response rates. Overall response rates are not always a good indicator of representativeness among subgroups. While an overall high response rate might give you confidence in your results, consider the data outlined in Table 1. The overall response rate (73%) appears reasonable, although further analysis shows that two subgroups are not well-represented in your returned surveys (Hispanic and American Indian/ Alaska Native families). Similar analyses could be done to determine if response rates differ for rural or urban families, surveys sent out in English vs. Spanish, etc.

**Table 2.** **Comparison of response rates among families in the**

**target population and families who responded**

|  |  |  |  |
| --- | --- | --- | --- |
| **Race/Ethnicity** | **Number of families sent the**  **survey** | **Number of families who**  **responded** | **Response**  ***rate*** |
| American Indian or Alaska Native | 16 | 3 | 19% |
| Asian or Pacific Islander | 21 | 17 | 81% |
| Black (not Hispanic) | 48 | 43 | 90% |
| Hispanic | 42 | 6 | 14% |
| White (not Hispanic) | 157 | 139 | 89% |
| **Total** | **284** | **208** | **73%** |

**Examine Response Proportions.** Many states also examine the *proportion* of responses in subgroups by comparing these to the target population. If the proportion of families that responded is the same as the proportion of the families in the target population, a state may conclude the data are representative for that group. If the proportions vary, the state must determine if the variation is enough for concern. Again, using multiple variables for comparison will increase the confidence in your results. Tables 3 and 4 provide examples of how a state might look at data proportions by race/ethnicity and eligibility criteria. Results highlight the same finding seen in the previous example: Hispanic and American Indian/ Alaska Native families are responding in lower proportions compared to other race/ethnicity groups.

**Table 3.** Comparison of proportions of responses by race/ethnicity

|  |  |  |
| --- | --- | --- |
| **Race/Ethnicity** | **Percent of families sent the**  **survey** | **Percent of families who**  **responded** |
| American Indian or Alaska Native | 5.6% | 1.4% |
| Asian or Pacific Islander | 7.4% | 8.2% |
| Black (not Hispanic) | 16.9% | 20.7% |
| Hispanic | 14.8% | 2.9% |
| White (not Hispanic) | 55.3% | 66.8% |
| **Total** | **100%** | **100%** |

It is a good idea to look at proportions of responses across several variables. Table 4 shows how subgroups of families are responding at different rates based on their child’s disability. This analysis shows us that fewer families of children with developmental delay are responding to the survey, while relatively more parents of children with an established condition are responding. When particular groups of families don’t complete the survey, we do not know if the data received accurately represents that group.

**Table 4.** Comparison of proportions of responses by eligibility category

|  |  |  |
| --- | --- | --- |
| **Eligibility Category** | **Percent of families sent the**  **survey** | **Percent of families who**  **responded** |
| Developmental delay | 47% | 25% |
| Established condition | 24% | 42% |
| Clinical opinion | 11% | 10% |
| At risk | 18% | 23% |
| **Total** | **100%** | **100%** |

**Examine across categories.** Looking at additional aspects within populations provides more in-depth information about the families that are responding or not responding to your survey. Table 5 shows the race/ethnicity results see in Tables 2 and 3, with the additional information about income. By adding this information, we add to the previous information about American Indian/ Alaska Native and Hispanic families responding at lower rates than other families. Adding the information about income shows that the specific families in these groups that were not responding were low-income families. In addition, we discover that low income Asian families also have not responded at the same rates as middle and upper income families, although this was not evident when looking at ethnicity alone. Looking at family demographics across variables in this case informed us that some low-income families are less frequently responding to the family survey and therefore may not be represented in our data.

**Table 5. Comparison of proportions of responses by**

**race/ethnicity and income**

| **Race/Ethnicity and Income** | **Families sent the**  **survey** | **Families who**  **responded** |
| --- | --- | --- |
| **American Indian or Alaska Native** | **16** | **3** |
| low income | 12 (75%) | 0 (0%) |
| middle/high income | 4 (25%) | 3 (100%) |
| **Asian or Pacific Islander** | **21** | **17** |
| low income | 9 (43%) | 5 (29%) |
| middle/high income | 12 (57%) | 12 (71%) |
| **Black (not Hispanic)** | **48** | **43** |
| low income | 18 (37%) | 15 (35%) |
| middle/high income | 30 (63%) | 28 (65%) |
| **Hispanic** | **42** | **6** |
| low income | 24 (57%) | 1 (17%) |
| middle/high income | 18 (43%) | 5 (83%) |
| **White (not Hispanic)** | **157** | **139** |
| low income | 59 (38%) | 50 (36%) |
| middle/high income | 98 (62%) | 89 (64%) |
| **Total** | **284** | **208** |

1. **What can I do if my data are not representative?**

Some states use data weights to adjust their data to make it more representative of the target population. If certain subgroups of families are either over- or underrepresented, a weight can be used to adjust a family’s contribution to the overall results. A weight is simply a multiplier that increases or decreases a family’s score. In order to determine what weight to apply to your data, consultation with a statistician is recommended.

Another possibility would be for a state to develop and implement improvement activities aimed at increasing either overall response rates, response rates for specific populations under-represented in the data, or both. To improve overall return rate, a state may decide to provide multiple options for families to respond to the survey—telephone interview, online survey, or paper/pencil survey. A state could tailor improvement activities for a specific population as well. For example, state data may show that the response rates are particularly low for families whose primarily language is not English. To increase the response rate for these families the state may decide to improve the availability of translated materials and/or provide interpretation services to support families in completing the survey in various languages. Additional information on improving response rates, along with state examples, can be found in the ECO handout *“Strategies for Increasing Survey Response Rates”*