PHASE IV

Analyzing And Interpreting The Data

Now is the time to make sense of all the data collected. It is important to organize collected data in a manner in which numerical information can be grouped, descriptive information can be summarized, and trends can be identified. Cleaning and coding data can begin before all collection activities are completed.

Step 1: Clean the raw data

- Check data for errors.
- Put aside unreliable and invalid data.

Step 2: Analyze program implementation data

- Code and categorize raw data.
- Review and summarize descriptions about program processes.
- Answer evaluation questions.

Step 3: Analyze data relating to participant outcomes

- Code data.
- Conduct descriptive analysis of raw data.
- Answer evaluation questions.

Step 4: Integrate and synthesize findings of the analysis

- Link patterns found in program implementation data with patterns found in program outcome data.
- Build tables and graphs to summarize findings.
- Decide what information is relevant.

NOTE

Since evaluation projects will differ across districts, this handbook will only outline broad analytical procedures. Refer to Appendix 4 for sources listing more detailed information.

Step 1

Clean The Raw Data

Raw data are the unchanged pieces of information that have been collected from or about students and teachers and/or other program staff.

Raw data include:

- Demographic information, e.g., gender, age, ethnicity, etc.,
- Test scores,
- Grades,
- School attendance and disciplinary records,
- Survey results (Pre and post test scores).

Before any type of description or analysis can begin, the data must be checked for errors. This is referred to as **cleaning the data**. It will require a visual inspection of completed surveys and interviews, school records or other documents under study.

Data cleaning includes asking:

- Are surveys filled out correctly, e.g., is there only one answer per survey question?
- Are survey, interview, and other document materials identified by school location or some other chosen identifier?
- Are interviews completely filled out? If hand written, are they legible?
- Are already existing document records complete and accurate?

Inaccurate or incomplete data must be corrected by gathering follow-up information when possible. Data that cannot be cleaned should be set aside. It is no longer valid. For example, an item on a student survey in which two responses were chosen where only one is needed makes that data item useless. It is not known whether the double response was accidental or intended; it is not possible to know which of the two answers would be the correct one.

Keep a record of the set aside data that cannot be used in the analysis. It will be necessary to report the original sample size that is determined after the data has been cleaned.

Step 2

Analyze Program Implementation Data

Raw data dealing with program processes should be coded, organized into categories, summarized, and interpreted in order to answer evaluation questions. The objective is to compare the program design plan with the actual delivery of the program. This comparison provides a way to determine whether program implementation objectives were met, and if not why.

Task Code Raw Data

To code data is to assign a numerical value to each response of every question on a survey, interview or other data collection instrument. The number assigned to each response of a question represents the answer. The assigned number can have a numerical value as in an age question or it can be representative in name only such as answer categories listed for questions asking for opinions.

- Data that can be tabulated numerically should be entered into a database or spreadsheet. This information comes from questions such as: "How many students were served?" "How much time was actually spent on program curricula within the classroom?" Responses to questions such as these can be easily coded into a database.

- Data derived from scaled-item responses can be assigned a numerical value and treated in a rank order fashion. In an opinion scale, for example, the response 'strongly agree' might be assigned the number 1, 'no opinion' the number 0, and 'strongly disagree' the number 5.

- Narrative data derived from face-to-face interviews with teachers or from teacher survey questions with open-ended space for answers should be compiled in some fashion. Response information can be sorted into broad categories. If appropriate to the design plan, these categorizations can be further assigned numbers and then coded into the database.

Task Conduct a Descriptive Analysis of Raw Data

After data have been cleaned, coded and/or categorized, describe the data in summary form:

- Describe your sample or population. How many respondents are represented in this data? Run frequencies to count respondents by pertinent characteristics such as the number of teachers interviewed per school. If you are working with a sample, is it representative of the larger target population (the total number of teachers who implement the program)? Do any anomalies exist?

- Create frequency counts that answer evaluation questions concerning the number of participants served, the number of participants completing the program, the amount of time spent on delivery of activities, etc.

- Report numerical information in the form of rates, e.g., rate (i.e., per 1000) of attendance, rate of program completion, rate of teacher opinions, etc. It makes it easier to identify patterns that will emerge when further analyzing data.

- Review non-numerical data. If responses deal with idiosyncrasies or glitches observed throughout the program delivery process, describe them in narrative form.

Task Answer Evaluation Questions and Look For Emerging Patterns

Were implementation objectives met? Why or why not? This is the crux of this part of the analysis. Look for patterns that identify elements of 'best practices' and hindrances or other barriers encountered:

- **The identification of differences** between **the actual delivery of services** and the **program design plan** is of significant importance. Differences point to reasons why implementation objectives were met or not met.

- Numerical data can be reported in the form of charts, tables, or graphs. Good visual representations offer a lot of information quickly.

- Non-numerical data can be reported in the form of a narrative. Tell the story of how the program was actually implemented including all the deviations from the intended design plan.

Step3

Analyze Data Relating To Participant Outcomes

In large part raw data pertaining to participant outcome information will be numerical in form. Analysis will, therefore, largely consist of *number crunching*. *The overall objective is to capture any changes in participant knowledge, skills, perceptions, attitudes, intentions, and or behavior as a result of program participation.*

Task Code Data

It is most useful to code data and enter it into a **database**, even if the quantity of data is very small. After checking the data for errors, choose a way to code it. Every possible response for each question on a student survey or other data collection instrument must have a distinct assigned number that represents that answer. Do not repeat assigned numbers within the same question. Frequencies tabulated on responses to a given question count the number of times each answer appears.

Task Conduct a Descriptive Analysis on Raw Data

Once the raw data has been entered into a database, summarize it:

- Count the number of respondents surveyed.
- Run frequencies to count respondents by age, ethnicity, gender, grade level, etc. Is the sample of respondents representative of the target population? Do any anomalies exist?
- Make frequency counts of answers to survey questions.
- Sort answers to survey questions (items) by group characteristics such as gender, ethnicity, age, school, etc.
- Create tables comparing the results from pre and post test responses across various items on the survey.
- Convert number counts into percentages. A percentage reports what portion of the population is involved in the behavior or other measured item.

- Report measures of central tendency. These are measures of average rates in a given population, i.e., the mean, median, and mode of number of participants smoking cigarettes in the last 30 days.

These types of summary descriptions present the data in a format that can make patterns more readily apparent. As stated earlier, tables, graphs, and charts are good visual representations that offer a lot of quick information.

Task Answer Evaluation Questions and Look for Emerging Patterns

What program outcome results can you identify? Do they meet the program's expected outcome objectives? Why or why not. And, do these results indicate whether ATOD use and/or violent behaviors are increasing or decreasing within the district? These are the most important questions of this section of the analysis. Answering these questions as well as the specific evaluation questions will require:

- Looking for patterns that demonstrate changes in participants as a result of program participation;
- Looking at baseline and comparison data and comparing these with program outcome results;
- Looking at individual participants as a group or a subset of groups (i.e., a particular school) to determine whether or not, as a group, program outcome objectives were met.

Analyzing comparison data will show whether things are getting better or worse. Items to compare include:

- A comparison between pre and post test survey results across specific items of information.
- A comparison of program outcome results on specific items of information with previous years evaluation studies. Be sure items have been measured in the same manner.
- A comparison of program outcome results with local or state aggregate data collected on same items of information. For example, past 30-day use of marijuana.

Task: Perform Statistical Measures, When Appropriate

There is no scientific formula to specify whether a given association (an identified pattern) is substantively significant, strong, and important. Whether it is worth reporting ultimately rests on what the evaluation committee intends to report to its audience. There are, however, statistical measures that establish the credibility of identified patterns found within the data at hand.

There are two levels of statistical measures to be considered in order to further substantiate any identified patterns under analysis.

Measures of Association: These are a class of statistical computations and techniques that summarize the relationships (identified patterns) under study. **What is important to remember is that any conclusions drawn from these analyses make no assertions beyond the data at hand.** For example, to demonstrate that the changes in attitudes and/or behavior observed in the initial analysis occurred not because of errors in the sample drawn but because of participation in the program, certain statistical procedures could be used to assert that these observed patterns within the sample did not happen by chance or by some other influence.

Tests of Statistical Significance: These tests estimate the unlikelihood that a pattern identified in the sample could be attributed to sampling error alone.[8] Such tests can only be applied to samples based on random and representative sampling strategies; they do not hold for samples drawn on the basis of convenience. **Tests of significance are used to estimate how much the identified patterns within the sample surveyed are generalizable to the larger target population from which the sample was drawn.**

For example, if initial data analysis shows that self-reported past 30-day use of marijuana has declined from 15% to 14.2% over the past year, simply reporting the .8% difference is meaningless by itself. In order to make a claim that this observed decrease in marijuana use within the sample could be generalized to the larger target population, it must be substantiated by a statistical test of significance.

The complexity of the statistical procedures used depends entirely on the capabilities of the person(s) analyzing the information. If no one on the evaluation team is skilled in measures of statistical analysis, then it is recommended that someone with the necessary skills be hired for this purpose.

[8] Babbie, (1992) *The Practice of Social Research*. 6th ed. p.G8.

Step 4

Integrate And Synthesize Findings Of The Initial Analyses

The summaries of information and the display of patterns for separate items of information must be pulled together into a larger analysis. Of particular relevance is to link analysis of program implementation data with analysis of program outcome data. It is important to be able to tie together what program processes facilitated or hindered the desired program outcomes.

Task: Build Summary Tables and Graphs of Findings

Build a summary table of findings. For example[9],

Worksheet 4.1
Summary of Findings

Specific question	Gist of the information collected	Presence of/ absence of differences	Possible reasons for change	Barriers encountered	Factors that facilitated

[9] The Administration of Children Youth and Families (n.d). *The Program Manager's Guide to Evaluation*, Chap 8.

Task: Decide What Information is Relevant

When analyzing data, evaluators generally learn much more than they set out to learn. Unanticipated relationships might emerge, but remember to stay focused on the task at hand. **Answer your evaluation questions first**. Working on information that is outside the scope of your immediate questions is a separate task and should be set aside until the evaluation goals have been met.

Things To Remember

1. Routine monitoring of data collection will reduce the amount of raw data "cleaning" required.

2. Do not lose focus on the evaluation questions.

3. Select tables, graphs, and charts for further analysis that display the strongest relationships, patterns or indicators.

4. Do not confuse statistical significance with substantive significance; the latter meaning that an identified pattern or relationship is important and worthwhile to report.

5. Statistical tests of significance are useful procedures in the analysis and interpretation of data. However, be careful of trying to interpret statistical results too