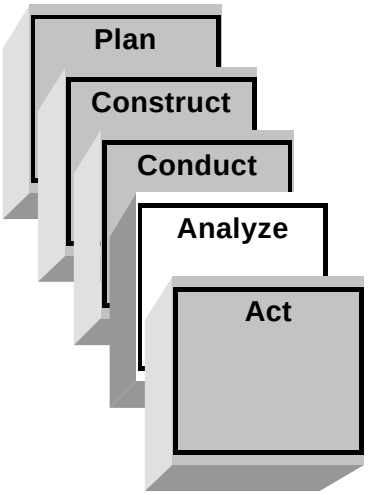
ANALYZE THE DATA



Throughout the customer feedback activity, the framework for analyzing findings should be established and modified. An *analysis plan* is a useful tool for organizing the data analysis. The analysis plan should specify how your organization will analyze the survey responses to produce the desired products. The plan is helpful for making sure that the data you collect will answer the overarching questions being posed, for ensuring that you do not gather extraneous data, and for setting forth expectations about the kinds of information that will result from the customer feedback activity.

You should include two important items in the analysis plan: 1) the designation of dependent and independent variables, and 2) the stipulation of the unit of analysis. A *dependent* variable is the phenomenon you are investigating. For EPA's feedback activities, the dependent variable will likely be the degree of customer satisfaction with a specific product or service. *Independent* variables help explain the observed level of the dependent variable, and may include factors such as differences in the nature of the product or service (e.g., customers were consistently more satisfied with one service than with another), frequency and type of interaction, and customer differences (e.g., educators, students, local planners, and small business owners using the same service). The *unit of analysis* is what you are studying. In customer feedback *surveys* at EPA, the unit of analysis will, in most cases, be the individual *person served*. When you use *continuous feedback* methods, the unit of analysis will generally be the individual *customer transaction*. For further discussion of unit of analysis, see **Factsheet VII**.

DATA CLEANUP

Once you have set up the database and entered all data, you must review the data and prepare data for analysis. This may entail a broad set of activities, such as deleting cases that left all answers blank on a mail survey and coding open-ended responses into categories. Generally, this is the time to a run a set of frequencies to show the number of responses of each kind to each question (the number of yeses and noes to a yes/no question) and the total number of responses of all kinds to each question. This quick analysis gives you a rough check on the completeness and accuracy of your data (the total number of responses to any one question cannot exceed the total number of respondents and rarely will differ greatly from the total number of responses for each of the other questions). Frequencies flag out-of-range values (i.e., responses to one question that are so different from responses to similar questions that you doubt their accuracy).

TYPES OF DATA AND ANALYSES

Data from focus groups tend to be qualitative in nature. Analysts may tabulate data from focus groups, such as "*x* percent of the participants expressed satisfaction." You should treat these numbers cautiously and not generalize them to the full set of customers because 1) focus groups usually have only a relatively small number of participants, and 2) participants may have been recruited because they had specific experiences or characteristics. You may review transcripts from focus groups to detect patterns and inconsistencies or you may apply more rigorous content analysis.

For quantitative you can produce a variety of statistics:

- Descriptions of central tendencies, such as the mean, median, or mode (i.e., the average value, the middle value (half are larger and half are smaller), or the most frequently occurring value).

- Other descriptive statistics, such as frequencies, percentiles, and percentages. In customer satisfaction surveys, the most commonly reported result is of this kind: the percentage of respondents who expressed satisfaction with a specific aspect of their interaction with EPA.

- Cross-tabulations that array independent variables against the dependent variable (for example, type of customer displayed against a summary measure of customer satisfaction, like the percentage of customers of each type who reported being satisfied with the product or service they received).

- Multivariate statistics—such as factor analysis, analysis of variance, and regression analysis—to determine the relationship between and among selected variables.

- Chi-square, z scores, t-tests, and other statistics to determine statistical significance.

- Time-series and trend analyses to determine long-term changes and seasonal and cyclical patterns in the data.

The following table contains information about the statistical techniques that will most likely meet all the needs and expectations of the EPA program or project conducting feedback:

Statistical Technique	Use	Example
Mean	To determine the average response	The mean rating for overall satisfaction is an 8.4. (Sum of all scores divided by number of respondents)
Median	To identify the middle response	The median score for overall satisfaction is a 9. (When responses are listed in numerical order, the middle response [if odd number of respondents] or the average of the two middle responses [if even number of respondents]
Frequencies	The summarize the distribution of responses	67% of respondents rate overall satisfaction a 9 or a 10.
Cross-tabulations	To summarize the distribution of responses by another variable	78% of Maryland respondents rate overall satisfaction a 9 or a 10, compared to 60% of Virginia respondents.
T-Test	To test for statistically significant differences between *two* independent groups	Maryland respondents are significantly more satisfied overall than Virginia respondents.
ANOVA (analysis of variance)	To test for statistically significant differences between *three or more* independent groups	Overall satisfaction differs significantly among Maryland respondents, Virginia respondents, and D.C. respondents.
Correlation	To determine how much responses to one question predict responses to another question (Measures the strength of relationship between variables)	Of all aspects of the office, satisfaction with the cleanliness best predicts overall satisfaction. (Respondents who are satisfied with cleanliness tend to be satisfied overall, and respondents who are dissatisfied with cleanliness tend to be dissatisfied overall)
Regression	To analyze the effects of a relationship among responses to two or more questions (Measures the effects of one or more variables on another variable)	As satisfaction with cleanliness decreases, overall satisfaction decreases.

ANALYSIS: AN EXAMPLE

The following is a simple example of how you might analyze data from customer feedback. Suppose an EPA group has distributed several thousand copies of the ABC Booklet, and because you want to know how satisfied customers are with the booklet, you asked 450 respondents to a survey which included:

> *On a scale of 1 to 6 where 1 represents "highly dissatisfied" and 6 represents "highly satisfied," how would you rate your satisfaction with the ABC booklet you received from EPA?*

If one were to tabulate the all the scores, the average score would be 3. 5. Although an average score is a very important piece of information, there is a lot more you can do with the data from your customers. It is often useful to begin with a frequency distribution where you determine the number and percentage of respondents who gave each score between 1 and 6. Here is one way to present that distribution:

Customer satisfaction with the ABC Booklet (n = 450)		
Score	**Number**	**Percent of those expressing an opinion**
1—Highly dissatisfied	42	11
2	27	7
3	122	31
4	132	34
5	38	10
6 —Highly satisfied	32	8
Don't Remember	22	5
Don't Know	35	8
Total:	393	100

This example points out several items you need to consider. ***First***, of the 450 customers responding to a survey this question, 22 did not remember receiving the booklet and 35 said they had no opinion or did not know how they would rate their satisfaction with the booklet. In the example provided above, the information about those who do not remember or have no opinion is presented outside the table because the analyst decided that it was more important to focus attention on those who did have opinions to express. Thus, the percentages of those with opinions is based on the 393 respondents who expressed opinions. If it is important to determine

the percentage of customers who don't remember or who have no opinion about the booklet, you would calculate those figures using 450—the total number who were asked the question—as the denominator. By including the sample size in the table (the information that n = 450), readers can do these calculations, should they be interested.

Second, the information presented may be at too great a level of detail for many audience members. The difference between a 2 and a 3 rating, for example, may not be meaningful for them. Thus, you may find it useful to *collapse* the information into some smaller number of categories. One possibility is to create three categories: dissatisfied, neutral, and satisfied.

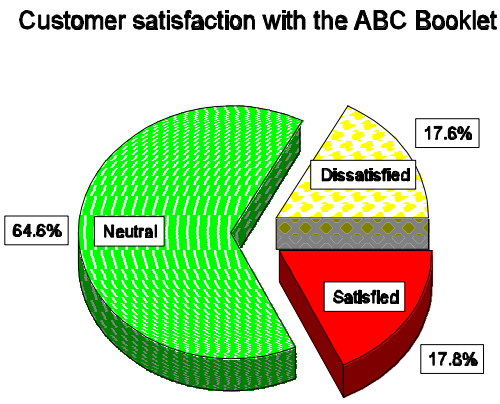
Scores of 1 to 2, 3 to 4, and 5 to 6 might be collapsed to create three categories and then report:

Customer satisfaction with the ABC Booklet (n =450)		
Rating	**Number**	**Percent of those expressing an opinion**
Dissatisfied	69	18
Neutral	254	65
Satisfied	70	18
Total:	393	101*

* Total is greater than 100 due to rounding

Don't remember receiving the ABC Booklet:	22 (5 percent of 450)
Don't know/no opinion	35 (8 percent of 450)

Note that the information can now be grasped much more immediately. It is reasonable to ask: If you will eventually collapse responses, why does the question posed to customers have six possible answers? Research has shown that people answering survey questions prefer to have a fairly wide range of responses because they don't like to feel forced into a limited set of options. In addition, analysts may have different approaches to collapsing categories.



The responsibility for reducing information to a manageable amount falls to the analyst. It is the analyst's task to identify sensible ways to collapse categories and to present these decisions to the audience (often as a footnote or technical appendix).

Third, as discussed in the next section, you should consider how to present the data. Although these

tables are simple and easy to interpret, compare them to a chart that summarizes the information instantly.

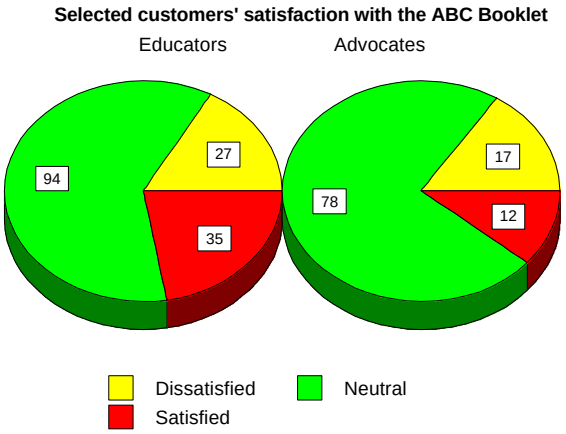
Fourth, the analysis you anticipated during the planning phase of the customer feedback activity should guide you whether you need to do subgroup analysis. Subgroup analysis examines whether different kinds of customers have different kinds of responses. Suppose you want to examine whether educators and representatives of advocacy organizations have the same or different opinions about the ABC Booklet. You could collapse categories and sort respondents by their status as educators or advocates (to be sure, some respondents may be both educators and advocates, but for simplicity, let us assume you had customers indicate their primary role), then present the findings:

Selected customers' satisfaction with the ABC Booklet (n = 450)				
Rating	Educators		Advocates	
	Number	Percent	Number	Percent
Dissatisfied	27	17	17	16
Neutral	94	60	78	73
Satisfied	35	22	12	11
Total:	156	99*	107	100

* Total is less than 100 due to rounding.

This table provides important information to the audience, but you might want to present it using charts for the two separate groups. You could also perform a statistical test to see if the two groups differ statistically in their satisfaction with the ABC Booklet.

The ***fifth*** item to consider is the adequacy of your findings. Be sure how strong your findings are before formulating recommendations. Many factors affect adequacy, such as the sample size, response rate, and objectivity of questions posed—plus the way you will use the findings. With a sufficient sample size, a good response rate (more than 75 percent for mail and telephone surveys, for example), and questions that are not biased, you can use the information with confidence. OMB requires an 80 percent response rate for survey results to be considered statistically valid. However, when less than 80 percent of those sampled return questionnaires in a customer



feedback and satisfaction measurement activity, the information gathered should still be used to improve customer service. Do not ignore the findings.

Let's say that in the above example, there was an additional group of people — small business owners — who were your customers, and that a total of 17 small-business owners responded to your survey. This is a small enough number that the sampling error for this one group of customers may be quite high. Nevertheless, pay attention to the results.

One final comment on this example: EPA has a large number of programs and offices, some of which may have customer bases much smaller than the thousands used in the example. If your customer base is quite small, you first must decide whether a statistical sample and quantitative survey is still viable because other techniques may be more suited for your purposes. If you decide to go ahead with a quantitative survey, recognize that the analyses you conduct should be carefully considered and constructed. If, for instance, you have 500 customers and survey 100 of them, you can perform the same analyses as in the example above, but you should examine the frequency distribution first.

Even if they do not adequately represent the larger group of small-business owners who were your customers, you can still

- Decide whether the findings are suggestive (rather than definitive). Should your office pay attention to the concerns suggested by these findings?
- Compare the findings to other similar data. Are small-business owners generally pleased or displeased with other EPA products?
- Compare the findings to information EPA gets from continuous feedback methods. If you call small-business owners after providing a service or product, what do they have to say in those conversations?
- How do the continuous feedback findings compare with the results of this survey?
- Discuss the findings with colleagues. Have they gotten similar reports? Is there a pattern emerging about small-business owners' level of satisfaction with EPA products?
- Raise the findings with program managers, being careful to note that this *might* be an area that requires attention to improve customers' satisfaction with EPA.
- Investigate the findings further. Should you use this as a starting point for more in-depth discussions with small-business owners? Conduct focus groups to see how products could produce higher levels of satisfaction?

In an extreme case, let's assume that 10 of your 100 respondents gave a score of 0, 60 gave a score of 3, and 30 gave a score of 6. Although the average score of 3. 0 may be close to the average of 3. 5 in the example, the distribution of responses is very different.

"A reasonable probability is the only certainty. "

E. W. Howe

Sixth, you need to consider how past responses compare with the new responses, and to ensure that you can compare the most current results with those you expect to from future questionnaires. This is time series or trends analysis and is vital to being able to measure change.

DRIVER ANALYSIS

An analytical approach that is very useful in customer research is *driver analysis*. Driver analysis identifies the service or services that most significantly affect respondents' satisfaction. This type of analysis provides decision makers with a tool to prioritize findings, which is important because customer feedback efforts often yield more information than an organization can deal with. Also, managers often do not have enough resources to adequately address all aspects of customer service that receive low satisfaction ratings. Driver analysis enables the study team to identify which areas deserve the highest levels of attention.

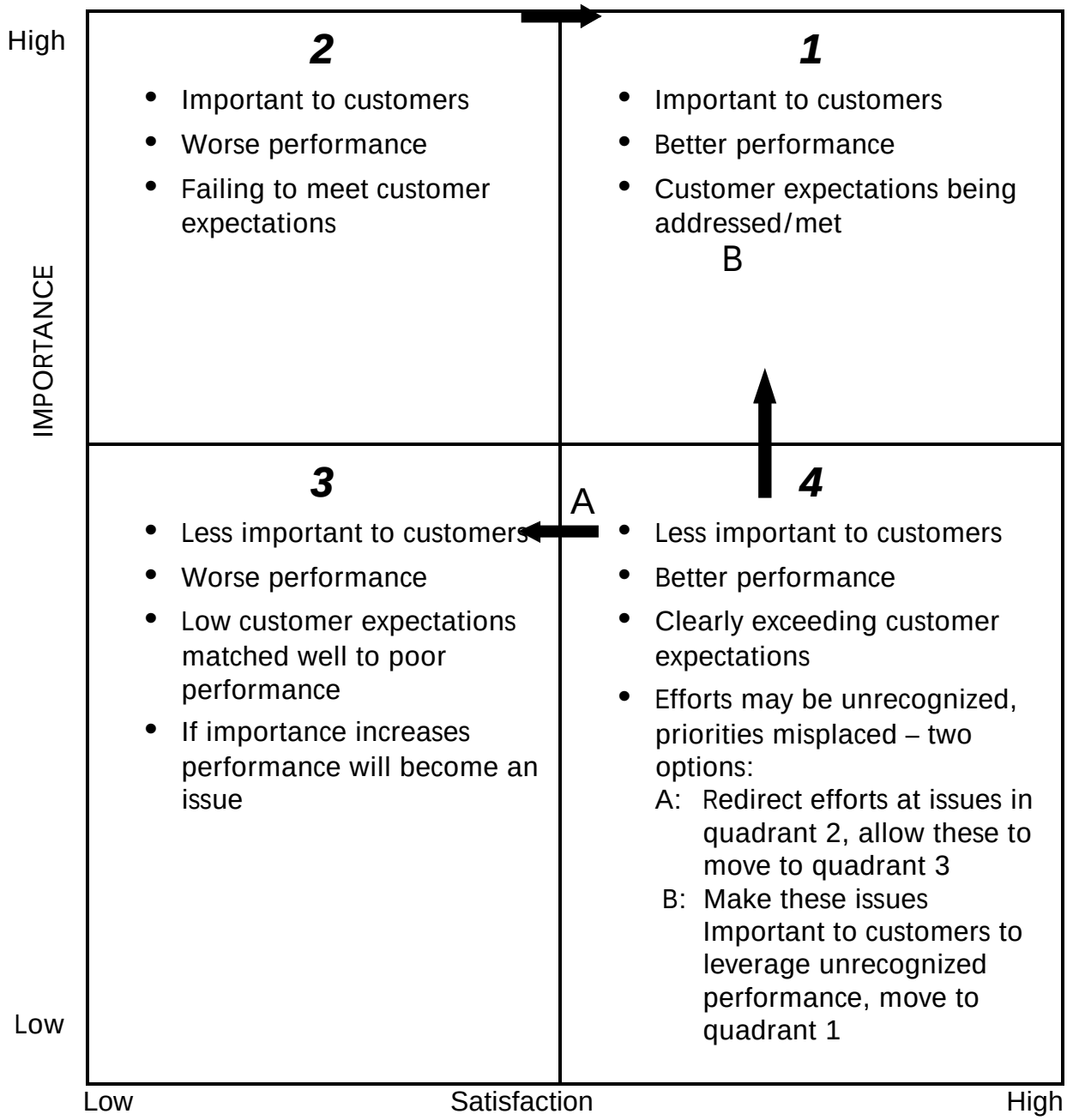
As an example, let us assume that an EPA program is assessing three ways of providing information: by telephone, by mail, and through published materials. Analysis of customer feedback can identify which of these methods results in the highest level of respondent satisfaction. This is the delivery system that most strongly drives satisfaction with the program's products and services. When you identify the method that significantly affects satisfaction, additional analysis can determine which factor within that method most significantly affects satisfaction. Continuing with the example, let us assume that you identify information received by telephone as the method producing highest satisfaction. You can also use driver analysis to identify the factor that most affects the respondent's opinion. Such factors may include one or more of the following: the accuracy of the information, the courtesy shown by the employee, or the accessibility of the correct person to answer the question. Identifying the driver in this way greatly enhances a manager's ability to set priorities for improvement efforts.

Two primary analytical techniques, *stated importance* and *derived importance,* are used in driver analysis:

Stated importance uses respondents' answers to specific questions regarding the importance of the services. Simply ask the respondent to rank or rate items on a prescribed scale (such as a scale from 1 to 6) according to their importance.

Derived importance uses multivariate analysis to identify the most important factors affecting satisfaction. In short, the overall level of satisfaction with the organization is compared to the levels of satisfaction with particular products or services received. Driver analysis will identify the degree to which variation in the overall level of satisfaction is explained by the variation in the product or service received. Those individual products or services that most adequately explain the variation in overall satisfaction are the drivers.

The following table relates a useful method of comparing importance data, such as from driver analysis, and satisfaction data. When the results from a question is plotted according to the levels of importance and satisfaction, some helpful inferences can be drawn.



PRESENTING THE DATA

One critical activity is to remove all identifying information from the data. To ensure credibility and confidentiality, your should never present findings that could be used to identify a specific customer. A typical practice is to strip names, addresses, and telephone numbers from the analytical database and keep them in a separate file that includes the unique identification number assigned during the data collection activity. If ever warranted, you can link the file with identifying information with customer feedback through the identification numbers.

Most people are interested in the "bottom line," presented as succinctly and clearly as possible. Therefore, it may be best to present the data reflecting survey results in simple, straightforward ways to most EPA audiences and save the mathematical details for an appendix or supplementary briefing. Many audience members want a brief summary of the study's findings. Two pages of text, with key findings presented as bullets, are usually sufficient.

Graphic representations of data are powerful displays of findings. It is very easy for audiences to grasp information presented in bar graphs, pie charts, and similar designs. The rapid growth of low-cost color printers means that these displays can be easily produced in color, adding to their ease of understanding. Examples of graphs are presented in **Factsheet VIII**.

FORMULATING RECOMMENDATIONS BASED ON THE DATA

Customer feedback may suggest many potential improvements or enhancements to consider. Narrowing down the list to those that will have the most direct effects on overall customer satisfaction is the ideal. Most organizations will have limited staff and other resources, so practical considerations must guide their choices. Usually, three to five targeted improvements are sufficient. Sometimes, a single improvement can present a significant challenge, and focusing on it can have a major impact.

Each organization will consider its own capacity for action. However, it is important to do something or customers may feel that their input was not valued and the effort they expended to respond was wasted. They may place even less trust in the surveying agency.

Recognize too that not everyone will be ready for the feedback results. Presenting them can raise sensitive issues for some individuals. Some people may feel threatened by anything but glowing results, or become defensive or emotional. Some may question the credibility of the findings, especially if they build logically to recommendations for changes that affect them.

To get buy-in and use the results to influence change, results must be honest, and presented in a constructive way that emphasizes the positives. Results, findings, and recommendations should be presented as opportunities for improvement. If the survey cannot be used to influence change or improvement, it did not meet its objective, no matter how carefully the whole feedback activity was conducted.

PRESENTING RECOMMENDATIONS—USING GRAPHICS

First, remember, at least 70 percent of the message is visual, so take advantage of how people take in information. Use the right visuals to communicate your message. You can

- Emphasize main numerical facts
- Uncover facts, trends, comparisons and relationships that might be overlooked in text or table
- Summarize, group or segment (stratify) data
- Add variety and interest to text, tables, and briefings.

It's best to use *pie* charts to display components or parts of a whole. Use *line* charts when you want to show independent or cumulative values when

- Your data cover a long period of time several series are compared on one chart
- You want to show change, not quantity
- To exhibit trends
- To show relationships
- The plot or the series fluctuates sharply.

Do not use column charts for comparing several data sets, for showing data with many plots, or to show many components. Finally, use picture graphs to demonstrate concepts or ideas. (See **Factsheet VII** for examples of graphics.)

ON DEVELOPING RECOMMENDATIONS

Whether you should develop recommendations depends on the purpose of the feedback activity, the significance of the issues, the quality and significance of the findings, and your audience. Your original purpose should be action-oriented; answers to your issue questions should naturally lead to ideas for actions that would improve program effectiveness.

If you develop and make recommendations, they should be feasible, supported by the findings (which are in turn supported by the data), and stated unambiguously. Providing a list of options for achieving a recommended improvement can increase the likelihood that it will be implemented. Another critical, although sometimes subtle, consideration in developing recommendations is the political climate. It's a fact of life that some

"Measuring fuzzy conclusions to three decimal places is akin to putting a caliper on a dust bunny. "

Kerry Patterson
The Balancing Act

recommendations, no matter how well you support them, will not be accepted by those in authority due to factors beyond your control. Just be aware of these factors so that you can develop an alternative recommendation or recognize that your recommendation may not be implemented until the climate changes or until others have helped tip the scale (from *Practical Evaluation for Public Managers*, Office of the Inspector General, Department of Health and Human Services, November 1994).

Analyze ✓ Checklist

_Clean up data
_Determine the appropriate types of analysis
_Perform analyses
_Present data
_Formulate conclusions and recommendations based on the data
_Present the recommendations