

Innovation *insights*

OFFICE OF PLANNING AND
Institutional Assessment

Innovation Insight Series Number 14

Using Surveys for Data Collection in Continuous Improvement

Across the University many people use data in assessment and decision-making. *Data-based decision-making*, an essential element of continuous quality improvement, helps individuals and teams to assess the efficiency and effectiveness of current processes. Numerous methods exist for collecting data. Focus groups, personal interviews, surveys, review of internal records, and counting events are some methods to gather data. This issue of *Innovation Insights* explains how to use surveys to obtain data.

What is a Survey?

A survey is a research method for collecting information from a selected group of people using standardized questionnaires or interviews. While many people think of a questionnaire as the “survey”, the questionnaire is just one part of the survey process. Surveys also require selecting populations for inclusion, pre-testing instruments, determining delivery methods, ensuring validity, and analyzing results.

In continuous quality improvement, surveys help to identify customer expectations, measure satisfaction levels, and determine specific areas for improvement.

Is a Survey Necessary?

Surveys that provide valid, usable results require thought, planning, logistical support, time and possibly, money. Consider first whether the data are available from other data sources or collection methods. For example:

- rather than asking employees how many times they used a certain service within the University, internal units may be able to provide the number from their records, or
- to determine the number of students using tutoring services, sign-in sheets could be located at each area.

Other sources of information might include: published reports, such as Penn State’s Factbook or the Common Data Set; data from Penn State’s Data Warehouse; previous surveys of students, staff or faculty (e.g., the 2004 Faculty/Staff Survey or the Penn State Pulse student surveys); and other internal data or records.

Richard Light, a nationally recognized proponent of conducting and using research, especially in the area of student assessment, believes that good research is one of the most important bases for sound decision making. Light and colleagues (1990) have argued that, “If used wisely, it [survey research] can lead to improvements throughout the entire fabric of an institution” (p. 234).

At the same time, they caution would-be data gatherers to think carefully before adding to the mountain of unused and often unusable survey data. Serious, targeted, and well-directed surveys that are sensibly integrated into an overall assessment and improvement approach can be a valuable tool, but individuals who are planning a survey should “weigh the costs against the benefits” (p. 10). Teams should not embark on a survey simply because it seems like the thing to do. A useful question to keep in mind is, “What do you want to come out of this?” (p. 232).

If no other data source is available, then a survey may be necessary.

Initial Decisions

To be useful in continuous quality improvement, survey results must be valid, that is, the survey must measure what it

intends to measure. Validity depends on how much error enters into a survey process. Some

An integral part of a well-designed survey is to “plan in” quality all along the way. One must devise ways to keep respondent mistakes and biases to a minimum.

Scheuren, 2004 (p. 18)

possible sources of error relate to:

- 1) whether the persons surveyed represent the true population under study;
- 2) how well the respondents understood the questions asked;
- 3) how willing people are to participate in the survey process; and,
- 4) how well the results are analyzed.

To lessen errors associated with these areas, a team undertaking a survey process needs to consider several basic questions at the start.

How will the survey results be used?

The purpose of the survey drives the collection method, the persons to be included in the survey process, the types of questions asked, and many other factors. The goal of the survey should be to collect objective, unbiased information from a representative group of stakeholders. One helpful step to take up front is to outline what you want to learn from the survey and how the results will help the team improve their processes. The survey should focus on the improvement of processes rather than individuals.

Who should be surveyed?

One way to increase validity of survey results is to ensure that participants in the survey process are the stakeholders who are affected by or involved in the processes under review in the survey. These persons will be the ones most knowledgeable about the outcomes and impacts of the process and have the most relevant input for improvement.

How many should be surveyed?

Some surveys include all persons within a population, while others sample just a subset of these individuals. For instance, the 2004 Student Satisfaction Survey at Penn State was based on a sample of only 6,877 students who represented the 70,000+ student population. At the other extreme, the 2004 Faculty/Staff Survey was distributed to all 15,000+ full-time employees.

Sampling, if done correctly, lowers the costs of surveys and decreases the logistics involved. For paper questionnaires, it decreases the costs of printing and mailing questionnaires and saves on the time necessary for entering data. Sampling also burdens fewer persons with completing a questionnaire or participating in an interview.

Who will design and administer the questionnaire and analyze the results?

Part of the answer to this question depends on how much time, money, expertise, and personnel resources are available. Survey administration requires a commitment of personnel time in developing the questionnaire, sending out questionnaires, following up with non-respondents, processing and entering data, and analyzing results. Preparation, administration, and analysis of a questionnaire takes time and may be better handled by a unit out-

side of the organization. In addition, questionnaires which include sensitive topics might be better administered by an outside organization. Recipients may be more willing to respond to what they perceive as a more neutral or objective party.

Several resources exist at Penn State for teams who want assistance in developing and administering questionnaires.

- The Office of Planning and Institutional Assessment is available to help teams determine if the survey process is the best method to capture the data in which they are interested. The Office can also assist in the development of questionnaires and depending on the scope, complexity, and regularity of surveys, may be able to administer surveys for teams.
- Management Engineering is an internal consultant division within the Office of Human Resources. Management Engineering has experience in survey design and survey data analysis techniques, and, for a nominal fee, can assist teams by working with them to develop questionnaires and administer them.
- The Social Science Research Institute's Survey Research Center provides survey research services to Penn State faculty and administrative units. For a fee, the Center provides services in the areas of survey design, sampling, and data collection (including telephone, mail, face-to-face, internet surveys and focus groups); survey data management; and data analysis.
- The Statistical Consulting Center (SCC) is an educational and service unit in Penn State's Department of Statistics. The Center provides consulting at no charge for graduate students working on projects directly related to their theses. For others, the SCC works on a fee basis and is able to consult with teams on questionnaire design and data analysis.

It is crucial that the persons most involved with the processes under study, and the persons using the results of the survey be included in the team that develops the questions. These persons have the most knowledge of the process and how the results will be used.

Do we need to submit an application to the Institutional Review Board?

University faculty, staff, administrators, or students who conduct research with human participants must submit an application to the Office for Research Protections (ORP) for review and approval by the Institutional Review Board

(IRB). The IRB defines research as a "systematic investigation designed to develop or contribute to generalizable knowledge." (See the ORP's website for more information <http://www.research.psu.edu/orp/about/whosubmits.asp>.) Typically, many surveys conducted by quality improvement teams are not considered research since the results are kept internal to the University and do not add to "generalizable knowledge." Teams may want to complete the *Non-Human/Non-Research Determination Form* found at <http://www.research.psu.edu/orp/areas/humans/applications/index.asp>.

Survey Format

Surveys can be conducted through:

- face-to-face interviews
- telephone interviews
- paper questionnaires
- online questionnaires
- a combination of these methods

Some factors to consider in deciding on a format are:

Cost - The cost of face-to-face and telephone interviews are generally higher than other formats and stem from interviewer expenses. Paper survey costs depend on the material, printing and postage costs. The cost of an electronic format survey depends on whether software is purchased or a free online survey service is used. Several companies offer limited web survey services at no cost. (These types of services can be found by using a Web search engine to search for "free web surveys".) Generally, web surveys tend to be the least expensive format.

Project Length - Across all survey formats, development of the questionnaire generally will take about the same amount of time. It is administration of the questionnaire and the data entry requirements that vary by format. Typically, online questionnaires require the least amount of time because the delivery is almost instantaneous and the time required for data entry is short.

Sampling Bias - To provide valid and useful results, the population being surveyed should fairly represent stakeholders. Some survey formats, such as online surveys, may not reach a fair representation of the audience since some members may not have access to the Web or e-mail. For example, since older persons are less likely to use the Internet and e-mail, an online survey of retired persons may not be the best choice. A mailed or phone survey may be a better choice for these types of audiences.

Developing a Questionnaire

Basic questionnaires consist of a cover letter, the questions, and a concluding page. The cover letter should be brief and appear at the front of the questionnaire. This cover letter should address:

- the purpose of the survey
- the persons conducting the survey
- the date by which the form should be returned
- an estimate of the time required for completion
- the confidentiality policy
- a person to contact

After completing the questions, respondents should be thanked for their participation and should again be given the name of a person to contact in case they have questions about the survey.

Types of Survey Questions

Questions fall into two categories: open-ended and closed. In **open-ended questions**, participants answer the questions in their own words. These types of questions are useful to elicit respondent feelings, and to provide depth to an issue. They should be written in a manner which precludes one-word responses, such as "Yes" or "No".

| |
|--|
| Instead of asking: "Are you satisfied with this service?" Ask: "With what aspects of this service are you satisfied?" |
|--|

Open-ended questions provide much information about the selected topics, but they are more difficult to analyze since they may cover a wide range of topics and need to be coded or grouped to provide some level of summary.

Closed questions provide the respondent with a defined set of answers. The response set can include categorical or scaled responses. **Categorical** question response sets have no inherent ordering within them. A question about gender is categorical (Male, Female), as is a question about type of transportation to work (Car, Bus, Taxi, Bicycle, Walk).

Scaled responses, on the other hand, have some type of progressive order. A question about age is one example of a scaled question while another is a question which asks respondents to rate their agreement or satisfaction. Although responses to the latter may not have a numerical value (i.e., Very Satisfied, Somewhat Satisfied, Neutral, etc.), they are scaled because they have a progressive order.

Writing Questions

Generally, even if an outside resource administers the survey process, individuals on the teams who are working on the process improvement are the persons to identify the topics and types of questions to be included. The following suggestions can help lessen survey error because they focus on increasing question clarity and participant motivation to complete questions:

The goal of writing a survey question for self-administration is to develop a query that every potential respondent will interpret in the same way, be able to respond to accurately, and be willing to answer.
Dillman, 2000 (p.32)

Begin the questionnaire with simple questions - Ease respondents into completing the form by asking simple questions that are not open-ended. This may make respondents more comfortable and they may be more likely to complete the questionnaire. However, the initial questions should also be sufficiently interesting to the respondent to attract their attention and make them want to complete the survey. Thus, questions about demographic information, such as gender or age, should not be the first questions on the page, even though they are straightforward.

Use concise sentences - Keep the questions as short as possible, but provide enough context that respondents can understand the question.

Use words and language that respondents will understand - Many experts recommend writing questions at the eighth-grade level. Dillman (2000, p.52) suggests "When a word exceeds six or seven letters, chances are that a shorter and more easily understood word can be substituted" and gives the example of using the word "tired" rather than "exhausted". Also avoid words, expressions, or acronyms that are unique to a specific unit, discipline or area.

Ask only one item per question - Focus the question on one action or attribute only. For example, "How satisfied are you with the knowledge and courtesy of our staff?" asks the respondent to rate staff on both their knowledge and manners, two different traits. Staff may be knowledgeable, but not courteous and responses would not reflect a discrepancy between the two.

Provide balance in responses - To prevent bias, an equal number of positive and negative responses should be offered in a question and the responses at each ends of the scale should be the same in weight.

For instance, if "Very Satisfied" appears at one end of a scaled range, the choice on the other end should be "Very Dissatisfied", not "Terribly Dissatisfied".

Offer neutral response as a choice - Questions which ask respondents to rate an area should include a "Neutral" response category.

Offer "not applicable" or "does not apply" as a choice - Including one of these two choices as an option is important for two reasons. First, respondents who have no knowledge of or experience with the question topic should not assess or comment on it. Second, the respondents who choose "not applicable" reflect any subgroups that do not use a service. This can provide a measure of the true usage of a service or process.

Use a natural flow for questions - The flow of the questions should be logical: Question 3 should follow Question 2, while Question 4 should follow Question 3. If possible, avoid asking the respondent to skip questions, especially on paper questionnaires, since it can confuse respondents. Online forms, however, can unobtrusively target questions to specific groups by automatic flowing. Many web survey packages allow questions to flow based on the categories respondents select in previous questions.

Create categories that cover all possible responses and are mutually exclusive - The responses for each question should cover all possible alternatives and should not overlap. For example, Columns A and B below show response categories to a question which asks students about the amount of time they study per day.

| Column A | Column B |
|--------------------|-------------------------------------|
| a) 0 hours | a) 0 hours |
| b) 1-2 hours | b) Less than 1 hour |
| c) 2-3 hours | c) At least 1 but less than 3 hours |
| d) 3-4 hours | d) At least 3 but less than 5 hours |
| e) 4 or more hours | e) 5 or more hours |

Under Column A, a student who studied for 45 minutes a day would not be able to select a correct response, while a student studying 3 hours a day could select c) or d). Column B displays preferable categories which cover all possible hours yet do not overlap.

Include plenty of space for written comments - Encourage respondents to write more by providing an adequate number of lines for responses. More lines may prompt respondents to write longer answers.

Prompt respondents with specific time frames - Providing a common reference point increases the likelihood that respondents will understand questions in a similar manner. Mentions of time may also help participants to better remember their actual activities.

Instead of asking:

“How often do you travel outside Pennsylvania?”

Ask:

“Since January 2006, how often have you traveled outside of Pennsylvania?”

Keep the questionnaire as short as possible – Response rate has a great impact on survey validity and minimizing questionnaire length increases response rates. Respondents to long questionnaires more frequently skip over questions, adding to non-response errors for individual questions. In creating questionnaires, developers should include those questions which are “must-haves” and reconsider those that are “nice-to-haves”.

Minimize identifying questions – Respondents feel more comfortable if they believe they can not be identified from their responses, and are more likely to complete questionnaires. For example, if the population under study includes few females over the age of fifty, older women may be less likely to complete the questionnaire because their individual responses could be identified. It may be important for the team to know the composition of the respondents and whether certain groups have different experiences, but to increase response rates, keep demographic questions to a minimum.

Be consistent in question wording when looking at trends – To accurately assess changes over time, the same questions that appeared on past questionnaires should be used. If the questions vary, apparent differences could be due to measurement issues and not reflect an actual change.

Pre-testing the Questionnaire

Prior to distributing the questionnaire, many survey researchers recommend recruiting a small group to complete the questionnaire and provide input on their understanding of the questions and ability to complete it. The group can include as few as 5-10 people. Items to cover with the persons pre-testing the questionnaire include: 1) any terms or words that were unfamiliar; 2) the clarity of the questions; 3) the flow of the questionnaire; 4) ability to access the form if online; and, 5) the actual time required to complete the questionnaire. Team members can sit with the participants as they complete the pre-test, can personally interview or survey each participant after she or he completes the pre-test, or can hold focus groups with the participants after they complete the pre-test. This type of feedback can improve the quality of the questionnaire.

Increasing Response Rates

Dillman (2000) recommends making five contacts with participants in order to increase response rates for mailed and online surveys. His research, along with that of others, has found that increasing the number of contacts can have a significant effect on return rates, thus increasing survey validity.

- The first contact occurs a few days prior to distributing the questionnaire and informs the participant that they will receive a questionnaire shortly.
- The participant receives the questionnaire two to three days later. Dillman recommends placing an identification code on each questionnaire so that responses can be tracked.
- Three days to a week later, the participant receives a thank you postcard. The note thanks participants for completing and returning the survey, and asks that persons who have not already completed the survey to do so.
- The fourth contact occurs two to four weeks after the initial questionnaire. It consists of a replacement questionnaire sent to those persons who have not yet responded.
- About a week after the fourth contact, Dillman recommends that a fifth contact be made through a delivery method that differs from the initial contacts, i.e., a phone call if the others were by e-mail, or a priority postal mailing if the others were sent by general delivery.

Using Dillman’s guidelines, from the time the pre-notice is sent to the time all completed questionnaires are received requires about two months. Other strategies which have been shown to increase response rates include personalizing all correspondence or including token financial or other incentives.

Data Tabulation

After all questionnaires are received, the results are tabulated. For very simple surveys, the results may be hand-tallied and the responses to each category for every question are counted. This approach is manageable only if the number of respondents is very small and the questions are limited.

Most surveys are more complex and require construction of a database. For paper formats and interviews, this means that responses to each question are entered into some type of data file for analysis. (Web surveys have the advantage here of having each respondent’s answers automatically

entered into a data file, thus saving on data entry time.) Spreadsheet, database and statistical analysis software is available for this step. Whatever software is used, the data entered should be double-checked for data entry errors.

Once the results are tabulated by hand or in a data base, the composition of the respondents should be compared to the composition of the entire population which received the survey. Characteristics for this type of comparison include gender, staff/faculty status, years employed at Penn State, unit, etc. This step indicates whether the respondents to the questionnaire are representative of the entire group. Discrepancies between the two groups need to be noted in the results; substantial differences limit the team’s ability to generalize the results.

Calculating Descriptive Statistics

Simple descriptive statistics such as frequency counts and percentage distributions provide the basic information teams need to answer the questions for which they initially started the survey process.

Frequency counts provide the number of respondents who selected each category in a question. For example as shown in the table below, 100 persons were included in each of two separate surveys. In Survey A, 40 persons responded “Yes” and 30 persons did so in Survey B. The counts for all other response categories also appear in the tables.

| EXAMPLE: Were you satisfied with the quality of the service you received during your last appointment? | | | |
|--|----------------|----------------------|-----------------------------------|
| SURVEY A | | | |
| Response | # of Responses | % of Total Responses | % of Responses (Not Including NA) |
| Yes | 40 | 40% | 50% |
| No | 40 | 40% | 50% |
| Not Applicable | 20 | 20% | ----- |
| TOTAL | 100 | 100% | 100% |
| SURVEY B | | | |
| Response | # of Responses | % of Total Responses | % of Responses (Not Including NA) |
| Yes | 30 | 30% | 50% |
| No | 30 | 30% | 50% |
| Not Applicable | 40 | 40% | ----- |
| TOTAL | 100 | 100% | 100% |

In addition to the frequencies, **percentage distributions** are useful, and should be calculated in a manner that reflects the data appropriately. The percentage distribution is the number of responses in each category divided by the total number of responses. For instance, 40 percent (40/100) of the persons in Survey A and 30 (30/100) percent in Survey B responded “Yes”. This is the straight percentage response.

The team should decide whether the percentage distributions based on the total number of responses is sufficient or whether the distribution should be based on the number of **valid responses**. Typically, valid responses exclude the categories of missing and “Not Applicable” responses. The calculation is shown in the last column in the table above.

It is important to know how many persons selected each response, but “Not Applicable”, along with missing responses, may need to be excluded from calculations of percentage distributions which are used for comparisons. For example, 20 persons in Survey A felt the question was not applicable as did 40 persons in Survey B. Excluding the “Not Applicable” responses shows that half of the respondents in both Survey A (40/80) and Survey B (30/60) responded “Yes”. Analyzing responses only for those persons who use a service better reflects true satisfaction levels.

Analysis of Results

Once the frequency tables and percentage distributions are available, team members can begin to analyze the data. One way to do this is to look for patterns within the data. Some specific areas include:

Comparison within a survey – Response patterns for certain questions may stand out from the others and may indicate an area for improvement. For example, 20 percent of respondents may be satisfied with the waiting time for a process while 60 percent may have reported they were satisfied with the quality of the process. Because of the relatively low satisfaction level with wait time, the team may identify this as an area for improvement.

Comparison across subgroups - Breaking out questionnaire responses by specific characteristics helps teams determine whether certain groups have different experiences. For example, employees who work on a part-time basis may rate availability of parking lower than employees who work full-time. Analysis of group differences depends on the types of attributes captured in the survey, for example, gender, employment status, department, etc.

Comparison across time – Comparisons across time are typically made in order to determine whether implemented process improvements are having the desired effect, and also to identify areas with declining performance or satisfaction for process improvement. The questions and sampling population need to be the same across both surveys to ensure that the apparent change is an actual change and not the result of differences in measurement.

Comparison with specific goals – Some teams may be working with processes that have established goals. In addition, teams may want to benchmark with other areas within the University or other institutions. Similar to comparisons across time, the same question format and sampling methodology should be used.

Summary

Valid, meaningful surveys provide an effective means for teams to obtain stakeholder input, but require much time and effort from team members. Before undertaking a survey, teams should consider whether the data can be collected in other ways. If surveys are necessary, it is important that potential sources of error be minimized. To achieve this, teams must ensure that: persons surveyed represent the true population under study; respondents accurately understand the questions asked; people are willing to participate in the survey process; and, the results are analyzed objectively.

Selected Resources:

University Areas:

Management Engineering

<http://www.ohr.psu.edu/MgtEng/home.cfm>

Office of Planning and Institutional Assessment

<http://www.psu.edu/president/pia/>

Social Science Research Institute Survey Research Center

<http://www.ssri.psu.edu/survey/index.html>

Statistical Consulting Center

<http://www.stat.psu.edu/~scc/index.html>

Additional Resources:

American Association for Public Opinion Research.

<http://www.aapor.org>

Dillman, Don. 2000. *Mail and Internet Surveys: The Tailored Design Method*. New York: John Wiley & Sons, Inc.

Light, Richard J., Judith D. Singer, and John B. Willett. 1990. *By Design: Planning Research on Higher Education*. Cambridge, MA: Harvard University Press.

Porter, Stephen R. 2004. "Pros and Cons of Paper and Electronic Surveys." *New Directions in Institutional Research*, no. 121:91-97.

Scheuren, Fritz. 2004. *What is a Survey?* American Statistical Association.

Umbach, Paul D. (2005.) "Getting back to the basics of survey research." *New Directions in Institutional Research*, no. 127: 91-100.

For more information, contact the Office of Planning and Institutional Assessment at 814-863-8721 or psupia@psu.edu, or visit our website: <http://www.psu.edu/president/pia>.

The services of the Office of Planning and Institutional Assessment are provided free of charge to The Pennsylvania State University and are adapted to the specific needs of departments or individuals.