

MP 05W0000077

MITRE PRODUCT

Fundamentals of Survey Research Methodology

April 2005

Priscilla A. Glasow (25988)

Division: W800
Department: W804

MITRE
Washington C3 Center
McLean, Virginia

MITRE Department

Approval:

Edward F. Gonzalez

Table of Contents

<u>Section</u>	<u>Page</u>
1. What is Survey Research?	1-1
1.1 Survey Strengths.....	1-1
1.2 Survey Weaknesses.....	1-2
1.3 Definition of Terms	1-2
2. The Survey Process	2-1
2.1 Survey Design.....	2-1
2.1.1 Sample Selection.....	2-1
2.1.2 Sample Size.....	2-2
2.1.2.1 Degree of Precision	2-2
2.1.2.2 Statistical Power	2-2
2.1.2.2.1 Effect Size as a Determinant of Power	2-3
2.1.2.2.2 Sample Size as a Determinant of Power.....	2-3
2.1.2.3 Access to Subjects	2-3
2.1.2.4 Stratification and Clustering.....	2-3
2.1.2.5 Units of Analysis	2-3
2.1.3 Choice of Survey Media	2-4
2.1.3.1 Written Surveys.....	2-4
2.1.3.2 Verbal Surveys	2-4
2.1.3.3 Mixed Mode Surveys	2-5
2.2 Survey Instrument Development	2-5
2.2.1 Standards for Good Survey Questions.....	2-5
2.2.1.1 Question Wording	2-5
2.2.1.2 Feasible and Ethical.....	2-6
2.2.1.3 Additional Considerations	2-6
2.2.1.4 Biased Wording.....	2-6
2.2.1.5 Biased Context.....	2-6
2.2.1.5.1 General Characteristics.....	2-6
2.2.1.5.2 Cognitive Implications of Biased Context.....	2-7
2.2.1.5.3 Facets of Biased Context	2-7
2.2.2 Types of Survey Questions.....	2-7
2.2.2.1 Open-ended Questions.....	2-7
2.2.2.2 Closed-ended Questions	2-7
2.2.2.2.1 Questions that Describe and Evaluate People, Places and Events.....	2-8
2.2.2.2.2 Questions that Measure Responses to Ideas, Analyses or Proposals.....	2-9
2.2.2.2.3 Questions that Measure Knowledge	2-9

<u>Section</u>	<u>Page</u>
2.2.3 Subjective Responses to Survey Questions	2-9
2.2.4 Cognitive Tasks Required for Survey Response	2-9
2.2.5 Sources of Measurement Error	2-10
2.3 Survey Execution.....	2-11
2.4 Data Analysis and Reporting Survey Results	2-11
3. Conclusion	3-1
List of References	RE-1
Appendix . A Checklist for Survey Assessment	A-1
A.1 Survey Design	A-1
A.2 Population and Sample.....	A-1
A.3 Survey Instrument	A-1
A.4 Study Variables	A-2
A.5 Data Analysis Plan.....	A-2
Distribution List	DI-1

Section 1

What is Survey Research?

Survey research is used:

“to answer questions that have been raised, to solve problems that have been posed or observed, to assess needs and set goals, to determine whether or not specific objectives have been met, to establish baselines against which future comparisons can be made, to analyze trends across time, and generally, to describe what exists, in what amount, and in what context.” (Isaac & Michael, 1997, p. 136)

Kraemer (1991) identified three distinguishing characteristics of survey research (p. xiii). First, survey research is used to quantitatively describe specific aspects of a given population. These aspects often involve examining the relationships among variables. Second, the data required for survey research are collected from people and are, therefore, subjective. Finally, survey research uses a selected portion of the population from which the findings can later be generalized back to the population.

In survey research, independent and dependent variables are used to define the scope of study, but cannot be explicitly controlled by the researcher. Before conducting the survey, the researcher must predicate a model that identifies the expected relationships among these variables. The survey is then constructed to test this model against observations of the phenomena.

In contrast to survey research, a *survey* is simply a data collection tool for carrying out survey research. Pinsonneault and Kraemer (1993) defined a survey as a “means for gathering information about the characteristics, actions, or opinions of a large group of people” (p. 77). Surveys can also be used to assess needs, evaluate demand, and examine impact (Salant & Dillman, 1994, p. 2). The term *survey instrument* is often used to distinguish the survey tool from the survey research that it is designed to support.

1.1 Survey Strengths

Surveys are capable of obtaining information from large samples of the population. They are also well suited to gathering demographic data that describe the composition of the sample (McIntyre, 1999, p. 74). Surveys are inclusive in the types and number of variables that can be studied, require minimal investment to develop and administer, and are relatively easy for making generalizations (Bell, 1996, p. 68). Surveys can also elicit information about attitudes that are otherwise difficult to measure using observational techniques (McIntyre, 1999, p. 75). It is important to note, however, that surveys only provide estimates for the true population, not exact measurements (Salant & Dillman, 1994, p. 13).

1.2 Survey Weaknesses

Pinsonneault and Kraemer (1993) noted that surveys are generally unsuitable where an understanding of the historical context of phenomena is required. Bell (1996) observed that biases may occur, either in the lack of response from intended participants or in the nature and accuracy of the responses that are received. Other sources of error include intentional misreporting of behaviors by respondents to confound the survey results or to hide inappropriate behavior. Finally, respondents may have difficulty assessing their own behavior or have poor recall of the circumstances surrounding their behavior.

1.3 Definition of Terms

- Verbal surveys are often known as *interviews* and written surveys are *questionnaires*.
- *Reliability* is often used to refer to the consistency of survey responses over time.
- *Item consistency* determines whether the responses for each question are consistent across constructs.
- *Test administration and scoring consistency* examines the possibility of errors caused by carelessness in administration or scoring (Creswell, 1994, p. 121).
- *Validity* is the extent to which the measurements of the survey provide the information needed to meet the study's purpose (Simon & Francis, 1998, p. 70). This definition is limited, however, to the *face validity* of the instrument.
- *Content validity* considers whether the questions measure the content they were intended to measure.
- *Predictive validity* examines whether the responses are able to predict a criterion measure.
- *Concurrent validity* addresses the correlation of survey results with results from other sources.
- *Construct validity* asks whether the survey questions capably measure hypothetical constructs.

Section 2

The Survey Process

2.1 Survey Design

According to Levy and Lemeshow (1999), survey design involves two steps. First, a sampling plan must be developed. The sampling plan is the methodology that will be used to select the sample from the population (p. 6). The sampling plan describes the approach that will be used to select the sample, how an adequate sample size will be determined, and the choice of media through which the survey will be administered. Survey media include telephone and face-to-face interviews, as well as mailed surveys using either postal or electronic mail (Salant & Dillman, 1994, p. 3).

Second, procedures for obtaining population estimates from the sample data and for estimating the reliability of those population estimates must be established. This process includes identification of the desired response rate and the preferred level of accuracy for the survey (Salant & Dillman, 1994, p. 3).

Survey design procedures require inputs from the people who will use the survey data and from those who will conduct the survey. The data users should identify the variables to be measured, the estimates required, the reliability and validity needed to ensure the usefulness of the estimates, and any resource limitations that may exist pertaining to the conduct of the survey (Levy & Lemeshow, 1999, p. 6). The people who conduct the survey should provide additional input regarding resource requirements and offer alternative sampling procedures that they deem feasible and appropriate to the task. Statisticians integrate these inputs to develop a survey design that will meet the data users' requirements within the specified resource constraints.

The following sections address three key elements of survey design: (a) considerations in the selection of the sample, requirements for determining the needed sample size, and considerations for choosing the appropriate survey media.

2.1.1 Sample Selection

Sample selection depends on the population size, its homogeneity, the sample media and its cost of use, and the degree of precision required (Salant & Dillman, 1994, p. 54). The people selected to participate in the sample must be selected at random; they must have an equal (or known) chance of being selected (p. 13).

Salant and Dillman (1994) observed that a prerequisite to sample selection is to define the target population as narrowly as possible (p. 58). It is often not possible, however, to know the true population. In such cases, Attewell and Rule (1991) suggested that a

theoretical sample may be used. Theoretical samples purposively select organizations that exhibit the desired features that are the focus of the researcher's study (p. 300). Although the theoretical sample is not randomly selected, individual respondents from within that sample can be selected at random to achieve an approximate effect.

2.1.2 Sample Size

Determination of sample size depends on five factors:

- Desired degree of precision
- Statistical power required
- Ability of the researcher to gain access to the study subjects
- Degree to which the population can be stratified
- Selection of the relevant units of analysis

2.1.2.1 Degree of Precision

A survey is used to establish that a postulated effect exists in the sample. The contention that this effect exists is the *alternate hypothesis*. The *null hypothesis* claims that this effect does not exist.

The sample must be "large enough to yield the desired level of precision" (Salant & Dillman, 1994, p. 5). Two measures of precision are discussed in the literature. First, the significance level is the amount of Type I error that the researcher will allow in the study. Type I error occurs when the null hypothesis is rejected when it is, in fact, true.

The second measure of precision is the *confidence interval*. A survey sample consists of data for which a mean and variance can be calculated. Confidence intervals can be constructed for each of these statistics, such that the researcher can state that he or she is, for example, "95 percent confident" that the corresponding statistic for the population falls within the specified range of the sample statistic.

Where the purpose of the study is to gain a general sense of a belief or attitude, a lower level of precision may be acceptable. A smaller sample size may then be drawn. Salant and Dillman (1994) noted that the researcher must ensure that the number of surveys distributed is sufficient to allow for no response and for unusable, illegible, and incomplete responses (p. 57).

2.1.2.2 Statistical Power

Statistical power is the probability that the researcher rejects the null hypothesis given that the alternate hypothesis is true (Attewell & Rule, 1991, p. 302). Where the null

hypothesis is erroneously accepted and the alternate hypothesis is true, a Type II error is said to exist. Statistical power is determined, in part, by effect size and sample size.

2.1.2.2.1 Effect Size as a Determinant of Power

As noted earlier, each random sample taken from a population has its own mean and variance. The *distribution of means* is the distribution of all possible means of samples that could be taken from the population. The distribution of means is the distribution given the null hypothesis.

Effect size is a determinant of statistical power that helps us “evaluate the overall importance of a result” (Aron & Aron, 1997, p. 137). In conducting a survey, a sample is taken from the population. The sample distribution, representing the alternate hypothesis, is compared to the distribution of means to determine whether the sample distribution differs significantly from that of the null hypothesis. The *effect size* is the extent to which the distributions of means for the null and alternate hypotheses do not overlap. The greater the difference in the distributions of means, the higher the power (p. 133).

2.1.2.2.2 Sample Size as a Determinant of Power

A second determinant of power is sample size. As the sample size increases, the distribution of means becomes narrower and the variance decreases, thereby reducing the overlap between the two distributions and increasing power (Aron & Aron, 1997, p. 140).

2.1.2.3 Access to Subjects

A third factor in determining sample size is the ability of the researcher to gain access to the desired subjects. One method for mitigating this problem is to use a theoretical sample, as previously described (Attewell & Rule, 1991, p. 301).

2.1.2.4 Stratification and Clustering

The fourth factor in determining sample size is whether the random sample can be stratified, for example, according to sector, size, or technology level (Attewell & Rule, 1991, p. 301). By using multiple levels to describe the technologies used in different organizations, for example, the researcher is able to statistically relate those differences in technology strata to differences (variances) in the impact of those technologies. Salant and Dillman (1994) noted, however, that the sample size must be increased when using stratified samples to maintain necessary precision (p. 57). Control and test groups are one example of stratified sampling.

2.1.2.5 Units of Analysis

The last factor in sample size determination is the selection of appropriate units of analysis. The researcher must decide whether the respondents to a survey will be

individuals, offices, or entire firms (Pinsonneault & Kraemer, 1993, p. 90). Attewell and Rule (1991) noted that workgroups may also be a useful unit of analysis. Aggregating individual questionnaire responses across a team helps to lessen the effects of idiosyncratic or individual attitudes. Such research must then try to explain the differences found across workgroups (p. 304).

2.1.3 Choice of Survey Media

Salant and Dillman (1994) noted that the choice of survey medium is determined by the resources that are available.

2.1.3.1 Written Surveys

Written surveys require minimum resources (staff, time, and cost) and are best suited to eliciting confidential information. Minimal sampling error occurs due to the relatively low cost per survey. There are also minimal interviewer and respondent measurement errors due to the absence of direct contact (Salant & Dillman, 1994, p. 35). Written surveys allow the respondent the greatest latitude in pace and sequence of response (p. 18).

Written surveys may be distributed using either postal or electronic mail. In some cases, written surveys are distributed in person to a group of respondents to evaluate a recent event. This approach is frequently used in military survey research where after action reports are used to evaluate an exercise. Although this method provides immediate results, the involuntary nature of an in-person written survey makes this medium prone to response biases.

Among the disadvantages of written surveys are their subjectivity to certain types of error. For example, written surveys are subject to coverage error where population lists are incomplete or out of date. They are also typically subject to nonresponse error. Less educated, illiterate, and disabled people are particularly less likely to respond to written surveys (Isaac & Michael, 1997, p. 138). Written surveys are also subject to bias where the intended respondent refers to others in completing the survey. Finally, written surveys are subject to item nonresponse where some questions may be inadvertently or intentionally skipped (Salant & Dillman, 1994, p. 35).

2.1.3.2 Verbal Surveys

Verbal surveys include telephone and face-to-face interviews. The face-to-face interview is a particularly flexible tool that can capture verbal inflexion, gestures, and other body language. A skilled interviewer can obtain additional insights into the answers provided by observing the respondent's body language (Isaac & Michael, 1997, p. 140).

Face-to-face interviews are useful where the true population is not known or when respondents are unable or unlikely to respond to written surveys (Salant & Dillman, 1994,

p. 40). They are also well suited to long or complex questionnaires and for reaching the correct respondents.

Verbal surveys are, however, subject to measurement error when untrained interviewers are used (Salant & Dillman, 1994, p. 42). They are also resource intensive in terms of staff, facilities, and time. Findings from face-to-face interviews, in particular, are difficult to summarize and incorporate in data analyses (Isaac & Michael, 1997, p. 140).

2.1.3.3 Mixed Mode Surveys

Salant and Dillman (1994) espoused the use of mixed mode surveys that combine survey media. This approach first uses the best method for achieving a high response rate at the lowest possible cost. Other media are then used to improve the response rate at increasing costs per survey. Written surveys are usually the first method used in mixed mode surveys, followed by verbal survey methods. The authors noted that mixed mode surveys generally reflect higher composite response rates than single medium surveys (p. 50).

2.2 Survey Instrument Development

Survey instrument development must be preceded by certain prerequisites. First, the focus of the study must be carefully defined. Second, the study objectives must be translated into measurable factors that contribute to that focus (Salant & Dillman, 1994, pp. 77-78). Third, the researcher must ensure that he or she is well versed in the topic (p. 99). Finally, the survey must be consistently administered (Fowler, 1995, p. 3).

Survey instruments should ideally be developed by experts in the measurement sciences. Levy and Lemeshow (1999) opined that a statistician should be called upon to provide input on the procedures that will be used to ascertain the quality of the data collected by the instrument, and to ensure that the instrument is conducive to easy data processing and manipulation for analysis.

2.2.1 Standards for Good Survey Questions

At a fundamental level, “a good question is one that produces answers that are reliable and valid measures of something we want to describe” (Fowler, 1995, p. 2).

2.2.1.1 Question Wording

Survey questions should use words that are consistent with the educational level of the intended respondents (McIntyre, 1999, p. 78). Both the question and any response options must be clear to both the respondent and the researcher (Fowler, 1995, p. 2; Salant & Dillman, 1994, p. 92). The wording should preclude alternative interpretations or incomplete sentences that would allow misinterpretation (Browne & Keeley, 1998, p. 115; Fowler, 1995,

p. 3; Salant & Dillman, 1994, p. 100). Survey questions should not be combined where the respondent may wish to answer affirmatively for one part, but negatively for another.

2.2.1.2 Feasible and Ethical

Good survey questions must be feasible to answer and respondents must be willing to answer them (Fowler, 1995, p. 3). Questions must be civil and ethical (McIntyre, 1999, p. 77). The researcher must avoid questions that ask the respondent for data they could not or do not have, including questions that assume the respondent knows something about the subject (Salant & Dillman, 1994, p. 98). Personal questions, objectionable statements that reflect the researcher's bias and questions that require difficult calculations should similarly be avoided (pp. 96-97).

2.2.1.3 Additional Considerations

McIntyre (1999) emphasized that the length of the survey should not be onerous. The researcher should avoid questions that involve double negatives and long questions that lose the respondent in the reading (p. 78). Undefined abbreviations, acronyms, and jargon should not be used (Salant & Dillman, 1994, p. 93). Similarly, the tone of survey questions should avoid biased wording that evokes an emotional response. Rating scales should be balanced to provide an equal number of positive and negative response options (p. 95).

Salant and Dillman (1994) also noted that open-ended questions that require precise answers are difficult for respondents to quickly answer (p. 93). They further cautioned against changing time references where the survey may be given at different times such that responses might reflect seasonal or temporal differences (p. 99).

2.2.1.4 Biased Wording

Biased wording is often observed where the question includes a predisposition either for or against a particular perspective (Salant & Dillman, 1994, p. 94). Such questions may be leading or include assumptions that may not be true (McIntyre, 1999, p. 78).

2.2.1.5 Biased Context

2.2.1.5.1 General Characteristics

Biased context results from the placement of questions in a particular order so that the respondent is already thinking along certain lines on the basis of previous questions (Browne & Keeley, 1998, p. 116). Biased context also occurs when the survey is long. For example, a given question might be used at both the beginning and end of a long survey to which respondents provide different responses. Fowler (1995) also noted that respondents are also more likely to use rankings on the left side of a continuum, regardless whether the continuum is decreasing or increasing from left to right (p. 75).

Salant and Dillman (1994) stated that response options should be limited to less than five choices and series of questions should be used to address complex and abstract issues (p. 87).

2.2.1.5.2 Cognitive Implications of Biased Context

Tourangeau (1999) identified two facets of biased context. First, the level of generality is the degree to which the survey question addresses an issue that is encompassed by a broader issue or attempts to generalize inappropriately about a narrow issue (p. 112). The second facet is the influence of consistency and contrast. Contrast is the degree to which a survey question is consistent with, or contrasts with, other survey questions that immediately precede or follow the question of interest.

2.2.1.5.3 Facets of Biased Context

Tourangeau (1999) identified two facets of biased context. First, the level of generality is the degree to which the survey question addresses an issue that is encompassed by a broader issue or attempts to generalize inappropriately about a narrow issue (p. 112). The second facet is the influence of consistency and contrast. Contrast is the degree to which a survey question is consistent with, or contrasts with, other survey questions that immediately precede or follow the question of interest.

2.2.2 Types of Survey Questions

2.2.2.1 Open-ended Questions

Open-ended survey questions allow respondents to answer in their own words. Open-ended questions also allow the researcher to explore ideas that would not otherwise be aired and are useful where additional insights are sought (Salant & Dillman, 1994, p. 81). They are also useful where the researcher is less familiar with the subject area and cannot offer specific response options.

Open-ended questions require greater thought and contemplation on the part of the respondent, and are, therefore, more time intensive to answer (Salant & Dillman, 1994, p. 79). The results obtained from open-ended questions are also more difficult to analyze. Finally, it is more difficult to identify a single course of action from the broad range of responses that are received to open-ended questions (p. 80).

2.2.2.2 Closed-ended Questions

In contrast, *closed-ended* questions require the respondent to choose from among a given set of responses (McIntyre, 1999, p. 75).

Closed-ended questions with ordered choices require the respondent to examine each possible response independent of the other choices. The choices form a continuum of

responses, such as those provided by Likert scales and numerical ranges. These types of questions are easiest for respondents to answer and for researchers to analyze the data.

The second type of closed-ended question is the *closed-ended question with unordered choices* (Salant & Dillman, 1994, p. 83). These questions ask the respondent to compare possible responses and select one. Multiple choice questions are an example of this type. The researcher must ensure that the respondent is given a comprehensive selection of responses. Closed-ended questions with unordered choices are useful for ranking items in order of preference.

The third type of closed-ended question is the *partial closed-ended question* in which the respondent is asked to compare possible responses and select one, or write in "other". Salant and Dillman (1994) observed that most respondents choose one of the given responses when this type of question is presented (p. 84).

Closed-ended questions may also be categorized as: (a) questions that describe and evaluate people, places, and events; (b) questions that measure responses to ideas, analyses, and proposals; and (c) questions that measure knowledge.

2.2.2.2.1 Questions that Describe and Evaluate People, Places and Events

This set of questions includes those that ask respondents to make self-assessments. The question should be carefully examined to ensure that it is not open to different interpretations (Fowler, 1995, p. 48).

- *Evaluative continua*

Evaluative continua are numerical or adjectival scales. Multiple choice questions should ideally offer five to seven (5-7) response options. The researcher must be aware, however, that numerical scales can be misinterpreted. For example, a respondent may begin at one end of a numerical scale and work up or down from there to arrive at his or her response. Similarly, the respondent may start at the midpoint of the numerical scale and formulate a response from that point.

- *Agreement continua*

Agreement continua are simply questions that require the respondent to agree or disagree with given statements. Such scales are subject to differences in interpretation. Fowler (1995) noted that less-educated respondents tend to agree to such questions more easily than do respondents having more education (p. 57).

- *Other rating scales*

Rankings are comparisons among given options, within pairs of options, by cardinality of importance (first, second, third), or that score items one at a time using a common scale

(Fowler, 1995, p. 58). Magnitude estimation asks respondents to place values on various subjects relative to some given baseline or point of reference (p. 61).

2.2.2.2.2 Questions that Measure Responses to Ideas, Analyses or Proposals

This type of question asked respondents to compare their own views to the ideas presented in the question statement. For this reason, questions must be clear and unambiguous. As noted earlier, the researcher must be careful to present one idea at a time.

The rating scales used to respond to these questions should avoid response options with emotional content, such as “strongly agree.” Fowler (1995) recommended the use of modifiers such as “completely”, “generally”, or “mostly.” He also suggested that researchers offer a “no response” option for those who have no strong feeling in either direction on the continuum, and a “not enough information” option, which differs in meaning and purpose from “no response” (p. 66).

2.2.2.3 Questions that Measure Knowledge

This type of question is often used to assess respondents’ familiarity with a subject. Such questions are used to gauge respondents’ ability to provide informed responses or to identify those respondents who believe they are informed and compare their responses to those who do not believe they are informed (Fowler, 1995, p. 68). The rating scales used for this type of question are usually of the true-false, yes-no, and multiple-choice variety. Fowler suggested that the researcher intentionally include some plausible, but incorrect answers in the multiple-choice format to distinguish those who know from those who think they know the correct answers.

2.2.3 Subjective Responses to Survey Questions

A respondent’s beliefs, attitudes, and behaviors are imprecise and apt to change over time. *Beliefs* are subjective opinions that indicate what people think. *Attitudes* are subjective opinions that identify what people want. *Behaviors* are objective facts of what people do. *Attributes* are objective facts that describe what people are. These also change, but over longer periods.

Beliefs, attitudes, and behaviors are also often inadequately contemplated. Salant and Dillman (1994) suggested that researchers use a series of related questions to gauge beliefs, attitudes, and behaviors, then examine the responses to identify patterns and consistencies in the answers (p. 88). The survey should ask for specific recommendations to be accepted or rejected, or to rank the relative importance of competing interests (p. 90).

2.2.4 Cognitive Tasks Required for Survey Response

Schwarz (1999) considered the cognitive tasks that respondents perform when asked to answer a survey question.

The first cognitive task is question interpretation. Specifically, the respondent must understand what the researcher is asking and determine what information will best meet that request (Schwarz, 1999, p. 66).

The second cognitive task is response formulation. Schwarz (1999) noted that respondents tend to construct new judgments as that is less cognitively demanding than determining whether previously-held judgments meet the specific constraints of the question (p. 66).

In the third cognitive task, the respondent communicates the response to the researcher. Schwarz (1999) observed that given response options may constrain cognitive activity so that the respondent only generates a response that directly fits the given options (p. 66). Additionally, the respondent may intentionally or unintentionally edit the response to meet unstated expectations of political correctness or social norms.

2.2.5 Sources of Measurement Error

Salant and Dillman (1994) cautioned interviewers to avoid leading respondents to specific answers, interpreting questions for them, or projecting an image that suggests certain answers are desired (p. 19). Each is a source of measurement error.

The respondent is another source of measurement error. Salant and Dillman (1994) observed that respondents may answer as they think the interviewer wants them to answer (p. 20). Additionally, responses to surveys may not reflect the true beliefs, attitudes, or behaviors of the respondents. Respondents may intentionally provide false responses to invalidate the survey's results or choose not to reveal their true insights for a host of personal reasons, reasons that may not be rational or even understood by the respondent (Browne & Keeley, 1998, p. 114).

Isaac and Michael (1997) identified three additional sources of bias associated with the respondent. First, the conduct of a survey is generally outside the daily routine of most respondents and their participation may invoke feelings of being special (p. 137). The Hawthorne effect, named after the Hawthorne Works of the Western Electric Company, is perhaps the most well-known example of this type of bias. The Hawthorne studies of worker performance in 1927 found that worker performance improved simply from the awareness that experimental attempts were being made to bring about improvement.

The second type of respondent bias noted by Isaac and Michael (1997) was the propensity of respondents to agree with bias inherent in the wording of the question, such that respondents more readily agreed with positively-worded questions.

Finally, respondents may give consistently high or low ratings, reflecting a rater bias that detracts from the validity of the results (Isaac & Michael, 1997, p. 137).

2.3 Survey Execution

The third phase of the survey process is the execution, or use, of the survey instrument. Salant and Dillman (1994) emphasized the importance of maintaining the confidentiality of individual responses and reporting survey results only in the aggregate. Another ethical consideration is recognizing that survey participation is a voluntary event that requires the researcher to encourage participation without undue pressure or coercion of the participants (p. 9).

A pilot survey must first be conducted to test both the instrument and the survey procedures before the actual survey is conducted (Levy & Lemeshow, 1999, p. 7). Surveys can be evaluated in two ways. First, survey questions can be evaluated using focus group discussions, cognitive interviews to determine how well respondents understand the questions and how they formulate their responses, and pilot tests of surveys under field conditions (Fowler, 1995, p. 5). Second, responses to surveys can be analyzed to reveal expected relationships among the answers given, and to ensure consistency of respondent characteristics across questions. Responses can be compared to alternatively worded questions and to official records when available. Surveys can also be evaluated by measuring the consistency of responses to given questions over time.

Field testing the survey instrument facilitates later data collection and analysis (Isaac & Michael, 1997, p. 137). Once field testing has been completed, the survey is conducted and the data are collected, coded, and processed.

2.4 Data Analysis and Reporting Survey Results

Finally, it is worthwhile to consider the resource requirements of surveys, data analysis, and effective presentation of results as important elements of a credible and successful survey. Isaac and Michael (1997) espoused the use of automated data collection tools to facilitate data tabulation and manipulation (p. 137). Lucas (1991) urged the use of nonparametric statistics where small sample sizes are involved (p. 278).

Section 3

Conclusion

This paper described the survey process, consisting of three phases: (a) survey design, (b) survey instrument development, and (c) survey execution. Data analysis and reporting of results was identified as an important fourth phase of the survey process. This paper provided specific guidance for the design and implementation of survey research. Additional guidance is provided in the Appendix—A Checklist for Survey Assessment.

List of References

- Aron, A., & Aron, E. N. (1997). Statistics for the behavioral and social sciences: A brief course. Upper Saddle River, NJ: Prentice Hall.
- Attewell, P., & Rule, J. B. (1991). Survey and other methodologies applied to IT impact research: Experiences from a comparative study of business computing. Paper presented at The Information Systems Research Challenge: Survey Research Methods.
- Bell, S. (1996). Learning with information systems: Learning cycles in information systems development. New York: Routledge.
- Browne, M. N., & Keeley, S. M. (1998). Asking the right questions: A guide to critical thinking. (5th Ed.). Upper Saddle River, NJ: Prentice Hall.
- Creswell, J. W. (1994). Research design: Qualitative and quantitative approaches. Thousand Oaks, CA: Sage.
- Fowler, J., Floyd J. (1995). Improving survey questions: Design and evaluation. (Vol. 38). Thousand Oaks, CA: Sage Publications.
- Isaac, S., & Michael, W. B. (1997). Handbook in research and evaluation: A collection of principles, methods, and strategies useful in the planning, design, and evaluation of studies in education and the behavioral sciences. (3rd Ed.). San Diego: Educational and Industrial Testing Services.
- Kraemer, K. L. (1991). Introduction. Paper presented at The Information Systems Research Challenge: Survey Research Methods.
- Levy, P. S., & Lemeshow, S. (1999). Sampling of populations: Methods and applications. (3rd ed.). New York: John Wiley and Sons.
- Lucas, J., Henry C. (1991). Methodological issues in information systems survey research. Paper presented at The Information Systems Research Challenge: Survey Research Methods.
- McIntyre, L. J. (1999). The practical skeptic: Core concepts in sociology. Mountain View, CA: Mayfield Publishing.

Pinsonneault, A., & Kraemer, K. L. (1993). Survey research methodology in management information systems: An assessment. Journal of Management Information Systems, 10, 75-105.

Salant, P., & Dillman, D. A. (1994). How to conduct your own survey. New York: John Wiley and Sons.

Schwarz, N. (1999). Cognitive research into survey measurement: Its influence on survey methodology and cognitive theory. In M. G. Sirken, D. J. Herrmann, S. Schechter, N. Schwarz, J. M. Tanur, & R. Tourangeau (Eds.), Cognition and Survey Research. New York: John Wiley and Sons.

Simon, M. K., & Francis, J. B. (1998). The dissertation cookbook: A practical guide to start and complete your dissertation. (2nd Ed.). Dubuque, IA: Kendall/Hunt.

Tourangeau, R. (1999). Interdisciplinary survey methods research. In M. G. Sirken, D. J. Herrmann, S. Schechter, N. Schwarz, J. M. Tanur, & R. Tourangeau (Eds.), Cognition and Survey Research. New York: John Wiley and Sons.

Appendix

A Checklist for Survey Assessment

A.1 Survey Design

- 1) Indicate why a survey is the preferred data collection method for this study. Consider its propensity to error and bias, resource requirements, and comparison to other data collection methods.
- 2) Indicate whether the survey will be cross-sectional (information collected at one time) or longitudinal (information collected over time).
- 3) Specify the survey media, the rationale for that choice, arguments for cost, availability, convenience, and accuracy of results.

A.2 Population and Sample

- 4) Describe the population, its size if known, and how it will be identified. Address questions of access and use of sampling frames if intended.
- 5) Determine whether the sampling design will be single stage (researcher contacts subjects directly) or multi-stage (researcher must first contact organizations, gather lists within organizational units, then sample within each cluster).
- 6) Identify how the sample will be selected. Aim for random sampling, using theoretical samples only if subject availability is problematic.
- 7) Discuss whether the population will be stratified to ensure that specific characteristics are represented in the sample. The sample must reflect the characteristics of the population in general.
- 8) Identify the strata that will be used (gender, age, education level, rank). For each stratum, state whether the sample will be chosen so that the characteristic of interest is proportionate or disproportionate to the size of the stratum.
- 9) Indicate the procedure for selecting the random sample from the population lists or the sampling frame. Random number selection is most rigorous.
- 10) Describe the sample size required and explain the sample size formula used.

A.3 Survey Instrument

- 11) Identify the survey instrument that will be used, whether it was designed for this study, modified from previous instruments, or developed by another researcher.

For modified and legacy instruments, indicate that permission has been granted to use the instruments (or parts thereof) for this study.

- 12) For legacy instruments, discuss the findings of previous researchers regarding the validity and reliability of the instrument.
- 13) Attach a copy of the survey instrument as an appendix to the test plan.
- 14) Describe the major sections of the instrument, including the cover letter, the question categories (demographics, belief questions, attitude questions, behavior questions, attribute questions), and the closing instructions. Describe the types of scales that are to be used (Likert ratings, binary choice, multiple choice, preference rankings).
- 15) Describe how the instrument will be pilot tested. Provide a rationale for this procedure, including the number of testers and how their comments will be incorporated into the final survey version.
- 16) Establish the follow-up actions for written surveys, such as postcard reminders, follow-up letters with replacement questionnaire forms, and telephone calls.

A.4 Study Variables

- 17) Develop a table that lists each variable (independent, dependent, and mediating) in the study, the research question or hypothesis that it responds to, and the survey questions that will elicit that information.

A.5 Data Analysis Plan

- 18) Discuss the method for determining nonresponse bias. Wave analysis is one method that examines the responses received from those who submitted their surveys late in the response period. This group almost did not respond and, therefore, approximates the nonrespondents. If responses from this late group are not substantially different from responses received earlier in the response period, "a strong case for absence of response bias can be established" (Creswell, 1994, pp. 123-124). Another method for determining nonresponse bias is to telephone a few nonrespondents to determine whether their responses to selected survey questions differ significantly from the responses that were received.
- 19) A descriptive analysis of all independent and dependent variables will be conducted. Summary statistics, such as means, variances, and standard deviations, will be reported for each variable.
- 20) If building one's own scales in an instrument, discuss how survey items will be combined into scales on the independent and dependent dimensions by using factor analysis. Report how the reliability of these scales will be checked statistically for

internal consistency, a measure also demonstrating the construct validity of the scales on the instrument.

- 21) Identify the statistics that will be used to compare groups or relate variables and provide evidence either in support of in refutation of the hypothesis. Provide a rationale for the choice of statistics and base that rationale on the unit of measurement of scales in the study, the intent of the research to either relate variables or compare groups, and whether the data meet the assumptions of the statistic.

Distribution List

Internal

W800

H. Carpenter

B. Moran

W804

P. Glasow

E. Gonzalez

R. Richards

R304

InfoCenter Services (M/S-H115)