SELECTION OF THE SAMPLE AND DATA COLLECTION

Introduction Data collection is one of the most important stage in conducting a research. You can have the best research design in the world but if you cannot collect the required data you will be not be able to complete your project. Data collection is a very demanding job which need thorough planning , hard work, patience, perseverance and more to be able to complete the task successfully. Data collection starts with determining what kind of data required followed

Introduction

Data collection is one of the most important stages in conducting a research. You can have the best research design in the world but if you cannot collect the required data you will be not be able to complete your project. Data collection is a very demanding job which needs thorough planning, hard work, patience, perseverance and more to be able to complete the task successfully. Data collection starts with determining what kind of data required followed by the selection of a sample from a certain population. After that, you need to use a certain instrument to collect the data from the selected sample.

Types of Data

Data can be divided into two types, namely quantitative and qualitative. Quantitative data is numerical in nature and can be mathematically computed . Quantitative data measure uses different scales, which can be classified as nominal scale, ordinal scale, interval scale and ratio scale. Nominal scale is used to categorize items into different group, for example male and female, different examination grades, different races in a country, different types of companies different shoe sizes and so on. The statistics that is used to analyze this type of data is mode. The second measure for quantitative data is the ordinal scale which other than providing information as the nominal scale does also permit ranking of the data. For example, shoe sizes can be ranked from big to small. A typical ordinal in ranking a set of data is poor, satisfactory, good, very god, excellent. On the other hand, an interval scale not only rank order of a set of data, it also measure the order in units of equal intervals. However, the starting point for the measurement is arbitrary rather than absolute. For example, Celsius in temperature measurement is an interval scale because 0o C is an arbitrary value, it is not absolute zero, it does not mean temperature is lacking at that point. For absolute quantities, we have to use the ratio scale. For example, measuring temperature in Kelvin is a ratio scale as it starts the measurement at absolute zero, which is about -273o Celsius.

Qualitative data are mostly non-numerical and usually descriptive or nominal in nature. This means the

Sources of data

Generally we can collect data from two sources, primary sources and secondary sources. Data collected from primary sources are known as primary data and data collected from secondary sources are called secondary data.

Primary data are also known as raw data. Data are collected from the original source in a controlled or an uncontrolled environment. Example of a controlled environment are experimental research where certain variables are being controlled by the researcher. On the other hand, data collected through observation or questionnaire survey in a natural setting are examples data obtained in an uncontrolled environment. Secondary data are data obtained from secondary sources such as reports, books, journals, documents, magazines, the web and more.

Data Collection Methods

There are many methods to collect data, depending on our research design and the methodologies employed. Some of the common methods are questionnaires , interview and observation.

Questionnaires Survey

According to Wikipedia, a questionnaire is a research instrument consisting of a series of questions and other prompts for the purpose of gathering information from respondents. Although they are often designed for statistical analysis of the responses, this is not always the case. The questionnaire was invented by Sir Francis Galton. Questionnaires are widely used for both quantitative and qualitative research

Questionnaires often use various measuring scales to obtain information from the respondents. To obtain the bio data of the respondents, we use nominal scale as it serves as label or identification such as gender and age, which often does not involve calculations. To gather respondents' preferences, we use ordinal scale which is used to arrange objects or alternatives according to their magnitude in an ordered relationship. To obtain information related to attitude, we use rating scales. Rating asks the respondents to estimate the magnitude of a characteristic or quality regarding certain object or thing.

Questionnaire is designed both for descriptive as well as analytical surveys. In a descriptive survey, the questionnaire will normally use nominal and ordinal scales because it concerns primarily with the particular characteristics of a specific population of subjects. It does not required the examination of dependent and independent variables. Examples of questions asked in a descriptive survey are shown below:

• State the location of your company

• How many workers are employed by your company?

• When was the company founded?

On the other hand, rating scale is always use to measure attitude or opinion of the respondents in an analytical survey. In an analytical survey, it normally needs to identify independent, dependent and extraneous variables based on a certain conceptual framework. A researcher needs to conduct a thorough literature review by paying attention to any existing research and theory relevant to the research problem. Once the variables are determined, they are built into a questionnaire using rating scale measures, the most popular one being the Likert scale. Likert scale questions can consist of three

items, four items, five item, six items, seven items and more. The ratings obtained from the respondents are then summarized or averaged up to reflect a certain variable, such as job satisfaction. They can be used to analysed any causal relationship with other variables. Example of Likert-type questions are shown below:

Leisure Research Methods

Once a research question has been determined the next step is to identify which method will be appropriate and effective.

The table below describes the basic characteristics of different methodologies.

Data Collection Methods   Examples

Documents

       Historical

       Literature review

       Meta-analysis

       Diaries

       Content Analysis

       Secondary Data(data mining)

       These methods identify trends in leisure research and practice. Participants keep diaries and journals researcher conducts content analysis of studies, reports and diaries.

Observations

       Interpretive

       Ethnographic

       Participant observer

       Case study

       How people behave and interact in public open spaces. Observe systematically, become a participant observer.

Survey

Questionnaire

Interview

Standardized Scales/Instruments

To learn what people think about leisure motivation. To identify

relationships between motivation and satisfaction. Use interviews, surveys and standardized scales.

Experimental

True designs

Quasi designs

Obtain information under controlled conditions about leisure attitudes and experience with virtual reality. Subjects may be randomly assigned to various tests and experiences then assessed via observation or standardized scales.

Other Field Methods

Nominal Group Technique

Delphi

To identify trends and issues about leisure services, management and delivery systems. Focus Group systems. Various group, question and pencil paper exercises are used by facilitators.

Multimethods Approach

Combination of methods shown

Interviews, journals and quantitative measures are combined to provide a more accurate definition and operationalization of the concept.

Source: Issac & Michael, 1985; Leedy, 1985; Dandekar, 1988; Thomas & Nelson, 1990.

Qualitative and Quantitative Research Methodologies

Quantitative research methods include:

Experiments: random treatment assignments and quasi experiments using nonrandomized treatments.

Surveys: which are cross-sectional or longitudinal

Qualitative research methods include:

ethnographies which are observations of groups

grounded theory which uses multi-staged data collection

phenomenological studies which studying subjects over a period of time through developing relationships with them and reporting findings based on research "experiences."

case studies which use various data to investigate the subject over time and by activity.

Each research method has it's strengths and weaknesses. When designing a research study it is important to decide what the outcome (data) the study will produce then select the best methodology to produce that desired information.

Data Collection Techniques

There are two sources of data. Primary data collection uses surveys, experiments or direct observations. Secondary data collection may be conducted by collecting information from a diverse source of documents or electronically stored information. U.S. census and market studies are examples of a common sources of secondary data. This is also referred to as "data mining."

Key Data Collection Techniques

Surveys

Questionnaires

Panel Questionnaire Designs

Interviews

Experimental Treatments

Writing an Introduction

In any research proposal the researcher should avoid the word "investigation." This word is perceived in a negative sense.

The key components of a good introduction include

1. a description of the purpose of the study,

2. identification of any sponsoring agency,

3. a statement regarding confidentiality,

4. a description of how sample or respondents were selected, and

5. an explanation of the results and their applications.

Experimental Treatments

Experimental designs are the basis of statistical significance. An example of the fundamentals of an experimental design is shown below.

A researcher is interested in the effect of an outdoor recreation program (the independent variable, experimental treatment, or intervention variable) on behaviors (dependent or outcome variables) of youth-at-risk.

In this example, the independent variable (outdoor recreation program) is expected to effect a change in the dependent variable. Even with a well designed study, an question remains, how can the researcher be confident that the changes in behavior, if any, were caused by the outdoor recreation program, and not some other, intervening or extraneous variable ? An experimental design does not eliminate intervening or extraneous variables; but, it attempts to account for their effects.

Experimental Control

Experimental control is associated with four primary factors (Huck, Cormier, & Bounds, 1974).

1.   The random assignment of individual subjects to comparison groups;

2.   The extent to which the independent variable can be manipulated by the researcher;

3.   The time when the observations or measurements of the dependent variable occur; and

4.   Which groups are measured and how.

Treatment Group: The portion of a sample or population that is exposed to a manipulation of the independent variable is known as the treatment group. For example, youth who enroll and participate in recreation programs are the treatment group, and the group to which no recreation services are provided constitutes the control group.

Validity Issues

There are two primary criteria for evaluating the validity of an experimental design.

        Internal validity. Determines whether the independent variable made a difference in the study? Can a cause-and-effect relationship be observed? To achieve internal validity, the researcher must design and conduct the study so that only the independent variable can be the cause of the results (Cozby, 1993).

        External validity, refers to the extent to which findings can be generalized or be considered representative of the population.

Confounding Errors

Errors: are conditions that may confuse the effect of the independent variable with that of some other variable(s).

1.   Premeasurement and interaction errors

2.   Maturation errors

3.   History errors

4.   Instrumentation errors

5.   Selection bias errors

6.   Mortality errors

EXPERIMENTAL DESIGNS

1.   True Designs

2.   Quasi Designs

3.   Ex Post Facto Designs

True Designs - Five Basic Steps to Experimental Research Design

1. Survey the literature for current research related to your study.

2. Define the problem, formulate a hypothesis, define basic terms and variables, and operationalize variables.

3. Develop a research plan:

    a. Identify confounding/mediating variables that may contaminate the experiment, and develop methods to control or minimize them.

    b. Select a research design (see Chapter 3).

    c. Randomly select subjects and randomly assign them to groups.

    d. Validate all instruments used.

    e. Develop data collection procedures, conduct a pilot study, and refine the instrument.

    f. State the null and alternative hypotheses and set the statistical significance level of the study.

4. Conduct the research experiment(s).

5. Analyze all data, conduct appropriate statistical tests and report results.

Quasi Designs

The primary difference between true designs and quasi designs is that quasi designs do not use random assignment into treatment or control groups since this design is used in existing naturally occurring settings.

Groups are given pretests, then one group is given a treatment and then both groups are given a post-test. This creates a continuous question of internal and external validity, since the subjects are self-selected. The steps used in a quasi design are the same as true designs.

Ex Post Facto Designs

An ex post facto design will determine which variables discriminate between subject groups.

Steps in an Ex Post Facto Design

1.   Formulate the research problem including identification of factors that may influence dependent variable(s).

2.   Identify alternate hypotheses that may explain the relationships.

3.   Identify and select subject groups.

4.   Collect and analyze data

Ex post facto studies cannot prove causation, but may provide insight into understanding of phenomenon.

OTHER FIELD METHODS/GROUP TECHNIQUES

Nominal Group Technique (NGT)

The NGT is a group discussion structuring technique. It is useful for providing a focused effort on topics. The NGT provides a method to identify issues of concern to special interest groups or the public at large. Ewert (1990) noted that the NGT is a collective decision-making technique for use in park and recreation planning and management. The NGT is used to obtain insight into group issues, behaviors and future research needs.

Five Steps of the NGT

1.   Members of the group identify their individual ideas in writing, without any group discussion;

2.   Each member lists his/her own ideas and then rank-orders them, again without any group discussion;

3.   A facilitator gives each participant an opportunity to state his/her ideas (one item per person at a time, in round-robin fashion) until all ideas are exhausted;

4.   As a group, participants discuss and consolidate ideas into a list; and

5.   Finally, members vote to select priority ideas. The final list of ideas becomes the focus of further research and discussion. These ideas can also be used to generate a work plan for a formal strategic planning process, a basis for a survey or interview, or the development of a scale.

Source: (Mitra & Lankford, 1999)

Delphi Method

The delphi method was developed to structure discussions and summarize options from a selected group to:

avoid meetings,

collect information/expertise from individuals spread out over a large geographic area, and

save time through the elimination of direct contact.

Although the data may prove to be valuable, the collection process is very time consuming. When time is available and respondents are willing to be queried over a period of time, the technique can be very powerful in identifying trends and predicting future events.

The technique requires a series of questionnaires and feedback reports to a group of individuals. Each series is analyzed and the instrument/statements are revised to reflect the responses of the group. A new questionnaire is prepared that includes the new material, and the process is repeated until a consensus is reached.

The reading below is a research study that used the delphi technique and content analysis to develop a national professional certification program.

 Sample Research Article: Job Competency Analyses of Entry-Level Resort and Commercial Recreation Professionals

Focus Groups

Richard Krueger (1988), describe the focus group as a special type of group in terms of purpose, size, composition, and procedures. A focus group is typically composed of seven to twelve participants who are unfamiliar with each other and conducted by a trained interviewer. These participants are selected because they have certain characteristics in common that relate to the topic of the focus group.

The researcher creates a permissive environment in the focus group that nurtures different perceptions and points of view, without pressuring participants to vote, plan, or reach consensus. The group discussion is conducted several times with similar types of participants to identify trends and patterns in

perceptions. Careful and systematic analysis of the discussions provide clues and insights as to how a product, service, or opportunity is perceived.

A focus group can be defined as a carefully planned discussion designed to obtain perceptions on a defined area of interest in a permissive, nonthreatening environment. It is conducted with approximately seven to twelve people by a skilled interviewer. The discussion is relaxed, comfortable, and often enjoyable for participants as they share their ideas and perceptions. Group members influence each other by responding to ideas and comments in the discussion.

CHARACTERISTICS OF FOCUS GROUPS

Focus group interviews typically have four characteristics:

1.   Identify the target market (people who possess certain characteristics);

2.   Provide a short introduction and background on the issue to be discussed;

3.   Have focus group members write their responses to the issue(s);

4.   Facilitate group discussion;

5.   Provide a summary of the focus group issues at the end of the meeting.

Other types of group processes used in human services (delphic, nominal, planning, therapeutic, sensitivity, or advisory) may have one or more of these features, but not in the same combination as those of focus group interviews.

Behavior/Cognitive Mapping

Cognitive and spatial mapping information provides a spatial map of:

current recreation use,

the most significant recreation resources, and

the approximate number of visitors to the recreation areas.

All types of recreation activities and travel involve some level of environmental cognition because people must identify and locate recreation destinations and attractions.

Cognitive mapping allows recreation resource managers the opportunity to identify where users and visitors perceive the best recreation areas are located. It is important to understand user perceptions in order to manage intensive use areas in terms of maintenance, supervision, budgeting, policy development and planning.

Cognitive maps grid the research site into zones. The zones identify existing geographic, climatic, landscape, marine resources, and recreation sites. The grids allow respondents to indicate primary recreation sites, and then a composite is developed to identify high impact areas. Researchers collect

data at recreation areas (beach, campground, marina, trailhead, etc.) by interviewing visitors and recreationists. During the data collection process, random sites, days, times, and respondents (every nth) should be chosen to increase the reliability and generalizability of the data.

Observations

Observational research is used for studying nonverbal behaviors (gestures, activities, social groupings, etc.).

Sommer & Sommer (1986) developed the list shown below to assist in observation research.

1.   Specify the question(s) of interest (reason for doing the study).

2.   Are the observational categories clearly described? What is being observed and why?

3.   Design the measurement instruments (checklists, categories, coding systems, etc.).

4.   Is the study designed so that it will be 'Valid (i.e., does it measure what it is supposed to measure, and does it have some generalizability)?

5.   Train observers in the use of the instruments and how to conduct observational research.

6.   Do a pilot test to (a) test the actual observation procedure and (b) check the reliability of the categories of observation using at least two independent observers.

7.   Revise the procedure and instruments in light of the pilot test results. If substantial changes are made to the instrument, run another pilot test to make sure changes will work under the field conditions.

8.   Collect, compile, and analyze the data and interpret results.

Casual observation is normally done like unstructured interviews. During the early stages of a research project, casual observation allows the researcher(s) to observe subjects prior to designing questionnaires and/or interview formats.

Types of Observation Studies

Participant observer

Windshield surveys

Case study

Documents (also called Secondary Data or Data Mining)

Data mining is commonly used in both qualitative and quantitative research. Secondary data provides data which provides a framework for the research project, development of research question(s), and validation of study findings.

Frequently used sources of secondary data are:

U.S. Census - Extensive demographic data including age, sex, distribution, education, ethnicity, migration patterns, service industry, etc.

Bureau of Labor Statistics - Extensive information on such things as employment, unemployment, types of employment, income, etc.

National Center for Health & Information Vital rates such as births, State Department of Health deaths, health, etc. marriage and divorce rates, etc.

State Employment Departments - Number employed by industry, projected levels of employment growth, available jobs skills and skill shortages

Federal Land Management - National Parks, historic sites, scenic areas, forests by acres, budget and visitation rates.

State Highway Departments - Miles and condition of highways, bike lanes, and streets, capital and maintenance costs of highways

Law Enforcement Agency - Number and types of motor vehicles, types of crimes and violations, number of police officers by county and city, law enforcement

Outdoor Recreation - Number and type of parks, number and type of Agency/Dept. campgrounds, location, and rates for parks, lakes, rivers, etc.

Welfare/Human Services - Number of families on various types of Department assistance such as Aid to families with Dependent Children, Social Security, and SSI. Number of alcohol and drug abuse counselors, number of family counselors. Number and cases of child abuse, spouse abuse, desertions, child adoptions rate.

Newspapers - Scanning local newspapers is an excellent means to become better acquainted with a community and its principal actors as well as the issues that have been of greatest local concern

Content Analysis

Content analysis systematically describes the form or content of written and/or spoken material. It is used to quantitatively studying mass media. The technique uses secondary data and is considered unobtrusive research.

The first step is to select the media to be studied and the research topic. Then develop a classification system to record the information. The techniques can use trained judges or a computer program can be used to sort the data to increase the reliability of the process.

Content analysis is a tedious process due to the requirement that each data source be analyzed along a number of dimensions. It may also be inductive (identifies themes and patterns) or deductive (quantifies frequencies of data). The results are descriptive, but will also indicate trends or issues of interest.

The reading below is a research study that used the delphi technique and content analysis to develop a national professional certification program.

Sample Research Article: Job Competency Analyses of Entry-Level Resort and Commercial Recreation Professionals

Meta-Analysis

Meta-analysis combines the results of studies being reviewed. It utilizes statistical techniques to estimate the strength of a given set of findings across many different studies. This allows the creation of a context from which future research can emerge and determine the reliability of a finding by examining results from many different studies. Researchers analyze the methods used in previous studies, and collectively quantify the findings of the studies. Meta-analysisfindings form a basis for establishing new theories, models and concepts.

Thomas and Nelson (1990) detail the steps to meta-analysis:

1.   Identification of the research problem.

2.   Conduct of a literature review of identified studies to determine inclusion or exclusion.

3.   A careful reading and evaluation to identify and code important study characteristics.

4.   Calculation of effect size. Effect size is the mean of the experimental group minus the mean of the control group, divided by the standard deviation of the control group. The notion is to calculate the effect size across a number of studies to determine the relevance of the test, treatment, or method.

5.   Reporting of the findings and conclusions.

Historical Research

Historical research in leisure studies may focus on:

        biographies of park and recreation professionals (Joseph Lee, Jane Adams, etc.),

        public, non-profit and private institutions (public parks and recreation, federal land management agencies),

        professional movements (playgrounds, leisure education), and

        related concepts (professionalism, certification and licensure, play).

Historical research is also referred to as analytical research. Common methodological characteristics include a research topic that addresses past events, review of primary and secondary data, techniques of criticism for historical searches and evaluation of the information, and synthesis and explanation of findings. Historical studies attempt to provide information and understanding of past historical, legal, and policy events.

Five basic procedures common to the conduct of historical research were identified by McMillan & Schumacher (1984). They provide a systematic approach to the process of historical research.

Step 1: Define the problem, asking pertinent questions such as: Is the historical method appropriate? Are pertinent data available ? Will the findings be significant in the leisure services field?

Step 2: Develop the research hypothesis (if necessary) and research objectives to provide a framework for the conduct of the research. Research questions focus on events (who, what, when, where), how an event occurred (descriptive), and why the event happened (interpretive). This contrasts with quantitative studies, in which the researcher is testing hypotheses and trying to determine the significance between scores for experimental and control groups or the relationships between variable x and variable y.

Step 3: Collect the data, which consists of taking copious notes and organizing the data. The researcher should code topics and subtopics in order to arrange and file the data. The kinds of data analysis employed in historical research include (based on McMillan & Schumacher, 1984):

a.　　　Analysis of concepts. Concepts are clarified by de- scribing the essential and core concepts beginning from the early developmental stages. Clarification allows other researchers to explore the topic in other fashions.

b.　　　Editing or compilation of documents, to preserve documents in chronological order to explain events. For ex- ample, an edition of Butler's park standards, the National Recreation and Park Association's first minutes, or letters from early pioneers in the field preserves the documents for future researchers.

c.　　　Descriptive narration tells the story from beginning to end in chronological order, utilizing limited generalizations and synthesized facts.

d.　　　Interpretive analysis relates one event to another event. The event is studied and described within a broader con- text to add meaning and credibility to the data. For example, an examination of the development of a local jurisdiction's ability to dedicate land for parks may be related to the urbanization and loss of open space in our communities.

e.　　　Comparative analysis examines similarities and differences in events during different time periods-for example, the budget-cutting priorities and procedures of the Proposition 13 era of the early 1980s in parks and recreation as compared to the budget-cutting priorities and procedures of today.

f.   Theoretical and philosophical analysis utilizes historical parallels, past trends, and sequences of events to suggest the past, present, and future of the topic being researched. Findings would be used to develop a theory or philosophy of leisure. For example, an analysis of public recreation agency goals and objectives of previous eras can be used to describe the future in the context of social, political, economic, technological, and cultural changes in society.

Step 4: Utilizing external and internal criticism, the re- search should evaluate the data. Sources of data include documents (letters, diaries, bills, receipts, newspapers, journals/magazines, films, pictures, recordings, personal and institutional records, and budgets), oral testimonies of participants in the events, and relics (textbooks, buildings, maps, equipment, furniture, and other objects).

Step 5: Reporting of the findings, which includes a statement of the problem, review of source material, assumptions, research questions and methods used to obtain findings, the interpretations and conclusions, and a thorough bibliographic referencing system.

Multimethod Approach

The multimethod approach encourages collecting, analyzing and integrating data from several sources and the use of a variety of different types of research methods.

Module 2: Study Design and Sampling

Study Design

Cross-sectional studies are simple in design and are aimed at finding out the prevalence of a phenomenon, problem, attitude or issue by taking a snap-shot or cross-section of the population. This obtains an overall picture as it stands at the time of the study. For example, a cross-sectional design would be used to assess demographic characteristics or community attitudes. These studies usually involve one contact with the study population and are relatively cheap to undertake.

Pre-test/post-test studies measure the change in a situation, phenomenon, problem or attitude. Such studies are often used to measure the efficacy of a program. These studies can be seen as a variation of the cross-sectional design as they involve two sets of cross-sectional data collection on the same population to determine if a change has occurred.

Retrospective studies investigate a phenomenon or issue that has occurred in the past. Such studies most often involve secondary data collection, based upon data available from previous studies or databases. For example, a retrospective study would be needed to examine the relationship between levels of unemployment and street crime in NYC over the past 100 years.

Prospective studies seek to estimate the likelihood of an event or problem in the future. Thus, these studies attempt to predict what the outcome of an event is to be. General science experiments are often

classified as prospective studies because the experimenter must wait until the experiment runs its course in order to examine the effects. Randomized controlled trials are always prospective studies and often involve following a "cohort" of individuals to determine the relationship between various variables.

Longitudinal studies follow study subjects over a long period of time with repeated data collection throughout. Some longitudinal studies last several months, while others can last decades. Most are observational studies that seek to identify a correlation among various factors. Thus, longitudinal studies do not manipulate variables and are not often able to detect causal relationships.

Sample

Once the researcher has chosen a hypothesis to test in a study, the next step is to select a pool of participants to be in that study. However, any research project must be able to extend the implications of the findings beyond the participants who actually participated in the study. For obvious reasons, it is nearly impossible for a researcher to study every person in the population of interest. In the example that we have been using thus far, the population of interest is "the developing world." The researcher must therefore make a decision to limit the research to a subset of that population, and this has important implications for the applicability of study results. The researcher must put some careful forethought into exactly how and why a certain group of individuals will be studied.(1)

Sampling Methods

Probability Sampling refers to sampling when the chance of any given individual being selected is known and these individuals are sampled independently of each other. This is also known as random sampling. A researcher can simply use a random number generator to choose participants (known as simple random sampling), or every nth individual (known as systematic sampling) can be included. Researchers also may break their target population into strata, and then apply these techniques within each strata to ensure that they are getting enough participants from each strata to be able to draw conclusions. For example, if there are several ethnic communities in one geographical area that a researcher wishes to study, that researcher might aim to have 30 participants from each group, selected randomly from within the groups, in order to have a good representation of all the relevant groups.

Non-Probability Sampling, or convenience sampling, refers to when researchers take whatever individuals happen to be easiest to access as participants in a study. This is only done when the processes the researchers are testing are assumed to be so basic and universal that they can be generalized beyond such a narrow sample.(2) For example, snowball sampling is an approach for

locating information-rich key informants.(3) Using this approach, a few potential respondents are contacted and asked whether they know of anybody with the characteristics that you are looking for in your research. Snowball sampling is not a stand-alone tool; the tool is a way of selecting participants and then using other tools, such as interviews or surveys.

Sampling Challenges

Because researchers can seldom study the entire population, they must choose a subset of the population, which can result in several types of error. Sometimes, there are discrepancies between the sample and the population on a certain parameter that are due to random differences. This is known as sampling error and can occur through no fault of the researcher.

Far more problematic is systematic error, which refers to a difference between the sample and the population that is due to a systematic difference between the two rather than random chance alone. The response rate problem refers to the fact that the sample can become self-selecting, and that there may be something about people who choose to participate in the study that affects one of the variables of interest. For example, in our eye care case, we may experience this kind of error if we simply sample those who choose to come to an eye clinic for a free eye exam as our experimental group and those who have poor eyesight but do not seek eye care as our control group. It is very possible in this situation that the people who actively seek help happen to be more proactive than those who do not. Because these two groups vary systematically on an attribute that is not the dependent variable (economic productivity), it is very possible that it is this difference in personality trait and not the independent variable (if they received corrective lenses or not) that produces any effects that the researcher observes on the dependent variable. This would be considered a failure in internal validity.

Another type of systematic sampling error is coverage error, which refers to the fact that sometimes researchers mistakenly restrict their sampling frame to a subset of the population of interest. This means that the sample they are studying varies systematically from the population for which they wish to generalize their results. For example, a researcher may seek to generalize the results to the "population of developing countries," yet may have a coverage error by sampling only heavily urban areas. This leaves out all of the more rural populations in developing countries, which have very different characteristics than the urban populations on several parameters. Thus, the researcher could not appropriately generalize the results to the broader population and would therefore have to restrict the conclusions to populations in urban areas of developing countries.(4)

First and foremost, a researcher must think very carefully about the population that will be included in the study and how to sample that population. Errors in sampling can often be avoided by good planning

and careful consideration. However, in order to improve a sampling frame, a researcher can always seek more participants. The more participants a study has, the less likely the study is to suffer from sampling error. In the case of the response rate problem, the researcher can actively work on increasing the response rate, or can try to determine if there is in fact a difference between those who partake in the study and those who do not. The most important thing for a researcher to remember is to eliminate any and all variables that the researcher cannot control. While this is nearly impossible in field research, the closer a researcher comes to isolating the variable of interest, the better the results.(5)