SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF SCIENCE AND HUMANITIES

DEPARTMENT OF VISUAL COMMUNICATION

UNIT – I – Introduction to Communication Research – SVC1301

I. INTRODUCTION TO RESEARCH

Meaning of research

The Department of Education and Training defines research as follows:

Research is defined as the creation of new knowledge and/or the use of existing knowledge in a new and creative way so as to generate new concepts, methodologies and understandings. This could include synthesis and analysis of previous research to the extent that it leads to new and creative outcomes.

Objectives of Research

• To gain familiarity with a phenomenon • To portray accurately the characteristics of a particular individual, situation or a group • To test a hypothesis • To find out the cause of a phenomenon • To establish the cause and effect relationship between variable • To explore new ideas and thoughts • To experiment new methods of problem solving.

Types of Research

• Descriptive research • Ex-post facto research • Applied research • Conceptual research • Qualitative research • Quantitative research • Empirical research • Exploratory research • Historical research

Although the research works and studies differ in their form and kind, they all still meet on the common ground of scientific methods employed by them. Hence, scientific research is expected to satisfy the following criteria:

i. The aim of the research should be clearly mentioned, along with the use of common concepts.

ii. The procedures used in the research should be adequately described, in order to permit another researcher to repeat the research for further advancement, while maintaining the continuity of what has already been done.

iii. The researchs procedural design should be carefully planned to obtain results that are as objective as possible.

iv. The flaws in the procedural design should be sincerely reported by the researcher to correctly estimate their effects upon the findings.

v. The data analysis should be adequate to reveal its significance.

vi. The methods used during the analysis should be appropriate.

vii. The reliability and validity of the concerned data should be checked carefully.

viii. The conclusions are needed to be confined and limited to only those data, which are justified and adequately provided by the research.

ix. In case, the researcher is experienced and has a good reputation in the field of research, greater confidence in research is warranted.

In other words, we can state the qualities of a good research" as following:

1) Systematic - This states that the research is structured with some specified steps, which are to be followed in a specified sequence, according to the well defined set of rules. Systematic characteristic of the research does not actually rule out creative thinking, but it does discourage the use of guessing and intuition in order to arrive at conclusions.

2) Logical - This states that the research is guided by the rules of logical reasoning, and that the logical process of induction and deduction are essential while conducting a research. Induction is the process of reasoning from a part to the whole; while, deduction is the process of reasoning from some premise to a conclusion that follows from that very premise. Besides, logical reasoning enables the research to be more meaningful in the context of decision making.

3) Empirical - This states that the research is basically related to one or more aspects of a real situation. Moreover, it deals with the concrete data, which provides a base for the external validity of research results.

4) Replicable - This states that the research results should be allowed verification by replicating their study, to thus build a sound basis for decisions.

For over three decades researchers and practitioners have depended on Communication

Research for the most up-to-date, comprehensive and important research on communication and its related fields. Important, In-Depth Research and Scholarship Communication processes are a fundamental part of virtually every aspect of human social life. Communication Research publishes articles that explore the processes, antecedents, and consequences of communication in a broad range of societal systems. Although most of the published articles are empirical, we also consider overview/review articles. These include the following:

mass media

interpersonal

health

political

entertainment

advertising/persuasive communication

new technology, online, computer-mediated and mobile communication

organizational

intercultural

group

family

Communication Research takes you to the cutting-edge of research and theory in all areaswithin the field of communication. It serves as the international forum aimed at the academic or professional interested in current research in communication and its related fields.

Why you need Communication Research

Research and theory presented in all areas of communication give you comprehensive coverage of the field Rigorous, empirical analysis provides you with research that's reliable and high in quality The multi-disciplinary perspective contributes to a greater understanding of communication processes and outcomes "Themed issues" bring you in-depth examinations of a specific area of importance, as thematically connected articles selected in the standard peer-review process are conveniently presented in a single issue Expert editorial guidance represents a wide range of interests from inside and outside the traditional boundaries of the communication discipline

What are the main types of quantitative approaches to research?

It is easier to understand the different types of quantitative research designs if you consider how the researcher designs for control of the variables in the investigation.

If the researcher views quantitative design as a continuum, one end of the range represents a design where the variables are not controlled at all and only observed. Connections amongst variable are only described. At the other end of the spectrum, however, are designs which include a very close control of variables, and relationships amongst those variables are clearly

established. In the middle, with experiment design moving from one type to the other, is a range which blends those two extremes together.

There are four main types of Quantitative research: Descriptive, Co relational, Causal-Comparative/Quasi-Experimental, and Experimental Research.

Types of Quantitative Design

Descriptive research seeks to describe the current status of an identified variable. These research projects are designed to provide systematic information about a phenomenon. The researcher does not usually begin with an hypothesis, but is likely to develop one after collecting data. The analysis and synthesis of the data provide the test of the hypothesis. Systematic collection of information requires careful selection of the units studied and careful measurement of Co relational research attempts to determine the extent of a relationship between two or more variables using statistical data. In this type of design, relationships between and among a number of facts are sought and interpreted.

This type of research will recognize trends and patterns in data, but it does not go so far in its analysis to prove causes for these observed patterns. Cause and effect is not the basis of this type of observational research. The data, relationships, and distributions of variables Causal-

comparative/quasi experimental research attempts to establish cause effect relationships among the variables. These types of design are very similar to true experiments, but with some key differences. An independent variable is identified but not manipulated by the experimenter, and effects of the independent variable on the dependent variable are measured. The researcher does not randomly assign groups and must use ones that are naturally formed or pre-existing groups.

Identified control groups exposed to the treatment Experimental research, often called true experimentation, uses the scientific method to establish the cause-effect relationship among a group of variables that make up a study. The true experiment is often thought of as a laboratory study, but this is not always the case; a laboratory setting has nothing to do with it. A true experiment is any study where an effort is made to identify and impose control over all other variables except one. An independent variable is manipulated to determine the effects on the dependent variables. Subjects are randomly assigned to each variable.

Examples of Descriptive Research:

• A description of how second-grade students spend their time during summer vacation

• A description of the tobacco use habits of teenagers

• A description of how parents feel about the twelvemonth school year

• A description of the attitudes of scientists regarding global warming

• A description of the kinds of physical activities that typically occur in nursing homes, and how frequently each occurs

• A description of the extent to which elementary teachers use math manipulative are studied only. Variables are not manipulated; they are only identified and are studied as they occur in a natural setting. *Sometimes correlational research is considered a type of descriptive research, and not as its own type of research, as no variables are manipulated in the study.

Examples of Correlational Research:

• The relationship between intelligence and self-esteem

• The relationship between diet and anxiety

• The relationship between an aptitude test and success in an algebra course

• The relationship between ACT scores and the freshman grades

• The relationships between the types of activities used in math classrooms and student achievement

• The covariance of smoking and lung disease variable are studied and compared to groups who are not. When analyses and conclusions are made, determining causes must be done carefully, as other variables, both known and unknown, could still affect the outcome. A causal comparative designed study, described in a New York Times article, "The Case for $320,00 Kindergarten Teachers," illustrates how causation must be thoroughly assessed before firm relationships amongst variables can be made.

Examples of Co relational Research:

• The effect of preschool attendance on social maturity at the end of the first grade

• The effect of taking multivitamins on a students' school absenteeism

• The effect of gender on algebra achievement

• The effect of part-time employment on the achievement of high school students

• The effect of magnet school participation on student attitude • The effect of age on lung capacity experimental treatments rather than identified in naturally occurring groups

Examples of Experimental Research:

• The effect of a new treatment plan on breast cancer

• The effect of positive reinforcement on attitude toward school

• The effect of teaching with a cooperative group strategy or a traditional lecture approach on students' achievement

• The effect of a systematic preparation and support system on children who were scheduled for surgery on the amount of psychological upset and cooperation

• A comparison of the effect of personalized instruction vs. traditional instruction on computational skill

What is the basic methodology for a quantitative research design?

The overall structure for a quantitative design is based in the scientific method. It uses deductive reasoning, where the researcher forms an hypothesis, collects data in an investigation of the problem, and then uses the data from the investigation, after analysis is made and conclusions are shared, to prove the hypotheses not false or false.

The basic procedure of a quantitative design is:

1. Make your observations about something that is unknown, unexplained, or new. Investigate current theory surrounding your problem or issue.

2. Hypothesize an explanation for those observations.

3. Make a prediction of outcomes based on your hypotheses. Formulate a plan to test your prediction.

4. Collect and process your data. If your prediction was correct, go to step 5. If not, the hypothesis has been proven false. Return to step 2 to form a new hypothesis based on your new knowledge.

5. Verify your findings. Make your final conclusions. Present your findings in an appropriate form for your audience.

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF SCIENCE AND HUMANITIES

DEPARTMENT OF VISUAL COMMUNICATION

UNIT – II – Introduction to Communication Research – SVC1301

II.RESEARCH DESIGN

WHAT IS RESEARCH DESIGN?

In this chapter, three research design types are introduced:

Exploratory

Descriptive

Causal

There are many definitions of research design, but no single definition imparts the full range of important

aspects:

Research design constitutes the blueprint for the collection, measurement, and analysis of data.

Research design aids the researcher in the allocation of limited resources by posing crucial choices in

methodology.

Research design is the plan and structure of investigation so conceived as to obtain answers to research

questions. The plan is the overall scheme or program of the research. It includes an outline of what the

investigator will do from writing hypotheses and their operational implications to the final analysis of

data.

Research design expresses both the structure of the research problem—the framework, organization, or

configuration of the relationships among variables of a study—and the plan of investigation used to obtain

empirical evidence on those relationships.

Together, these definitions give the essentials of research design:

An activity- and time-based plan.

A plan always based on a research question.

A guide for selecting sources and types of information.

A framework for specifying the relationships among the study's variables.

A procedural outline for every research activity.

One of the project management tools used in mapping a research design is critical path method (CPM).

CPM depicts sequential and simultaneous activities and estimates schedules or timetables for each activity

and phase of the research project.

CLASSIFICATION OF DESIGNS

Early in any research study, one faces the task of selecting the design to use.

Following are the eight different descriptors of research design.

1- Degree of Research Question Crystallization

A study may be exploratory or formal.

The distinctions between the two are the (a) degree of structure, and (b) the immediate objective of the

study.

Exploratory studies tend toward loose structures, with the objective of discovering future research tasks.

– The immediate purpose is usually to develop hypotheses or questions for future research.

The formal study begins where the exploration leaves off—with a hypothesis or research question.

– It also involves precise procedures and data source specifications.

– The goal of a formal design is to test the hypotheses or answer the research questions posed.

2- Method of Data Collection

The method of data collection distinguishes between monitoring and communication processes.

Monitoring includes studies in which the researcher inspects the activities of a subject or the nature of some

material, without attempting to elicit responses from anyone.

Examples of monitoring include:

– Traffic counts at intersections

– License plates recorded in a restaurant parking lot

– A search of the library collection

– The State Farm Dangerous Intersection Study

In each case, the researcher notes and records the information available from observation.

In a communication study, the researcher questions the subjects and collects their responses by personal or

impersonal means.

Collected data may result from:

– Interview or telephone conversations.

– Self-administered or self-reported instruments through the mail, left in convenient locations, or

transmitted electronically, or by other means.

– Instruments presented before and/or after a treatment or stimulus condition in an experiment.

3- Researcher Control of Variables

In an experiment, the researcher attempts to control and/or manipulate the variables in the study.

Experimental design is appropriate when one wishes to discover whether certain variables produce effects

in other variables.

Experimentation provides the most powerful support possible for a hypothesis of causation.

With an ex post facto design, investigators have no control over the variables in the sense of being able to

manipulate them.

They can only report what has happened, or what is happening.

Researchers using this design must not influence the variables; doing so introduces bias.

The researcher is limited to holding factors constant by judicious selection of subjects, according to strict

sampling procedures and by statistical manipulation of findings.

4- The Purpose of the Study

The essential difference between reporting, descriptive, causal-explanatory and causal-predictive studies lies

in their objectives.

A reporting study provides a summation of data, often recasting data to achieve a deeper understanding

or to generate statistics for comparison.

A descriptive study is concerned with finding out who, what, where, when, or how much.

A causal-explanatory study is concerned with learning why. That is, how one variable produces changes

in another variable.

A causal-predictive study attempts to predict the effect on one variable by manipulating another variable

while holding all other variables constant.

5- The Time Dimension

Cross-sectional studies are carried out once, and represent a snapshot of one point in time.

Longitudinal studies are repeated over an extended period.

The advantage of a longitudinal study is that it can track changes over time.

In longitudinal panel studies, researchers may study people over time.

In marketing, panels are set up to report consumption data.

These data provide information on relative market share, consumer response to new products, and new

promotional methods.

Some types of information cannot be collected a second time from the same person without the risk of bias.

Some benefits of a longitudinal study can be revealed in a cross-sectional study by adroit questioning about

past attitudes, history, and future expectations.

6- The Topical Scope

Statistical studies are designed for breadth, rather than depth.

They attempt to capture a population's characteristics by making references from a sample's

characteristics.

Generalizations about findings are based on the relativity of the sample and the validity of the design.

Case studies place more emphasis on a full contextual analysis of fewer events or conditions, and their

interrelations.

The reliance on qualitative data makes support or rejection more difficult.

An emphasis on detail provides valuable insight for problem solving, evaluation, and strategy.

– This detail is secured from multiple sources of information.

– It allows evidence to be verified and avoids missing data.

Although they have a significant scientific role, case studies have been maligned as "scientifically

worthless" because they do not meet the minimum requirements for comparison.

– Important scientific propositions have the form the universals, which can be falsified by a single

counter-instance.

– A single, well-designed case study can provide a major challenge to a theory, and provide a source of

new hypotheses and constructs simultaneously.

7- The Research Environment

Designs differ as to whether they occur under actual environmental conditions (field conditions) or under

staged/manipulated conditions (laboratory conditions).

Simulations, which replicate the essence of a system or process, are increasingly used in research, especially

in operations research.

Conditions and relationships in actual situations are often represented in mathematical models.

Role-playing and other behavioral activities may also be viewed as simulations.

8- Participants' Perceptual Awareness

Participant's perceptual awareness refers to when people in a disguised study perceive that research is being

conducted.

Participant's perceptual awareness may reduce the usefulness of a research design.

Participants' perceptual awareness influence the outcomes of the research.

When participants believe that something out of the ordinary is happening, they may behave less naturally.

There are three levels of perceptional awareness:

a- Participants perceive no deviations from everyday routines (non-aware, unaffected).

b- Participants perceive deviations, but as unrelated to the researcher (aware, consciously unaffected).

c- Participants perceive deviations as researcher-induced (aware, consciously affected).

I- EXPLORATORY STUDIES

Exploration is particularly useful when researchers lack a clear idea of the problems they will meet during the

study.

Exploration allows researchers to:

– Develop clearer concepts

– Establish priorities

– Develop operational definitions

– Improve the final research design

– Possibly save time and money

If exploration reveals that a problem is not as important as first thought, more formal studies can be

cancelled.

Exploration serves other purposes as well:

The area of investigation may be so new or vague that the researcher needs to do an exploration just to

learn something about the dilemma.

Important variables may not be known or well defined.

A hypothesis for the research may be needed.

The researcher may need to determine if it is feasible to do a formal study.

Researchers and managers alike give exploration less attention than it deserves.

There is often pressure for a quick answer.

There may be a bias about qualitative research.

Subjectiveness

Non-representation

Non-systematic design.

Exploration can save time and money, so it should not be slighted.

Qualitative Techniques

Although both qualitative and quantitative techniques are applicable, exploration relies more heavily on

qualitative techniques.

There are multiple ways to investigate a management question, including:

Individual depth interviews: usually conversational, rather than structured.

Participant observation: perceive firsthand what participants experience

Films, photographs, and videotape: to capture the life of the group under study.

Projective techniques and psychological testing: such as a Thematic Apperception Test, projective

measures, games, or role-playing.

Case studies: for an in-depth contextual analysis of a few events or conditions.

Street ethnography: to discover how a cultural subgroups described and structures its world at street level.

Elite or expert interviewing: for information from influential or well-informed people.

Document analysis: to evaluate historical or contemporary confidential or public records, reports,

government documents, and opinions.

Proxemics and kinesics: to study the use of space and body-motion communication, respectively.

When these approaches are combined, four exploratory techniques emerge:

Secondary data analysis.

Experience surveys.

Focus groups.

Tow-stage designs.

Secondary Data Analysis

The first step in an exploratory study is a search of the secondary literature.

Studies made by others, for their own purposes, represent secondary data.

It is inefficient to discover anew through the collection of primary data or original research what has

already been done and reported.

Start with an organization's own archives.

By reviewing prior studies, you can identify methodologies that proved successful and unsuccessful.

Solutions that didn't receive attention in the past may reveal subjects for further study.

Avoid duplication in instances where prior data can help resolve the current dilemma.

The second source of secondary data is published documents prepared by authors outside the sponsor

organization.

Data from secondary sources help us decide what needs to be done, and can be a rich source of

hypotheses.

In many cases, you can conduct a secondary search in libraries, or via your computer and an online

service or an Internet gateway.

If we confine an investigation to obvious subjects in bibliographic sources, some of the best information

may be missed.

We provide a detailed list of Business Reference Sources on the website.

Experience Survey

Published data are seldom more than a fraction of the existing knowledge in a field.

A significant portion of what is known on a topic is proprietary to a given organization, and therefore

unavailable to an outside researcher.

Also, internal data archives are rarely well organized, making secondary sources difficult to locate.

Thus, it is beneficial to seek information from persons experienced in the field of study, tapping into their

memories and experiences.

In an experience survey, we seek a person's ideas about important issues or aspects of the subject and

discover what is important across the subject's range of knowledge.

Avenues to explore:

What is being done?

What has been tried in the past without success? With success?

How have things changed?

What are the change-producing elements of the situation?

Who is involved in decisions and what role does each person play?

What problem areas and barriers can be seen?

What are the costs of the processes under study?

Whom can we count on to assist and/or participate in the research?

What are the priority areas?

The product of such questioning may be:

A new hypothesis

The discarding of an old one

Information about the practicality of doing the study

Discovery of what facilities are available

What factors need to be controlled, and how

Who will cooperate in the study.

Discovery is more easily carried out if we analyze cases that provide special insight.

People who might provide insightful information include:

Newcomers to the scene

Marginal or peripheral individuals

Individuals in transition

Deviants and isolates

"Pure" cases

Those who fit well and those who do not

Those who represent different positions in the system

During the early phase of the MindWriter research study, Jason and Myra plan to interview:

Managers of the service facility

Managers of the call center

Managers of the contract courier service

Long-term employees of various departments

Individuals associated with engineering and production

Focus Groups

Focus groups became widely used in the 1980s.

A focus group is a group of people (typically 6 to 10), led by a trained moderator, who meet for 90 minutes to

2 hours.

The facilitator or moderator uses group dynamics to focus or guide the group in an exchange of ideas,

feelings, and experiences on a specific topic.

One typical objective of a focus group might be a new product or product concept, a new employee

motivation program, or improved production-line organization.

The basic output of the session is a list of ideas and behavioral observations, with recommendations by the

moderator, that are often used for later quantitative testing.

In another application, a large title insurance company ran focus groups with its branch office

administrations to discover their preferences for distributing files on the company's intranet.

Two-Stage Design

With a two-stage design approach, exploration becomes s separate first stage with limited objectives:

Clearly define the research question

Develop the research design

Argument for a two-stage approach: we need to know more about the problem before resources are

committed.

This approach is particularly useful when the research budget in inflexible.

A limited exploration for a specific, modest cost carries little risk for both the sponsor and the researcher.

An exploratory study is finished when researchers have achieved the following:

Major dimensions of the research task have been established.

A set of subsidiary investigative questions that can guide a detailed research design have been defined.

Several hypotheses about possible causes of a management dilemma have been developed.

Certain hypotheses have been identified as being so remote that they can be safely ignored.

A conclusion that additional research is not needed or is not feasible has been reached.

II- DESCRIPTIVE STUDIES

Formalized studies are typically structured with clearly states hypotheses or investigative questions.

Research objectives:

Descriptions of phenomena or characteristics associated with a subject population (the who, what, when,

where, and how of a topic).

Estimates of the proportions of a population that have these characteristics.

Discover of associations among different variables (sometimes labeled a correlational study)

A descriptive study may be simple or complex, and it may be done in many settings.

The simplest study concerns a univariate question or hypothesis in which we ask about (or state something

about) the size, form, distribution, or existence of a variable.

In the account analysis at BankChoice, we might want to develop a profile of savers. Examples of other

variables include:

– Number of accounts opened in the last six months

– Amount of account activity

– Size of accounts

– Number of accounts for minors

Our task is to determine if the variables are interdependent or unrelated. If they are, we must determine

the strength or magnitude of the relationship.

Descriptive studies are often much more complex than the BankChoice example.

III-CAUSAL STUDIES

Statistically untrained undividual sometimes mistake correlation (the simultaneous occurance of two

phenomena as causation.

The essential element of causation is that A "produces" B or A "forces" B to occur.

Empirically, we can never demonstrate an A-B causality with certainty.

– Empirical conclusions are inferences (inductive conclusions).

– As such, they are based on what we observe and measure.

– We cannot observe and measure all the processes that may account for the A-B relationship.

In Chapter 3 we discussed the example of sales failing to increase following a promotion. Having ruled

out other causes, we were left with the inference that was probably (but not certainly) the cause: a poorly

executed promotion.

Meeting the ideal standard of causation requires that one variable always causes another variable, and no

other variable has the same causal effect.

The method of agreement states that "When two or more cases of a given phenomenon have one and only one

condition in common, then that condition may be regarded as the cause (or effect) of the phenomenon."

The method of agreement helps rule out some variables as irrelevant. A, B, D, and E are unlikely to be

causes of Z. However, there is an implicit assumption that there are no variables to consider.

– No one can accept this supposition with certainty because the number of variables is infinite.

– In addition, while C may be the cause, it may instead function only in the presence of some other

variable, which was not included.

– The negative canon of agreement states that where the absence of C is associated with the absence of

Z, there is evidence of a causal relationship between C and Z.

– Together with the method of agreement, this forms the basis for the method of difference: "If there are

two or more cases, and in one of them observation Z can be made, while in the other it cannot; and if

variable C occurs when observation Z can be made, while in the other it cannot; and if variable C

occurs when observation Z is made, and does not occur when observation Z is not made; then it can

be asserted that there is a causal relationship between C and Z.

No one can ever be certain that variable A causes variable B to occur, but one can gather evidence that

increases the belief that A leads to B.

We seek three types of evidence of causation:

Covariation between A and B.

– Do we find that A and B occur together in the way hypothesized?

– When A does not occur, is there also an absence of B?

– When there is more or less of A, does one also find more of less of B?

Time order of events moving in the hypothesized direction.

– Does A occur before B?

No other possible causes of B.

– No one can determine that C, D, and E do not covary with B in a way that suggests possible causal

connections?

In addition to the three conditions above, successful inference making from experimental designs must meet

two additional requirements:

Control: all factors, with the exception of the independent variable, must be held constant and not

confound with another variable that is not part of the study.

Random assignment: each person must have an equal chance for exposure to each level of the

independent variable.

If we consider the possible relationships that can occur between two variables, we can conclude that there are

three possibilities:

Symmetrical

Reciprocal

Asymmetrical

A symmetrical relationship is one in which two variables fluctuate together, but we assume the changes in

neither variable are due to changes in the other.

Symmetrical conditions are most often found when two variables are alternate indicators of another cause

or independent variable.

We might conclude that a correlation between low work attendance and active participation in a camping

club is the result of (dependent on) another factor, such as a lifestyle preference.

A reciprocal relationship exists when two variables mutually influence or reinforce each other.

This could occur if reading an advertisement leads to the use of a product. The usage, in turn, sensitizes

the person to notice and read more of the advertising for that product.

Most research analysts look for asymmetrical relationships.

With these, we postulate that changes in one independent variable (IV) are responsible for changes in a

dependent variable (DV).

The identification of the IV and DV is often obvious, but sometimes the choice is not clear. In these cases,

dependence and independence should be evaluated on the basis of:

– The degree of which each variable may be altered.

– The time order between the variables.

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF SCIENCE AND HUMANITIES

DEPARTMENT OF VISUAL COMMUNICATION

UNIT – III – Introduction to Communication Research – SVC1301

III.METHOD OF DATA COLLECTION

Quantitative data collection methods rely on random sampling and structured data collection instruments that fit diverse experiences into predetermined response categories. They produce results that are easy to summarize, compare, and generalize. Quantitative research is concerned with testing hypotheses derived from theory and/or being able to estimate the size of a phenomenon of interest. Depending on the research question, participants may be randomly assigned to different treatments. If this is not feasible, the researcher may collect data on participant and situational characteristics in order to statistically control for their influence on the dependent, or outcome, variable. If the intent is to generalize from the research participants to a larger population, the researcher will employ probability sampling to select participants. A probability sampling method is any method of sampling that utilizes some form of random selection (http://www.socialresearchmethods.net/kb/sampprob.php). In order to have a random selection method, you must set up some process or procedure that assures that the different units in your population have equal probabilities of being chosen. Simple examples of random selection are picking a name out of a hat or choosing the short straw. Computers are used for generating random selection in more complex projects. Typical quantitative data gathering strategies include: Administering surveys with closed-ended questions (e.g., face-to face and telephone • interviews, mail questionnaires, etc.) (http://www.achrn.org/quantitative_methods.htm) Experiments/clinical trials. • Observing and recording well-defined events (e.g., counting the number of patients • waiting in emergency at specified times of the day). Obtaining relevant data from management information systems. • INTERVIEWS In Quantitative research (survey research), interviews are more structured than in Qualitative research. http://www.stat.ncsu.edu/info/srms/survpamphlet.html In a structured interview, the researcher asks a standard set of questions and nothing more. (Leedy and Ormrod, 2001) 5 Telephone interviews Advantages: Less time consuming• Less expensive • Researcher has ready access to anyone who has a landline telephone. • Higher response rate than the mail questionnaire. • Can be fully automated using CATI (Computer Assisted Telephone Interviewing) saving• data processing time. Disadvantages: The response rate is not as high as the face-to-face interview. • The sample may be biased as only those people who have landline phones are• contacted (excludes people who do not have a phone, or only have cell phones). Face-to-face interviews (Leedy and Ormrod, 2001) Advantages: Enables the researcher to establish rapport with potential participants and therefore • gain their cooperation. Yields the highest response rates in survey research. • Allows the researcher to clarify ambiguous answers and when appropriate, seek follow- • up information. Disadvantages: Impractical when large samples are involved• Can be time consuming and expensive. • Computer Assisted Personal Interviewing (CAPI): is a form of personal interviewing, but instead of completing a questionnaire, the interviewer brings along a laptop or hand-held computer to enter the information directly into the database. Advantages: Saves time involved in processing the data. • Saves the interviewer from carrying around hundreds of questionnaires. • Disadvantages: Can be expensive to set up.• Requires that interviewers have computer and typing skills. • QUESTIONNAIRES Questionnaires often

make use of checklist and rating scales. These devices help simplify and quantify people's behaviors and attitudes. A checklist is a list of behaviors, characteristics, or other entities the researcher is looking for. Either the researcher or survey participant simply checks whether each item on the list is observed, present or true or vice versa. A rating scale is 6 more useful when a behavior needs to be evaluated on a continuum. They are also known as Likert scales. (Leedy and Ormrod, 2001) Mail questionnaires Advantages: Can be sent to a large number of people.• Saves the researcher time and money compared to interviewing. • People are more truthful while responding to the questionnaires regarding controversial• issues in particular due to the fact that their responses are anonymous. Allow the respondent to answer at their leisure.• Disadvantages: In most cases, the majority of people who receive questionnaires don't return them. • Therefore: o Over-sampling may be necessary if doing a one-time mail out in order to get enough completed questionnaires to be generalizable to the population. o Follow-up reminders to participants encouraging them to complete the questionnaire may be necessary, thereby increasing the time and cost to conduct the study. o May need to offer incentives to increase response rate. Time – mail surveys take longer than other types of surveys. • Web-based questionnaires: A new and inevitably growing methodology is the use of Internet based research. This would mean receiving an e-mail on which you would click on an address that would take you to a secure web-site to fill in a questionnaire. Advantages: This type of research is often quicker and less detailed. • Very cost effective.• Disadvantages: Excludes people who do not have a computer or are unable to access a computer. • Need to have access to email addresses.• Many worksites have screening mechanisms in place blocking access to employee• emails. The validity of such surveys may be in question as people might be in a hurry to• complete it and so might not give accurate responses. (http://www.statcan.ca/english/edu/power/ch2/methods/methods.htm)

Quantitative data collection methods rely on random sampling and structured data collection instruments that fit diverse experiences into predetermined response categories. They produce results that are easy to summarize, compare, and generalize. Quantitative research is concerned with testing hypotheses derived from theory and/or being able to estimate the size of a phenomenon of interest. Depending on the research question, participants may be randomly assigned to different treatments. If this is not feasible, the researcher may collect data on participant and situational characteristics in order to statistically control for their influence on the dependent, or outcome, variable. If the intent is to generalize from the research participants to a larger population, the researcher will employ probability sampling to select participants.

A probability sampling method is any method of sampling that utilizes some form of random selection (http://www.socialresearchmethods.net/kb/sampprob.php). In order to have a random selection method, you must set up some process or procedure that assures that the different units in your population have equal probabilities of being chosen. Simple examples of random selection are picking a name out of a hat or choosing the short straw. Computers are

used for generating random selection in more complex projects.

Typical quantitative data gathering strategies include:

Administering surveys with closed-ended questions (e.g., face-to face and telephone interviews, mail questionnaires, etc.)

(http://www.achrn.org/quantitative_methods.htm)

Experiments/clinical trials.

Observing and recording well-defined events (e.g., counting the number of patients waiting in emergency at specified times of the day).

Obtaining relevant data from management information systems.

INTERVIEWS

In Quantitative research (survey research), interviews are more structured than in Qualitative

research. http://www.stat.ncsu.edu/info/srms/survpamphlet.html

In a structured interview, the researcher asks a standard set of questions and nothing more. (Leedy and Ormrod, 2001)

5

Telephone interviews

Advantages:

Less time consuming

Less expensive

Researcher has ready access to anyone who has a landline telephone.

Higher response rate than the mail questionnaire.

Can be fully automated using CATI (Computer Assisted Telephone Interviewing) saving data processing time.

Disadvantages:

The response rate is not as high as the face-to-face interview.

The sample may be biased as only those people who have landline phones are contacted (excludes people who do not have a phone, or only have cell phones).

Face-to-face interviews (Leedy and Ormrod, 2001)

Advantages:

Enables the researcher to establish rapport with potential participants and therefore gain their cooperation.

Yields the highest response rates in survey research.

Allows the researcher to clarify ambiguous answers and when appropriate, seek follow-up information.

Disadvantages:

Impractical when large samples are involved

Can be time consuming and expensive.

Computer Assisted Personal Interviewing (CAPI): is a form of personal interviewing, but instead of completing a questionnaire, the interviewer brings along a laptop or hand-held computer to enter the information directly into the database.

Advantages:

Saves time involved in processing the data.

Saves the interviewer from carrying around hundreds of questionnaires.

Disadvantages:

Can be expensive to set up.

Requires that interviewers have computer and typing skills.

QUESTIONNAIRES

Questionnaires often make use of checklist and rating scales. These devices help simplify and quantify people's behaviors and attitudes. A checklist is a list of behaviors, characteristics, or other entities the researcher is looking for. Either the researcher or survey participant simply

checks whether each item on the list is observed, present or true or vice versa. A rating scale is

6

more useful when a behavior needs to be evaluated on a continuum. They are also known as Likert scales. (Leedy and Ormrod, 2001)

Mail questionnaires

Advantages:

Can be sent to a large number of people.

Saves the researcher time and money compared to interviewing.

People are more truthful while responding to the questionnaires regarding controversial issues in particular due to the fact that their responses are anonymous.

Allow the respondent to answer at their leisure.

Disadvantages:

In most cases, the majority of people who receive questionnaires don't return them.

Therefore:

o Over-sampling may be necessary if doing a one-time mail out in order to get enough completed questionnaires to be generalizable to the population.

o Follow-up reminders to participants encouraging them to complete the questionnaire may be necessary, thereby increasing the time and cost to conduct the study.

o May need to offer incentives to increase response rate.

Time – mail surveys take longer than other types of surveys.

Web-based questionnaires:

A new and inevitably growing methodology is the use of Internet based research. This would mean receiving an e-mail on which you would click on an address that would take you to a secure web-site to fill in a questionnaire.

Advantages:

This type of research is often quicker and less detailed.

Very cost effective.

Disadvantages:

Excludes people who do not have a computer or are unable to access a computer.

Need to have access to email addresses.

Many worksites have screening mechanisms in place blocking access to employee emails.

The validity of such surveys may be in question as people might be in a hurry to complete it and so might not give accurate responses.

(http://www.statcan.ca/english/edu/power/ch2/methods/methods.htm)

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF SCIENCE AND HUMANITIES

DEPARTMENT OF VISUAL COMMUNICATION

UNIT – IV – Introduction to Communication Research – SVC1301

IV.SAMPLING

Sampling

Sampling is the process of selecting units (e.g., people, organizations) from a population of interest so that by studying the sample we may fairly generalize our results back to the population from which they were chosen. Let's begin by covering some of the key terms in sampling like "population" and "sampling frame." Then, because some types of sampling rely upon quantitative models, we'll talk about some of the statistical terms used in sampling. Finally, we'll discuss the major distinction between probability and Nonprobability sampling methods and work through the major types in each.

Sampling Terminology

As with anything else in life you have to learn the language of an area if you're going to ever hope to use it. Here, I want to introduce several different terms for the major groups that are involved in a sampling process and the role that each group plays in the logic of sampling.

The major question that motivates sampling in the first place is: "Who do you want to generalize to?" Or should it be: "To whom do you want to generalize?" In most social research we are interested in more than just the people who directly participate in our study. We would like to be able to talk in general terms and not be confined only to the people who are in our study. Now, there are times when we aren't very concerned about generalizing. Maybe we're just evaluating a program in a local agency and we don't care whether the program would work with other people in other places and at other times. In that case, sampling and generalizing might not be of interest. In other cases, we would really like to be able to generalize almost universally. When psychologists do research, they are often interested in developing theories that would hold for all humans. But in most applied social research, we are interested in generalizing to specific groups. The group you wish to generalize to is often called the population in your study. This is the group you would like to sample from because this is the group you are interested in generalizing to. Let's imagine that you wish to generalize to urban homeless males between the ages of 30 and 50 in the United States. If that is the population of interest, you are likely to have a very hard time developing a reasonable sampling plan. You are probably not going to find an accurate listing of this population, and even if you did, you would almost certainly not be able to mount a national sample across hundreds of urban areas. So we probably should make a distinction between the population you would like to generalize to, and the population that will be accessible to you. We'll call the former the theoretical population and the latter the accessible population. In this example, the accessible population might be homeless males between the ages of 30 and 50 in six selected urban areas across the U.S.

Once you've identified the theoretical and accessible populations, you have to do one more thing before you can actually draw a sample – you have to get a list of the members of the accessible population. (Or, you have to spell out in detail how you will contact them to assure representativeness). The listing of the accessible population from which you'll draw your sample is called the sampling frame. If you were doing a phone survey and selecting names from the telephone book, the book would be your sampling frame. That wouldn't be a great way to sample because significant subportions of the population either don't have a phone or

have moved in or out of the area since the last book was printed. Notice that in this case, you might identify the area code and all three-digit prefixes within that area code and draw a sample simply by randomly dialingnumbers (cleverly known as random-digit-dialing). In this case, the sampling frame is not a list per se, but is rather a procedure that you follow as the actual basis for sampling. Finally, you actually draw your sample (using one of the many sampling procedures). The sample is the group of people who you select to be in your study. Notice that I didn't say that the sample was the group of people who are actually in your study. You may not be able to contact or recruit all of the people you actually sample, or some could drop out over the course of the study. The group that actually completes your study is a subsample of the sample – it doesn't include nonrespondents or dropouts. The problem of nonresponse and its effects on a study will be addressed when discussing "mortality" threats to internal validity.

People often confuse what is meant by random selection with the idea of random assignment. You should make sure that you understand the distinction between random selection and random assignment.

At this point, you should appreciate that sampling is a difficult multi-step process and that there are lots of places you can go wrong. In fact, as we move from each step to the next in identifying a sample, there is the possibility of introducing systematic error or bias. For instance, even if you are able to identify perfectly the population of interest, you may not have access to all of them. And even if you do, you may not have a complete and accurate enumeration or sampling frame from which to select. And, even if you do, you may not draw the sample correctly or accurately. And, even if you do, they may not all come and they may not all stay. Depressed yet? This is a very difficult business indeed. At times like this I'm reminded of what Donald Campbell used to say (I'll paraphrase here): "Cousins to the amoeba, it's amazing that we know anything at all!"

Statistical Terms in Sampling

Let's begin by defining some very simple terms that are relevant here. First, let's look at the results of our sampling efforts. When we sample, the units that we sample – usually people – supply us with one or more responses. In this sense, a response is a specific measurement value that a sampling unit supplies. In the figure, the person is responding to a survey instrument and gives a response of 4. When we look across the responses that we get for our entire sample, we use a statistic. There are a wide variety of statistics we can use – mean, median, mode, and so on. In this example, we see that the mean or average for the sample is 3.75. But the reason we sample is so that we might get an estimate for the population we sampled from. If we could, we would much prefer to measure the entire population. If you measure the entire population and calculate a value like a mean or average, we don't refer to this as a statistic, we call it a parameter of the population.

The Sampling Distribution

So how do we get from our sample statistic to an estimate of the population parameter? A crucial midway concept you need to understand is the sampling distribution. In order to understand it, you have to be able and willing to do a thought experiment. Imagine that instead of just taking a single sample like we do in a typical study, you took three independent samples of the same population. And furthermore, imagine that for each of your three samples, you collected a single response and computed a single statistic, say, the mean of the response. Even though all three samples came from the same population, you wouldn't expect to get the exact same statistic from each. They would differ slightly just due to the random "luck of the draw" or to the natural fluctuations or vagaries of drawing a sample. But you would expect that all three samples would yield a similar statistical estimate because they were drawn from the same population. Now, for the leap of imagination! Imagine that you did an infinite number of samples from the same population and computed the average for each one. If you plotted them on a histogram or bar graph you should find that most of them converge on the same central value and that you get fewer and fewer samples that have averages farther away up or down from that central value. In other words, the bar graph would be well described by the bell curve shape that is an indication of a "normal" distribution in statistics. The distribution of an infinite number of samples of the same size as the sample in your study is known as the sampling distribution.

We don't ever actually construct a sampling distribution. Why not? You're not paying attention! Because to construct it we would have to take an infinite number of samples and at least the last time I checked, on this planet infinite is not a number we know how to reach. So why do we even talk about a sampling distribution? Now that's a good question! Because we need to realize that our sample is just one of a potentially infinite number of samples that we could have taken. When we keep the sampling distribution in mind, we realize that while the statistic we got from our sample is probably near the center of the sampling distribution (because most of the samples would be there) we could have gotten one of the extreme samples just by the luck of the draw. If we take the average of the sampling distribution – the average of the averages of an infinite number of samples – we would be much closer to the true population average – the parameter of interest. So the average of the sampling distribution is essentially equivalent to the parameter. But what is the standard deviation of the sampling distribution (OK, never had statistics? There are any number of places on the web where you can learn about them or even just brush up if you've gotten rusty. This isn't one of them. I'm going to assume that you at least know what a standard deviation is, or that you're capable of finding out relatively quickly). The standard deviation of the sampling distribution tells us something about how different samples would be distributed. In statistics it is referred to as the standard error (so we can keep it separate in our minds from standard deviations. Getting confused? Go get a cup of coffee and come back in ten minutes…OK, let's try once more… A standard deviation is the spread of the scores around the average in a single sample. The standard error is the spread of the averages around the average of averages in a sampling distribution. Got it?)

Sampling Error

In sampling contexts, the standard error is called sampling error. Sampling error gives us some idea of the precision of our statistical estimate. A low sampling error means that we had relatively less variability or range in the sampling distribution. But here we go again – we never actually see the sampling distribution! So how do we calculate sampling error? We base our calculation on the standard deviation of our sample. The greater the sample standard deviation, the greater the standard error (and the sampling error). The standard error is also related to the sample size. The greater your sample size, the smaller the standard error. Why? Because the greater the sample size, the closer your sample is to the actual population itself. If you take a sample that consists of the entire population you actually have no sampling error because you don't have a sample, you have the entire population. In that case, the mean you estimate is the parameter.

The 68, 95, 99 Percent Rule

You've probably heard this one before, but it's so important that it's always worth repeating… There is a general rule that applies whenever we have a normal or bell-shaped distribution. Start with the average – the center of the distribution. If you go up and down (i.e., left and right) one standard unit, you will include approximately 68% of the cases in the distribution (i.e., 68% of the area under the curve). If you go up and down two standard units, you will include approximately 95% of the cases. And if you go plus-and-minus three standard units, you will include about 99% of the cases. Notice that I didn't specify in the previous few sentences whether I was talking about standard deviation units or standard error units. That's because the same rule holds for both types of distributions (i.e., the raw data and sampling distributions). For instance, in the figure, the mean of the distribution is 3.75 and the standard unit is .25 (If this was a distribution of raw data, we would be talking in standard deviation units. If it's a sampling distribution, we'd be talking in standard error units). If we go up and down one standard unit from the mean, we would be going up and down .25 from the mean of 3.75. Within this range – 3.5 to 4.0 – we would expect to see approximately 68% of the cases. This section is marked in red on the figure. I leave to you to figure out the other ranges. But what does this all mean you ask? If we are dealing with raw data and we know the mean and standard deviation of a sample, we can predict the intervals within which 68, 95 and 99% of our cases would be expected to fall. We call these intervals the – guess what – 68, 95 and 99% confidence intervals.

Now, here's where everything should come together in one great aha! experience if you've been following along. If we had a sampling distribution, we would be able to predict the 68, 95 and 99% confidence intervals for where the population parameter should be! And isn't that why we sampled in the first place? So that we could predict where the population is on that variable? There's only one hitch. We don't actually have the sampling distribution (now this

is the third time I've said this in this essay)! But we do have the distribution for the sample itself. And we can from that distribution estimate the standard error (the sampling error) because it is based on the standard deviation and we have that. And, of course, we don't actually know the population parameter value – we're trying to find that out – but we can use our best estimate for that – the sample statistic. Now, if we have the mean of the sampling distribution (or set it to the mean from our sample) and we have an estimate of the standard error (we calculate that from our sample) then we have the two key ingredients that we need for our sampling distribution in order to estimate confidence intervals for the population parameter.

Perhaps an example will help. Let's assume we did a study and drew a single sample from the population. Furthermore, let's assume that the average for the sample was 3.75 and the standard deviation was .25. This is the raw data distribution depicted above. now, what would the sampling distribution be in this case? Well, we don't actually construct it (because we would need to take an infinite number of samples) but we can estimate it. For starters, we assume that the mean of the sampling distribution is the mean of the sample, which is 3.75. Then, we calculate the standard error. To do this, we use the standard deviation for our sample and the sample size (in this case N=100) and we come up with a standard error of .025 (just trust me on this). Now we have everything we need to estimate a confidence interval for the population parameter. We would estimate that the probability is 68% that the true parameter value falls between 3.725 and 3.775 (i.e., 3.75 plus and minus .025); that the 95% confidence interval is 3.700 to 3.800; and that we can say with 99% confidence that the population value is between 3.675 and 3.825. The real value (in this fictitious example) was 3.72 and so we have correctly estimated that value with our sample.

Probability Sampling

A probability sampling method is any method of sampling that utilizes some form of random selection. In order to have a random selection method, you must set up some process or procedure that assures that the different units in your population have equal probabilities of being chosen. Humans have long practiced various forms of random selection, such as picking a name out of a hat, or choosing the short straw. These days, we tend to use computers as the mechanism for generating random numbers as the basis for random selection.

Some Definitions

Before I can explain the various probability methods we have to define some basic terms. These are:

N is the number of cases in the sampling frame

n is the number of cases in the sample

NCn = the number of combinations (subsets) of n from N

f = n/N is the sampling fraction

That's it. With those terms defined we can begin to define the different probability sampling methods.

Simple Random Sampling

The simplest form of random sampling is called simple random sampling. Pretty tricky, huh? Here's the quick description of simple random sampling:

Objective: To select n units out of N such that each NCn has an equal chance of being selected.

Procedure: Use a table of random numbers, a computer random number generator, or a mechanical device to select the sample.

A somewhat stilted, if accurate, definition. Let's see if we can make it a little more real.

How do we select a simple random sample? Let's assume that we are doing some research with a small service agency that wishes to assess clients' views of quality of service over the past year. First, we have to get the sampling frame organized. To accomplish this, we'll go through agency records to identify every client over the past 12 months. If we're lucky, the agency has good accurate computerized records and can quickly produce such a list. Then, we have to actually draw the sample. Decide on the number of clients you would like to have in the final sample. For the sake of the example, let's say you want to select 100 clients to survey and that there were 1000 clients over the past 12 months. Then, the sampling fraction is f = n/N = 100/1000 = .10 (or 10%). Now, to actually draw the sample, you have several options. You could print off the list of 1000 clients, tear then into separate strips, put the strips in a hat, mix them up real good, close your eyes and pull out the first 100. But this mechanical procedure would be tedious and the quality of the sample would depend on how thoroughly you mixed them up and how randomly you reached in. Perhaps a better procedure would be to use the kind of ball machine that is popular with many of the state lotteries. You would need three sets of balls numbered 0 to 9, one set for each of the digits from 000 to 999 (if we select 000 we'll call that 1000). Number the list of names from 1 to 1000 and then use the ball machine to select the three digits that selects each person. The obvious disadvantage here is that you need to get the ball machines. (Where do they make those things, anyway? Is there a ball machine industry?).

Neither of these mechanical procedures is very feasible and, with the development of inexpensive computers there is a much easier way. Here's a simple procedure that's especially useful if you have the names of the clients already on the computer. Many computer programs can generate a series of random numbers. Let's assume you can copy and paste the list of client names into a column in an EXCEL spreadsheet. Then, in the column right next to it paste the function =RAND() which is EXCEL's way of putting a random number between 0 and 1 in the cells. Then, sort both columns – the list of names and the random number – by the random numbers. This rearranges the list in random order from the lowest to the highest random number. Then, all you have to do is take the first hundred names in this sorted list. pretty simple. You could probably accomplish the whole thing in under a minute.

Simple random sampling is simple to accomplish and is easy to explain to others. Because simple random sampling is a fair way to select a sample, it is reasonable to generalize the

results from the sample back to the population. Simple random sampling is not the most statistically efficient method of sampling and you may, just because of the luck of the draw, not get good representation of subgroups in a population. To deal with these issues, we have to turn to other sampling methods.

Stratified Random Sampling

Stratified Random Sampling, also sometimes called proportional or quota random sampling, involves dividing your population into homogeneous subgroups and then taking a simple random sample in each subgroup. In more formal terms:

Objective: Divide the population into non-overlapping groups (i.e., strata) N1, N2, N3, … Ni, such that N1 + N2 + N3 + … + Ni = N. Then do a simple random sample of f = n/N in each strata.

There are several major reasons why you might prefer stratified sampling over simple random sampling. First, it assures that you will be able to represent not only the overall population, but also key subgroups of the population, especially small minority groups. If you want to be able to talk about subgroups, this may be the only way to effectively assure you'll be able to. If the subgroup is extremely small, you can use different sampling fractions (f) within the different strata to randomly over-sample the small group (although you'll then have to weight the within-group estimates using the sampling fraction whenever you want overall population estimates). When we use the same sampling fraction within strata we are conducting proportionate stratified random sampling. When we use different sampling fractions in the strata, we call this disproportionate stratified random sampling. Second, stratified random sampling will generally have more statistical precision than simple random sampling. This will only be true if the strata or groups are homogeneous. If they are, we expect that the variability within-groups is lower than the variability for the population as a whole. Stratified sampling capitalizes on that fact.

For example, let's say that the population of clients for our agency can be divided into three groups: Caucasian, African-American and Hispanic-American. Furthermore, let's assume that both the African-Americans and Hispanic-Americans are relatively small minorities of the clientele (10% and 5% respectively). If we just did a simple random sample of n=100 with a sampling fraction of 10%, we would expect by chance alone that we would only get 10 and 5 persons from each of our two smaller groups. And, by chance, we could get fewer than that! If we stratify, we can do better. First, let's determine how many people we want to have in each group. Let's say we still want to take a sample of 100 from the population of 1000 clients over the past year. But we think that in order to say anything about subgroups we will need at least 25 cases in each group. So, let's sample 50 Caucasians, 25 African-Americans, and 25 Hispanic-Americans. We know that 10% of the population, or 100 clients, are African-American. If we randomly sample 25 of these, we have a within-stratum sampling fraction of 25/100 = 25%. Similarly, we know that 5% or 50 clients are Hispanic-American. So our

within-stratum sampling fraction will be 25/50 = 50%. Finally, by subtraction we know that there are 850 Caucasian clients. Our within-stratum sampling fraction for them is 50/850 = about 5.88%. Because the groups are more homogeneous within-group than across the population as a whole, we can expect greater statistical precision (less variance). And, because we stratified, we know we will have enough cases from each group to make meaningful subgroup inferences.

Systematic Random Sampling

Here are the steps you need to follow in order to achieve a systematic random sample:

number the units in the population from 1 to N

decide on the n (sample size) that you want or need

k = N/n = the interval size

randomly select an integer between 1 to k

then take every k^{th} unit

All of this will be much clearer with an example. Let's assume that we have a population that only has N=100 people in it and that you want to take a sample of n=20. To use systematic sampling, the population must be listed in a random order. The sampling fraction would be f = 20/100 = 20%. in this case, the interval size, k, is equal to N/n = 100/20 = 5. Now, select a random integer from 1 to 5. In our example, imagine that you chose 4. Now, to select the sample, start with the 4th unit in the list and take every k-th unit (every 5th, because k=5). You would be sampling units 4, 9, 14, 19, and so on to 100 and you would wind up with 20 units in your sample.

For this to work, it is essential that the units in the population are randomly ordered, at least with respect to the characteristics you are measuring. Why would you ever want to use systematic random sampling? For one thing, it is fairly easy to do. You only have to select a single random number to start things off. It may also be more precise than simple random sampling. Finally, in some situations there is simply no easier way to do random sampling. For instance, I once had to do a study that involved sampling from all the books in a library. Once selected, I would have to go to the shelf, locate the book, and record when it last circulated. I knew that I had a fairly good sampling frame in the form of the shelf list (which is a card catalog where the entries are arranged in the order they occur on the shelf). To do a simple random sample, I could have estimated the total number of books and generated random numbers to draw the sample; but how would I find book #74,329 easily if that is the number I selected? I couldn't very well count the cards until I came to 74,329! Stratifying wouldn't solve that problem either. For instance, I could have stratified by card catalog drawer and drawn a simple random sample within each drawer. But I'd still be stuck counting cards. Instead, I did a systematic random sample. I estimated the number of books in the entire collection. Let's imagine it was 100,000. I decided that I wanted to take a sample

of 1000 for a sampling fraction of 1000/100,000 = 1%. To get the sampling interval k, I divided N/n = 100,000/1000 = 100. Then I selected a random integer between 1 and 100. Let's say I got 57.

Next, I did a little side study to determine how thick a thousand cards are in the card catalog (taking into account the varying ages of the cards). Let's say that on average I found that two cards that were separated by 100 cards were about .75 inches apart in the catalog drawer. That information gave me everything I needed to draw the sample. I counted to the 57th by hand and recorded the book information. Then, I took a compass. (Remember those from your high-school math class? They're the funny little metal instruments with a sharp pin on one end and a pencil on the other that you used to draw circles in geometry class.) Then I set the compass at .75", stuck the pin end in at the 57th card and pointed with the pencil end to the next card (approximately 100 books away). In this way, I approximated selecting the 157th, 257th, 357th, and so on. I was able to accomplish the entire selection procedure in very little time using this systematic random sampling approach. I'd probably still be there counting cards if I'd tried another random sampling method. (Okay, so I have no life. I got compensated nicely, I don't mind saying, for coming up with this scheme.)

Cluster (Area) Random Sampling

The problem with random sampling methods when we have to sample a population that's disbursed across a wide geographic region is that you will have to cover a lot of ground geographically in order to get to each of the units you sampled. Imagine taking a simple random sample of all the residents of New York State in order to conduct personal interviews. By the luck of the draw you will wind up with respondents who come from all over the state. Your interviewers are going to have a lot of traveling to do. It is for precisely this problem that cluster or area random sampling was invented.

In cluster sampling, we follow these steps:

- **divide population into clusters (usually along geographic boundaries)**
- **randomly sample clusters**
- **measure all units within sampled clusters**

For instance, in the figure we see a map of the counties in New York State. Let's say that we have to do a survey of town governments that will require us going to the towns personally. If we do a simple random sample state-wide we'll have to cover the entire state geographically. Instead, we decide to do a cluster sampling of five counties (marked in red in the figure). Once these are selected, we go to every town government in the five areas. Clearly this strategy will help us to economize on our mileage. Cluster or area sampling, then, is useful in situations like this, and is done primarily for efficiency of administration. Note also, that we probably don't have to worry about using this approach if we are conducting a mail or telephone survey because it doesn't matter as much (or cost more or raise inefficiency) where we call or send letters to.

Multi-Stage Sampling

The four methods we've covered so far – simple, stratified, systematic and cluster – are the simplest random sampling strategies. In most real applied social research, we would use sampling methods that are considerably more complex than these simple variations. The most important principle here is that we can combine the simple methods described earlier in a variety of useful ways that help us address our sampling needs in the most efficient and effective manner possible. When we combine sampling methods, we call this multi-stage sampling.

For example, consider the idea of sampling New York State residents for face-to-face interviews. Clearly we would want to do some type of cluster sampling as the first stage of the process. We might sample townships or census tracts throughout the state. But in cluster sampling we would then go on to measure everyone in the clusters we select. Even if we are sampling census tracts we may not be able to measure everyone who is in the census tract. So, we might set up a stratified sampling process within the clusters. In this case, we would have a two-stage sampling process with stratified samples within cluster samples. Or, consider the problem of sampling students in grade schools. We might begin with a national sample of school districts stratified by economics and educational level. Within selected districts, we might do a simple random sample of schools. Within schools, we might do a simple random sample of classes or grades. And, within classes, we might even do a simple random sample of students. In this case, we have three or four stages in the sampling process and we use both stratified and simple random sampling. By combining different sampling methods we are able to achieve a rich variety of probabilistic sampling methods that can be used in a wide range of social research contexts.

Nonprobability Sampling

The difference between nonprobability and probability sampling is that nonprobability sampling does not involve random selection and probability sampling does. Does that mean that nonprobability samples aren't representative of the population? Not necessarily. But it does mean that nonprobability samples cannot depend upon the rationale of probability theory. At least with a probabilistic sample, we know the odds or probability that we have represented the population well. We are able to estimate confidence intervals for the statistic. With nonprobability samples, we may or may not represent the population well, and it will often be hard for us to know how well we've done so. In general, researchers prefer probabilistic or random sampling methods over nonprobabilistic ones, and consider them to be more accurate and rigorous. However, in applied social research there may be circumstances where it is not feasible, practical or theoretically sensible to do random sampling. Here, we consider a wide range of nonprobabilistic alternatives.

We can divide nonprobability sampling methods into two broad types: accidental or purposive. Most sampling methods are purposive in nature because we usually approach the sampling problem with a specific plan in mind. The most important distinctions among these types of sampling methods are the ones between the different types of purposive sampling approaches.

Accidental, Haphazard or Convenience Sampling

One of the most common methods of sampling goes under the various titles listed here. I would include in this category the traditional "man on the street" (of course, now it's probably the "person on the street") interviews conducted frequently by television news programs to get a quick (although nonrepresentative) reading of public opinion. I would also argue that the typical use of college students in much psychological research is primarily a matter of convenience. (You don't really believe that psychologists use college students because they believe they're representative of the population at large, do you?). In clinical practice,we might use clients who are available to us as our sample. In many research contexts, we sample simply by asking for volunteers. Clearly, the problem with all of these types of samples is that we have no evidence that they are representative of the populations we're interested in generalizing to – and in many cases we would clearly suspect that they are not.

Purposive Sampling

In purposive sampling, we sample with a purpose in mind. We usually would have one or more specific predefined groups we are seeking. For instance, have you ever run into people in a mall or on the street who are carrying a clipboard and who are stopping various people and asking if they could interview them? Most likely they are conducting a purposive sample (and most likely they are engaged in market research). They might be looking for Caucasian females between 30-40 years old. They size up the people passing by and anyone who looks to be in that category they stop to ask if they will participate. One of the first things they're likely to do is verify that the respondent does in fact meet the criteria for being in the sample. Purposive sampling can be very useful for situations where you need to reach a targeted sample quickly and where sampling for proportionality is not the primary concern. With a purposive sample, you are likely to get the opinions of your target population, but you are also likely to overweight subgroups in your population that are more readily accessible.

All of the methods that follow can be considered subcategories of purposive sampling methods. We might sample for specific groups or types of people as in modal instance, expert, or quota sampling. We might sample for diversity as in heterogeneity sampling. Or, we might capitalize on informal social networks to identify specific respondents who are hard to locate otherwise, as in snowball sampling. In all of these methods we know what we want – we are sampling with a purpose.

Modal Instance Sampling

In statistics, the mode is the most frequently occurring value in a distribution. In sampling, when we do a modal instance sample, we are sampling the most frequent case, or the "typical" case. In a lot of informal public opinion polls, for instance, they interview a "typical" voter. There are a number of problems with this sampling approach. First, how do we know what the "typical" or "modal" case is? We could say that the modal voter is a person who is of average age, educational level, and income in the population. But, it's not clear that

using the averages of these is the fairest (consider the skewed distribution of income, for instance). And, how do you know that those three variables – age, education, income – are the only or even the most relevant for classifying the typical voter? What if religion or ethnicity is an important discriminator? Clearly, modal instance sampling is only sensible for informal sampling contexts.

Expert Sampling

Expert sampling involves the assembling of a sample of persons with known or demonstrable experience and expertise in some area. Often, we convene such a sample under the auspices of a "panel of experts." There are actually two reasons you might do expert sampling. First, because it would be the best way to elicit the views of persons who have specific expertise. In this case, expert sampling is essentially just a specific subcase of purposive sampling. But the other reason you might use expert sampling is to provide evidence for the validity of another sampling approach you've chosen. For instance, let's say you do modal instance sampling and are concerned that the criteria you used for defining the modal instance are subject to criticism. You might convene an expert panel consisting of persons with acknowledged experience and insight into that field or topic and ask them to examine your modal definitions and comment on their appropriateness and validity. The advantage of doing this is that you aren't out on your own trying to defend your decisions – you have some acknowledged experts to back you. The disadvantage is that even the experts can be, and often are, wrong.

Quota Sampling

In quota sampling, you select people non-randomly according to some fixed quota. There are two types of quota sampling: proportional and non-proportional. In proportional quota sampling you want to represent the major characteristics of the population by sampling a proportional amount of each. For instance, if you know the population has 40% women and 60% men, and that you want a total sample size of 100, you will continue sampling until you get those percentages and then you will stop. So, if you've already got the 40 women for your sample, but not the sixty men, you will continue to sample men but even if legitimate women respondents come along, you will not sample them because you have already "met your quota." The problem here (as in much purposive sampling) is that you have to decide the specific characteristics on which you will base the quota. Will it be by gender, age, education race, religion, etc.?

Non-proportional quota sampling is a bit less restrictive. In this method, you specify the minimum number of sampled units you want in each category. here, you're not concerned with having numbers that match the proportions in the population. Instead, you simply want to have enough to assure that you will be able to talk about even small groups in the population. This method is the non-probabilistic analogue of stratified random sampling in that it is typically used to assure that smaller groups are adequately represented in your sample.

Heterogeneity Sampling

We sample for heterogeneity when we want to include all opinions or views, and we aren't concerned about representing these views proportionately. Another term for this is sampling for diversity. In many brainstorming or nominal group processes (including concept mapping), we would use some form of heterogeneity sampling because our primary interest is in getting broad spectrum of ideas, not identifying the "average" or "modal instance" ones. In effect, what we would like to be sampling is not people, but ideas. We imagine that there is a universe of all possible ideas relevant to some topic and that we want to sample this population, not the population of people who have the ideas. Clearly, in order to get all of the ideas, and especially the "outlier" or unusual ones, we have to include a broad and diverse range of participants. Heterogeneity sampling is, in this sense, almost the opposite of modal instance sampling.

Snowball Sampling

In snowball sampling, you begin by identifying someone who meets the criteria for inclusion in your study. You then ask them to recommend others who they may know who also meet the criteria. Although this method would hardly lead to representative samples, there are times when it may be the best method available. Snowball sampling is especially useful when you are trying to reach populations that are inaccessible or hard to find. For instance, if you are studying the homeless, you are not likely to be able to find good lists of homeless people within a specific geographical area. However, if you go to that area and identify one or two, you may find that they know very well who the other homeless people in their vicinity are and how you can find them.

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF SCIENCE AND HUMANITIES

DEPARTMENT OF VISUAL COMMUNICATION

UNIT – V – Introduction to Communication Research – SVC1301

IV. Classification and Tabulation

After the collection of data is completed, it is prepared for analysis. As the data is raw, it needs to be transformed in such a way, that it is appropriate for analysis. The form of data, highly influences the result of analysis and so, to get positive results, the data preparation should be proper. There are various steps of data preparation, which include editing, coding, classification, tabulation, graphical representation and so on. • For a layperson, classification and tabulation are same, but the fact is they are different, as the former is a means to sort data, for further analysis while the latter is used to present data.

CLASSIFICATION

Classification is that the process of classifying data into groups

TABULATION

Tabulation is the act of presenting data in tabular form, for better interpretation.

DEFINITION OF CLASSIFICATION

• Classification refers to a process, wherein data is arranged based on the characteristic under consideration, into classes, or groups, as per resemblance of observations. Classification puts the data in a condensed form, as it removes unnecessary details that helps to easily comprehend data. • The data collected for the first time is raw data and so it is arranged in haphazard manner, which does not provide a clear picture. The classification of data reduces the large volume of raw data into homogeneous groups, i.e. data having common characteristics or nature are placed in one group and thus, the whole data is bifurcated into a number of groups. there are four types of classification: • Qualitative Classification or Ordinal Classification • Quantitative Classification • Chronological or Temporal Classification • Geographical or Spatial Classification DEFINITION OF TABULATION • Tabulation refers to a logical data presentation, wherein raw data is summarized and displayed in a compact form, i.e. in statistical tables. In other words, it is a systematic arrangement of data in columns and rows, that represents data in concise and attractive way. One should follow the given guidelines for tabulation. • A serial number should be allotted to the table, in addition to the self explanatory title. • The statistical table is required to be divided into four parts, i.e. Box head, Stub, Caption and Body. The complete upper part of the table that contains columns and sub-columns, along with caption, is the Box Head. The left part of the table, giving description of rows is called stub. The part of table that contains numerical figures and other content is its body. • Length and Width of the table should be perfectly balanced. • Presentation of data should be such that it takes less time and labor to make comparison between various figures. • Footnotes, explaining the source of data or any other thing, are to be presented at the bottom of the table.

Significance of Report Writing In general, reporting writing is very helpful for making the record of documentation. With the help of reports, we can easily recognize our work. For example, reports play a vital role in schools and colleges for knowing how many students have joined in this year. Report writing also helps the director of the industry, business or any

organization in order to make quick decisions and planning of anything. The importance of report writing is that it also helps to communicate within the company that is workers, to discuss the problems of the business and to give investor details of everyday running. A report can be good when it can be written in the manner of proper communication and written communication. SIGNIFICANCE OF REPORT WRITING • Easy Tool for Making Decision • Analysis • Evaluation • Quick Source • Improvement of Skill • Explain Facts • Professional Improvements • Fully Control • Easy Tool • Handle Complex Situation • Skills You Need!

EASY TOOL FOR MAKING DECISION In this modern world, huge companies need a large number of information. The only solution to manage this is to make reports. With the help of report writing, necessary judgments can be made in business. Report writing provides easy, updated and helpful details in a document. ANALYSIS Report is very important because whenever any issue occurs, a group of committees try to find the reason for that issue and provide the whole scenario and results with or without the suggestion in the shape of report writing. EVALUATION A company that is based on a large scale are involved in many different activities. Due to this reason, it is impossible for the management to keep an eye on every one of what everybody is doing. Therefore, the management finds an easy way and writes a report to highlight the acts of every department. QUICK SOURCE It is a fact that marketing managers require details in order to make a quick decision. In this case, sometimes a senior manager seems to be very busy for many reasons that's why it requires authentic sources to get information. These types of sources can be in the form of report writing. IMPROVEMENT OF SKILL Report writing helps you to improve the skills of designing, judgments, and means of communication. Therefore, these skills make a way to promote you. EXPLAIN FACTS Facts can be considered in a way that can be obtained naturally. That is why this type of presentation is verified with the help of a report because it examines and explores, calculates and many other things about any fact. PROFESSIONAL IMPROVEMENTS Report writing is very important if we talk about the advancement and improvement of the professional. For example, in a company, if anyone is selected for promotion, so before this, you have to make report writing document in which you will write about the position, performance satisfaction and level of working, with the help of all these points you can be able to make a report and submitted to the manager so that promotion can be given to that particular person fulfilling all the requirements areas. FULLY CONTROL No matter, the activities are fully achieved in accordance with the plan or not. That is why in order to control such activities it depends mainly on report writing. EASY TOOL Many of the reports make an easy way of activities for the managers. In order for planning, promoting, controlling and managing. No doubt a report writing plays a vital role to help a manager as the report refers to as a source of information. HANDLE COMPLEX SITUATION In an organization that is based on a large scale, there might always have a problem of labor which may result in complex conditions. In order to handle this situation, managers make a report. SKILLS YOU NEED! If you are looking for some extra help and want new ideas? A firm named, Best Assignment Writers gives you help in any kind of field of study that you currently need. You can submit all your work to online report writers in order to get high-quality results that your teacher will surely surprise and amazed. You must feel free to get any information regarding your assignment topic and can contact us at any time! This is quite helpful for students to get free from any problematic report and assignment writing.

TYPES OF REPORTS • Types of reports include • memos, meeting minutes, • expense reports, • audit reports, • closure reports, • progress reports, • justification reports, • compliance reports, • annual reports, • feasibility reports.