Arba Minch University

Continuing and Distance Education Center

Research Methods in Accounting and Finance Module

Prepared
by:
Mr. Taye Tadesse (Asst. Professor)

Arba Minch University
College of Business and Economics
Department of Accounting and Finance

June, 2019

Research Methods in Accounting and Finance Module (AcFn 2131)

Prepared
by:
Mr. Taye Tadesse (Asst. Prof.)

Edited
by:
Mr. TamratGetahun (MBA)

June, 2019

Acknowledgement

First and foremost, I thank the Almighty God for giving me the patience in all my accomplishments.

I am also indebted to Ato Yimam Damtie who read the manuscript and pointed out areas of confusion so that I could make concepts more understandable, and who edited that much of the module and give their constructive criticism.

Lastly, I would like to express my heartfelt gratitude to my friends being there for me in all endeavors.

Table of Contents

Preface

"Research Methods in Accounting and Finance" module is one component of the intensive courses to be taught in Accounting and Finance field of study. The module is designed with the aim of imparting the basic concepts, principles and processes of research methods. The module is organized in three broad categories: viz, knowing what research is, planning research and conducting research. The nine units are sequenced in a way that closely follows the process from inception to conclusion of a research project.

Definition of business research, reasons for studying research methods, language of research and ethical issues related to research work has been discussed briefly in the first unit.

The second unit begins with the management-research question hierarchy in the research process and continues to elaborate about designing the study through reporting the results i.e., framework of the research process.

The third unit briefly presents what research proposal mean, its components (the structure of research proposal), and the processes for evaluating the quality of proposals.

A research design is the strategy for study and the plan by which the strategy is to be carried out. And hence, the fourth unit specifies the methods and procedures for the collection, measurement and analysis of data (i.e., related to exploratory, descriptive and causal studies).

Researching secondary sources is complex and challenging. Nevertheless, the fifth unit depicts the levels and types of information sources and shows the way how to evaluate these sources. It has given a great coverage about data mining from internal sources too.

After defining qualitative research, the sixth unit clearly shows the controversy and distinction between qualitative and quantitative research. Even though it places greater emphasis to qualitative research, at last it describes the strategies for combining qualitative and quantitative methodologies.

The communication approach involves surveying or interviewing people and recording their responses for analysis. Hence, personal interviews, telephone interviews, or self-administered surveys, with each method's specific strengths and weaknesses; they have been discussed in the seventh unit.

When field studies are undertaken in a practical life, consideration of time and cost almost invariably lead to a selection of respondents i.e., selection of only a few items. In relation to this, discussion has been made about probability and non-probability sampling techniques in detail with illustrative examples.

The last unit but the crucial one talk about analysis and presentation of data. It has been presented into three parts: Data preparation and description (editing, coding, and entering of data); Exploring, displaying and examining data (statistical analysis); and Reporting research findings (how to write the research report and conduct a successful oral presentation).

Finally, the module consists of not only different activities in each topic but also review questions at the end of each unit. Please read the module before you attempt all the questions.

Reference:

Alley, Michael (2008), *The Craft of Scientific Presentations: Critical Steps to Succeed and Critical Errors to Avoid*. New York, Springer Verlag, LLC.

De Marrais, k. & Lapan, S.D. (eds.) (2004), *Foundations for Research Methods of Inquiry in Education and Social Sciences*. Mahwah: Lawrence Erlbaum Associates.

Brown Lee A., Nchinda T.C., and Yolande M.G. (1984) *How to develop Proposals and Design Research to Solve Priority Health Problems. Boston*: WHO/ Boston University Health Policy Institute

Bell J. (1993), *Doing Your Research Project. A Guide for First-Time Researchers in Education and Social Science*, Buckingham: Open University Press.

Miller F.G. (2003) *Ethical Issues in Research with Health Volunteers: Risk-Benefit Assessment* Clinical Pharmacy Therapeutics.

Unit 1: Introduction to Research in Accounting and Finance

Learning Objectives:

After completion of this unit, you should be able to:

- Identify the stimulating factors for studying research methods;
- Define what business research is;
- Distinguish between good business research and research that fails short of professional quality;
- Differentiate the terminology used by professional researchers employing scientific thinking; and
- Discuss what issues are covered in research ethics.

Introduction

Dear Learner! This unit deals with the definition of research and an attempt has been to introduce the research methods in accounting and finance with the rationales behind studying it. The languages used in research have been briefly discussed since they play a crucial role while conducting a research work. More over, the ethical issues related to business research have been also covered in this unit.

1. Research Defined

People have long strived to grips with their environmental and to understand the nature of the phenomena it presents to their senses. One of the means by which they set out to achieve these ends is research. Research is an often-misused term; its usage in everyday language is very different from the strict scientific meaning. In the field of science, it is important to move away from the looser meaning and use it only in its proper context. Scientific research adheres to set of strict protocols and long established structures.

Research is defined as human activity based on intellectual application in the investigation of matter. In other words, research is the systematic process of collecting and analyzing information to increase our understanding of the phenomenon under study. It is the function of the researcher to contribute to the understanding of the phenomenon and communicate that understanding to others. It may be said that the general aims of research are to observe and describe, to predict, to determine causes and explain.

2. Why Study Business Research?

You are about to begin your study of business research, both the process and the tools needed to reduce risk in managerial decision making. **Accounting and Finance research**, as we use the term in this module, is a systematic inquiry that provides information to guide managerial decisions. More, specifically, it is a process of planning, acquiring, analyzing, and disseminating relevant data, information, and insights to decision makers in ways that mobilize the organization to take appropriate actions that, in turn, maximize business performance. A variety of different types of research projects are grouped under the label "accounting and finance research" and we will explore them all later in this unit.

As the opening vignette and the early decision scenarios reveal, decision makers can be found in every type of organization: business, not-for-profit organizations, and public agencies. Regardless of where these decision makers are found or whether their resources are abundant or limited, they all rely on information to make more efficient and effective use of their budgets. Thus, in this module, we will take the broadest perspective of managing and its resulting application to accounting and finance research.

At no other time in our history has so much attention been placed on measuring and enhancing **return on investment** (**ROI**). At its most simplistic, when we measure ROI we calculate the financial return for all expenditures. Increasingly organizational managers want to know what strategies and tactics capture the highest return. In the last dozen years, as technology has improved our measurement and tracking capabilities, managers have realized that they need a better understanding of employee, stockholder,

and customer behavior in order to influence the desired metrics. Accounting and finance research plays an important role in this new measurement environment. Not only does it help managers choose better strategies and tactics, but accounting and finance research expenditures are increasingly scrutinized for their contribution to ROI.

The research methods course recognizes that students preparing to manage any function-regardless of the setting-need training in a disciplined process for conducting an inquiry of a **management dilemma**, the problem or opportunity that requires a management decision. Several factors should stimulate your interest in studying research methods:

1. *Explosive growth and influence of the internet*: the explosive growth of the company Web sites, e-commerce, and electronic publications brings extensive amounts of new information-but its quality and its credibility are increasingly suspect.
2. *Stakeholders demanding greater influence*: customers, workers, shareholders, and the general public demand to be included in company decision making; armed with extensive information, they are more sensitive to their own self-interests than ever before and more resistant to an organization's stimuli.
3. *More vigorous competition:* competition, both global and domestic, is growing and often coming from unexpected sources; many organizations refocus on primary competencies, while they seek to improve operations by reducing costs and converting customers to advocates.
4. *More government intervention:* government continues to show concern with all aspects of society, becoming increasingly aggressive in protecting its various publics by posing restrictions on the use of accounting and finance research tools.
5. *More complex decisions:* managers have more variables to consider in every decision, increasing the manager's need for more and better information and for greater insights from that information.
6. *Greater computing power and speed:*
    - *Lower cost of data collection*: computer and telecommunications lowered the costs of data collection, drastically changing knowledge about consumers at both store and household levels; employees at the position, team, and

department levels; suppliers and distributors at the transaction, division, and company levels; and machines at the part, process and production-run levels.

- *Better visualization tools*: high speed downloads of images allow researchers to help people visualize complex concepts, which enrich measurement capabilities.
- *Powerful computations*: sophisticated techniques of quantitative analysis are emerging to take advantage of increasingly powerful computing capabilities.
- *More integration of data*: computer advances permit businesses to create and manage a **data warehouse**, an electronic storehouse where vast arrays of collected, integrated data are ready for mining.
- *More and faster access to information*: the power and ease of the use of today's computers offer us the capability to analyze more data more quickly to deal with complex managerial problems. Yet the quantity of collected raw data overwhelms users, necessitating a means to manage it. Early efforts to provide a flow of information to managers used an accounting information system (AIS). As time passed, the challenge of database management from an AIS perspective include removing obstacles like resistance to use, reluctance of mangers to disclose fully their information needs and decision criteria, costs of single-user report generation, system design time, slow adaptation to changing organizational structures, and decision relevance (standard versus tailored reports).
- *Advanced analytical tools for enhanced insights*: organizations increasingly practice **data mining**, applying mathematical models to extract meaningful knowledge from volumes of data contained within internal databases. Enormous quantities of research data are reduced to relatively straight forward equations with statistical models. Expert systems, an outgrowth of artificial intelligence, and data mining entered the 21st century as important tools for research. Advanced analytical tools are available to answer a variety of research questions. Traditional topics open to modeling-market share, price elasticity, the cannibalization of one product's sales by the introduction of another product, the effects on productivity of changing an employee

compensation system, to name a few-create decision support models that reflect the behavior of individuals, households, and industries. Programs that combine modeling and decision support systems evolve in the later part of the 20th century to provide the most utility to users.

- *Customized reporting*: while routine AIS reports are useful for well-structured problems and those amenable to a standardized set of procedures, data must be more than timely and standardized; reporting must be customized to be truly meaningful to users.

7. *New perspectives on established research methodologies*: older tools and methodologies once limited to exploratory research are gaining wider acceptance in dealing with a broader range managerial problems.

To do well in such an environment, you will need to understand how to identify quality information and to recognize the solid, reliable accounting and finance research on which high-risk decisions as a manager can be based. You also need to know how to conduct such research. Developing these skills requires understanding the scientific method as it applies to the decision-making environment. This module addresses your needs as information collector, processor, evaluator, and user.

3. What is Good Research?

Good research generates dependable data that are derived by professional conducted practices and that can be used reliably for decision making. In contrast, poor research is carelessly planned and conducted, resulting in data that a manager can't use to reduce his/her decision-making risks. Good research follows the standards of the **scientific method**: systematic, empirically based procedures for generating replicable research.

Here are the several defining characteristics of the scientific method with the management dimensions of each.

1. *Purpose clearly defined*: the purpose of the business research-the problem involved or the decision to be made-should be clearly defined and sharply delineated in terms as unambiguous as possible. Getting this in writing is valuable

in instances where the same person serves as researcher and decision maker. The statement of the decision problem should include its scope, its limitations and the precise meanings of all words and terms significant to the research. Failure of the researcher to do this adequately may raise legitimate doubts in the minds of research report readers as to whether the researcher has sufficient understanding of the problem to make a sound proposal attacking it. This characteristic is comparable to developing a strategic plan for achieving an objective before developing a tactical plan or an action map.

2. *Research process detailed*: the research procedures used should be described in sufficient detail to permit another researcher to repeat the research. Except when secrecy is imposed, research reports should reveal with candor the sources of data and the means by which they were obtained. Omission of significant procedural details makes it difficult or impossible to estimate the validity and reliability of the data and justifiably weakens the confidence of the reader in the research itself as well as any recommendations based on the research. This characteristic is comparable to developing a tactical plan.
3. *Research design thoroughly planned*: the procedural design of the research should be carefully planned to yield results that are as objectives as possible. When a sampling of the population is involved, the report should include evidence concerning the degree of representativeness of the sample. A survey of opinions or recollections ought not to be used when more reliable evidence is available from documentary sources or by direct observation. Bibliographic searches should be a thorough and complete as possible. Experiments should have satisfactory controls. Direct observations should be recorded in writing in as soon as possible after the event. Efforts should be made to minimize the influence of personal bias in selecting and recording data. This characteristic is comparable to developing detailed action plans for each tactic.
4. *High ethical standards applied*: researchers often work independently and have significant latitude in designing and executing research projects. A research design that includes safeguards against causing mental or physical harm to participants and makes data integrity a first priority should be highly valued.

Ethical issues in research reflect important moral concerns about the practices of responsible behavior in society.

Researchers frequently find themselves precariously balancing the rights of their subjects against the scientific dictates of their chosen method. When this occurs, they have a responsibility to guard the welfare of the participants in the studies and also the organizations to which they belong, their clients, their colleagues, and themselves. Careful consideration must be given to those research situations in which there is a possibility of physical or psychological harm, exploitation, invasion of privacy, and/or loss of dignity. The research need must be weighted against the potential for adverse effects. Typically, you can redesign a study, but sometimes you can not. The researcher should be prepared for this dilemma.

5. *Limitations frankly reveled:* the researcher should report, with complete frankness, flaws in procedural design and estimate their effect on the findings. There are very few perfect research designs. Some of the imperfections may have little effect on the validity and reliability of the data; others may invalidate them entirely. A competent researcher should be sensitive to the effects of imperfect design. The researcher's experience in analyzing data should provide a basis for estimating the influence of design flaws. As a decision maker, you should question the value of research where no limitations are reported.
6. *Adequate analysis for decision maker's needs:* analysis of the data should be extensive enough to reveal its significance, what managers call "insights." The methods of analysis used should be appropriate. The extent to which this criterion is met is frequently a good measure of the competence of the researcher. Adequate analysis of the data is the most difficult phase of the research for the novice. The validity and reliability of data should be checked carefully. The data should be classified in ways that assist the researcher in reaching pertinent conclusions and clearly reveal the findings that have led to those conclusions. When statistical methods are used, the probability of error should be estimated and the criteria for statistical significance applied.

7. *Findings presented unambiguously*: some evidence of the competence and integrity of the researcher may be found in the report itself. For example, language that is restrained, clear, and precise; assertions that are carefully drawn and hedged with appropriate reservations; and an apparent effort to achieve maximum objectivity tend to leave a favorable impression of the researcher with the decision maker. Generalizations that outrun the evidence on which they are based, exaggerations, and unnecessary verbiage tend to be leave an unfavorable impression. Such reports are not valuable to managers wading through the minefields of organizational decision making. Presentation of data should be comprehensive, easily understood by the decision maker, organized so that the decision maker can readily locate critical findings.
8. *Conclusions justified*: conclusions should be limited to those for which the data provide an adequate basis. Researchers are often tempted to broaden the basis of induction by including personal experiences and their interpretations-data not subject to the controls under which the research data were gathered. Equally undesirable is the all-too-frequent practice of drawing conclusions form a study of a limited population and applying them universally. Researchers also may be tempted to rely too heavily on data collected in a prior study and use it in the interpretation of a new study. Such practice sometimes occurs among research specialists who confine their work to clients in a small industry. These actions tend to decrease the objectivity of the research and weaken reader's confidence in the findings. Good researchers always specify the conditions under which their conclusions seem to be valid.
9. *Researcher's experience reflected*: greater confidence in the research is warranted if the researcher is experienced, has a good reputation in research, and is a person of integrity. Were it possible for the reader of a research report to obtain sufficient information about the researcher, this criterion perhaps would be one of the best bases for judging the degree of confidence a piece of research warrants and the value of any decision based upon it. For this reason the research report should contain information about the qualifications of the researcher.

Good business research has an inherent value only to the extent that it helps management make better decisions that help achieve organizational goals. Interesting information about consumers, employees, competitors, or the environment might be pleasant to have, but its value is limited if the information cannot be applied to a critical decision. If a study does not help management select more effective, more efficient, less risky, or more profitable alternatives than otherwise would be the case, its use should be questioned. Alternatively, management may have insufficient resources (time, money, or skill) to conduct an appropriate study or may face a low level of risk associated with the decision at hand. In these situations, it is valid to avoid business research and its associated costs in time and money. Business research finds its justification in the contribution it makes to the decision maker's task and to the bottom line.

In the units that follow, we discuss scientific research procedures and show their application to pragmatic problems of the manager. At a minimum, our objective is to make you a more intelligent consumer of research products prepared by others as well as to enable you perform quality research for your own decisions and those of others to whom you report.

Activity 1

Dear learner, based on the above discussions

1. Define business research.
2. Write the characteristics of good business research.
3. Write the reasons why we study business research.

4. Research and the Scientific Method

Good business research is based on sound reasoning. Competent researchers and smart managers alike practice thinking habits that reflect sound reasoning-finding correct premises, testing the connections between their facts and assumptions, making claims based on adequate evidence.

The scientific method, as practiced in business research, guides our approach to problem solving. The essential tenets of the scientific method are:

- Direct observation of phenomena.
- Clearly defined variables, methods, and procedures.
- Empirically testable hypotheses.
- The ability to rule out rival hypotheses.
- Statistical rather than linguistic justifications of conclusions.
- The self-correcting process.

The scientific method and scientific inquiry generally, is described as a puzzle-solving activity. For the researcher, puzzles are solvable problems that may be clarified or resolved through reasoning processes. The steps that follow represent one approach to assessing the validity of conclusions about observable events. They are particularly appropriate for accounting and finance researchers whose conclusions result from empirical data. The researcher:

1. Encounters a curiosity, doubt, barrier, suspicion, or obstacle.
2. Struggles to state the problem-asks questions, contemplates existing knowledge, gathers facts, and moves from an emotional to an intellectual confrontation with the problem.
3. Proposes a hypothesis, a plausible explanation, to explain the facts that are believed to be logically related to the problem.
4. Deduces outcomes or consequences of the hypothesis-attempts to discover what happens if the results are in the opposite direction of that predicted or if the results support the expectations.
5. Formulates several rival hypotheses.
6. Devices and conducts a crucial empirical test with various possible outcomes, each of which selectively excludes one or more hypotheses.
7. Draws a conclusion (an inductive inference) based on acceptance or rejection of the hypotheses.
8. Feeds information back into the original problem, modifying it according to the strength of the evidence.

Clearly, reasoning is pivotal to much of the researcher's success: gathering facts consistent with the problem, proposing and eliminating rival hypotheses, deducing outcomes, developing crucial empirical tests, and deriving the conclusion.

Researchers think of the doing of science as an orderly process that combines induction, deduction, observation, and hypothesis testing into a set of reflective thinking activities. Although the scientific method consists of neither sequential nor independent stages, the problem-solving process that it reveals provides insight into the way research is conducted.

5. The Language of Research

When we do research, we seek to know what is in order to understand, explain, and predict phenomena. We might want to answer the question "What will be the department's reaction to the new flexible work schedule?" or "Why did the stock market price surge higher than when all normal indicators suggested it would go down?" When dealing with such questions, we must agree on definitions. Which member of the department: clerical or professional? What kind of reaction? What are normal indicators? These questions require the use of concepts, constructs, definitions, variables, propositions, theory and model.

5.1. Concepts

To understand and communicate information about objects and events, there must be a common ground on which to do it. Concepts serve this purpose. A **concept** is a generally accepted collection of meanings or characteristics associated with certain events, objects, conditions, situations and behaviors. Classifying and categorizing objects or events that have a common characteristic beyond any single observation creates concepts. When you think of a spreadsheet or a warranty card, what comes to mind is not a single example but your collected memories of all spreadsheets and warranty cards, from which you abstract a set of specific and definable characteristics.

We abstract such meanings from our experiences and use words as labels to designate them. For example, we see a man passing and identifying that he is running, walking, skipping, crawling, or hopping. These movements all represent concepts. We also have abstracted certain visual elements by which we identify that the moving object is an adult male, rather than an adult female or a truck or a horse. We use numerous concepts daily in our thinking, conversing, and other activities.

Sources of Concepts

Concepts that are infrequent and general have been developed over time through shared language usage. We acquire them through personal experience. Ordinary concepts make up the bulk of communication even in research, but we often run into difficulty trying to deal with an uncommon concept or a newly advanced idea. One way to handle this problem is to borrow from other languages (for example, *gestalt*) or to borrow from other fields (for example, from art, *impressionism*). The concept of gravitation is borrowed from physics and used in marketing in an attempt to explain why people shop where they do. The concept of distance is used in attitude measurement to describe degree of variability between the attitudes of two or more persons. Threshold is used effectively to describe a concept about the way we perceive. Sometimes we need to adopt new meanings for words (make a word cover different concept) or develop new labels for concepts. The recent broadening of the meaning of *model* is an example of the first instance; the development of concepts such as *sibling* and *status-stress* are examples of the second. When we adopt new meanings or develop new labels, we begin to develop a specialized jargon or terminology. Jargon no doubt contributes to efficiency of communication among specialists, but it excludes everyone else.

Importance of Concept to Research

In research, special problems grow out of the need for concept precision and inventiveness. We design hypotheses using concepts. We devise measurement concepts by which to test these hypothetical statements. We gather data using these measurement concepts. The success of research hinges on (1) how clearly we conceptualize and (2) how well others understand the concepts we use. Foe example, when we survey people

on the question of customer loyalty, the questions we used need to tap faithfully the attitudes of the participants. Attitudes are abstract, yet we must attempt to measure them using carefully selected concepts.

The challenge is to develop concepts that others will clearly understand. We might for example, ask participants for an estimate of their family's total income. This may seem to be a simple, unambiguous concept, but we will receive varying and confusing answers unless we restrict or narrow the concepts by specifying:

- Time period, such as weekly, monthly, or annually.
- Before or after income taxes.
- For head of family only or for all family members.
- For salary and wages only or also for dividends, interest, and capital gains.
- Income in kind, such as free rent, employee discounts, or food stamps.

5.2. Constructs

Concepts have progressive levels of abstraction-that is, the degree to which the concept does or does not have something objective to refer to. *Table* is an objective concept. We can point to a table, and we have images of the characteristics of all tables in our mind. An abstraction like *personality* is much more difficult to visualize. Such abstract concepts are often called constructs. A **construct** is an image or abstract idea specifically invented for a given research and/or theory-building purpose. We build constructs by combining the simpler, more concrete concepts, especially when the idea or image we intend to convey is not subject to direct observation.

5.3. Definitions

Confusion about the meaning of concepts can destroy a research study's value without the researcher or client even knowing it. If words have different meanings to the parties involved, then the parties are not communicating well. Definitions are one way to reduce this danger.

Researchers struggle with two types of definitions: dictionary definitions and operational definitions. In the more familiar dictionary definition, a concept is defined with a synonym. For example, a customer is defined as a patron; a patron, in turn, is defined as a customer or a client of an establishment; a client is defined as one who employs the services of any professional and, loosely, as a patron of any shop. Circular definitions may be adequate for general communication but not for research. In research, we measure concepts and constructs, and this requires more rigorous definitions.

Operational Definitions

An operational definition is a definition stated in terms of specific criteria for testing or measurement. These terms must refer to empirical standards (that is, we must be able to count, measure, or in some other way gather the information through our senses). Whether the object to be defined is physical (e.g., a can of soup) or highly abstract (e.g., achievement motivation), the definition must specify the characteristics and how they are to be observed. The specification and procedures must be so clear that any competent person using them would classify the object in the same way.

Suppose college undergraduates are classified by class. No one has much trouble understanding such terms as *freshman*, *sophomore*, and so forth. But the task may not be that much simple if you must determine which students fall in each class. To do this, you need operational definitions.

Operational definitions may vary, depending on your purpose and the way you choose to measure them. Here are two different situations requiring different definitions of the same concepts:

1. You conduct a survey among students and wish to classify their answers by their class levels. You merely ask them to report their class status, and you record it. In this case, class is freshman, sophomore, junior, or senior, and you accept the answer each respondent gives as correct. This is a rather casual definition process but nonetheless an operational definition. It is probability adequate even though some of the respondents report inaccurately.

2. You make a tabulation of the class level of students from the university registrar's annual report. The measurement task here is more critical, so your operational definition needs to be more precise. You decide to define class levels in terms of semester hours of credit completed by the end of the spring semester and recorded in each student's record in the registrar's office:

Freshman	Fewer than 30 hours' credit
Sophomore	30 to 59 hours' credit
Junior	60 to 89 hours' credit
Senior	90 or more hours' credit

Those examples deal with relatively concrete concepts, but operational definitions are even more critical for treating abstract ideas. Suppose one tries to measure a construct called "consumer socialization." We may intuitively understand what this means, but to attempt to measure it among consumers is difficult. We would probability develop questions on skills, knowledge, and attitudes, or we may use a scale that has already been developed and validated by someone else. This scale then operationally defines the construct.

Whether you use a definitional or operational definition, its purpose in research is basically the same- to provide an understanding and measurement of concepts. We may need to provide operational definitions for only a few critical concepts, but these will almost always be the definitions used to develop the relationships found in hypotheses and theories.

5.4. Variables

In practice, the term **variable** is used as a synonym to *construct* or the property being studied. In this context, a variable is a symbol of an event, act, characteristic, trait or attribute that can be measured and to which we assign categorical values.

For purposes of data entry and analysis, we assign numerical value to a variable based on the variable's properties. For example, some variables, said to be *dichotomous*, have only two values, reflecting the presence or absence of a property: employed-unemployed or

male-female have two values, generally 0 and 1. Variables also take on values representing added categories, such as the demographic variables of race or religion. All such variables that produce data that fit into categories are said to be discrete, since only certain values are possible. An automobile variable, for example, where "Chevrolet" is assigned a 5 and "Honda" is assigned a 6, provides no option for a 5.5.

Income, temperature, age, or a test score are examples of *continuous* variables. These variables may take the values within a given range or, in some case, an infinite set. Your test score may range from 0 to 100, your age may be 23.5, and your present income could be Birr 7,000.

Independent and Dependent Variables

Researchers are most interested in relationships among variables. For a example, does a newspaper coupon (independent variable) influence product purchase (dependent variable), or can a salesperson's ethical standards influence her ability to maintain customer relationships?

Many text books use the term *predictor variable* as a synonym for **independent variable** (**IV**). This variable is manipulated by the researcher, and the manipulation causes an effect on the dependent variable. We recognize that there are often several independent variables and that they are probably at least somewhat "correlated" and therefore not independent among themselves. Similarly, the term *criterion variable* is used synonymously with **dependent variable** (**DV**). This variable is measured, predicted, or otherwise monitored and is expected to be affected by manipulation of an independent variable.

Moderating Variables

In each relationship, there is at least one independent variable and a dependent variable. It is normally hypothesized that in some way the IV "causes" the DV to occur. For simple relationships, all other variables are considered extraneous and are ignored.

In actual study situations-for example, in a typical sales office- a simple one-on-one relationship needs to be revised to take over variables into account. Often we use another type of explanatory variable in here-the moderating variable. A **moderating variable** (**MV**) is a second independent variable that is included because it is believed to have a significant contributory or contingent effect on the originally stated IV_DV relationship. For example, one might hypothesize:

> The switch to commission from a salary compensation system (IV) will lead to increased sales productivity (DV) per worker, especially among younger workers (MV).

In this case, there is differential pattern of relationship between the compensation system and productivity that is the result of age differences among the workers.

Whether a given variable is treated as an independent or as a moderating variable depends on the hypothesis.

Extraneous Variables

An almost infinite number of **extraneous variables** (**EVs**) exist that might conceivably affect a given relationship. Some can be treated as independent or moderating variables, but most must be either assumed or excluded from the study. Fortunately, these variables have little or no effect on a given situation. Most can be safely ignored. Others may be important, but their impact occurs in such a random fashion as to have little effect. However, there may be other extraneous variables to consider as possible confounding variables to our hypothesized IV-DV relationship. One might think that the *type of customers* would have an effect on a compensation system's impact on sales productivity. This might lead to our introducing a control variable as follows:

> With new customers (EV-control), a switch to commission from a salary compensation system (IV) will lead to increased sales productivity (DV) per worker, especially among younger workers (MV).

In our salesperson compensation example, we would attempt to control for type of customers by studying the effect of the switch in compensation within groups having different types of customers (new versus established).

Intervening Variables

The variables mentioned with regard to casual relationships are concrete and clearly measurable; they can be seen, counted, or observed in some way. Sometimes, however, one may not be completely satisfied by the explanations they offer. An intervening variable is a conceptual mechanism through which the IV and MV might affect the DV. The intervening variable (IVV) may be defined as "that factor which theoretically affects the observed phenomenon but cannot be seen, measured, or manipulated; its effect must be inferred from the effects of the independent and moderate variables on the observed phenomenon."

In the case of the compensation hypothesis, one might view the intervening variable to the job satisfaction, giving a hypothesis such as:

> The switch to commission compensation system (IV) will lead to higher sales productivity (DV) by increasing overall compensation (IVV).

5.5. Propositions and Hypotheses

We define a **proposition** as a statement about observable phenomena (concepts) that may be judged as true or false. When a proposition is formulated for empirical testing, we call it a **hypothesis**. As a declarative statement about the relationship between two or more variables, a hypothesis is of a tentative and conjectural nature.

Hypotheses have also been described as statements in which we assign variables to cases. A **case** is defined in this sense as the entity or thing the hypothesis talks about. The variable is the characteristic, trait, or attributes that, in the hypothesis, is imputed to the case.

> Brand Manager Ephrem (case) has a higher-than-average achievement motivation (variable).

If our hypothesis was based on more than one case, it would be a generalization. For example:

> Brand managers in company F (case) have a higher-than-average achievement motivation (variable).

Descriptive Hypotheses

Both of the above hypotheses are examples of descriptive hypotheses. They state the existence, size, form, or distribution of some variable. Researchers often use a research question rather than a descriptive hypothesis. For example:

Descriptive Hypothesis Format	*Research Question Format*
Ethiopian cities (cases) are experiencing budget difficulties (variable).	Are Ethiopian cities experiencing budget difficulties?
Eight percent of Company F stockholders (cases) favor increasing the company's cash dividend (variable).	Do stockholders of Company F favor an increased cash dividend?

Either format is acceptable, but the descriptive hypothesis format has several advantages:

- It encourages researchers to crystallize their thinking about the likely relationships to be found.
- It encourages them to think about the implications of a supported or rejected finding.
- It is useful for testing statistical significance.

Relational Hypotheses

The research question format is less frequently used with a situation calling for **relational hypotheses**. These are statements that describe a relationship between two variables with respect to some case. For example, "Foreign (variable) clothes are perceived by Ethiopian consumers (case) to be a better quality (variable) than domestic clothes." In this instance, the nature of the relationship between the two variables ("country of origin" and "perceived quality") is not specified. Is there only an implication that the variable occur in some predictable relationship, or is one variable somehow responsible for the other? The first interpretation (unspecified relationship) indicates a correlational relationship; the second (predictable relationship) indicates an explanatory, or casual, relationship.

Correlational hypotheses state that the variables occur together in some specified manner without implying that one causes the other. Such weak claims are often made

when we believe that there are more basic casual forces that affect both variables or when we have not developed enough evidence to claim a stronger linkage. Here are two sample correlational hypotheses:

Young women (under 35 years of age) purchase fewer units of our product than women who are 35 years of age or older.

The number of suits sold varies directly with the level of the business cycle.

By labeling these as correlational hypotheses, we make no claim that one variable causes the other to change or take on different values.

With explanatory (casual) hypotheses, there is an implication that the existence of or a change in one variable causes or leads to a change inn the other variable. as we noted previously, the causal variable is typically called the independent variable (IV) and the other the dependent variable (DV). Cause means roughly to "help make happen." So the IV need not be the reason for the existence of or change in the DV. Here are three examples of explanatory hypotheses:

An increase in family income (IV) leads to an increase in the percentage of income saved (DV).

Exposure to the company's messages concerning industry problems (IV) leads to more favorable attitudes (DV) by employees toward the company.

Loyalty to a particular grocery store (IV) increases the probability of purchasing the private brands (DV) sponsored by that store.

In proposing or interpreting causal hypotheses, the researcher must consider the direction influence. In many cases, the direction is obvious from the nature of the variables. Thus, one would assume that family income influence savings rate rather than the reverse. Sometimes our ability to identify the direction influence depends on the research design. In the worker attitude hypothesis, if the exposure to the message clearly precedes the attitude management, then the direction of exposure to attitude seems clear. If information about both exposure and attitude was collected at the same time, the researcher might be justified in saying that different attitudes led to selective message perception or non-perception. Store loyalty and purchasing of store brands appear to be

interdependent. Loyalty to a store may increase the probability of buying the store's private brands, but satisfaction with the store's private brand may also lead to greater store loyalty.

The Role of the Hypothesis

In research, a hypothesis serves several important functions:

- It guides the direction of the study.
- It identifies facts that are relevant and those that are not.
- It suggests which form of research design is likely to be most appropriate,
- It provides a frame work for organizing the conclusions that result.

Unless the researcher curbs the urge to include additional elements, a study can be diluted by trivial concerns that do not answer the basic questions posed by the management dilemma. The virtue of the hypothesis is that, if taken seriously, it limits what shall be studied and what shall not. To consider specifically the role of the hypothesis in determining the direction of the research, suppose we use this:

Husbands and wives agree in their perceptions of their roles in purchase decisions.

The hypothesis specifies who shall be studied (married couples), in what context they shall be studied (their consumer decision making), and what shall be studied (their individual perceptions of their roles).

The nature of this hypothesis and the implications of the statement suggest that the best research design is a communication-based study, probably a survey or interview. We have at this time no other practical means to ascertain perceptions of people except to ask about them in one way or another. In addition, we are interested only in the roles that are assumed in the purchase or consumer decision-making situation. The study should not, therefore, involve itself in seeking information about other types of roles husbands and wives might play. Reflection upon this hypothesis might also reveal that husbands and wives disagree on their perceptions of roles, but the differences may be explained in terms of additional variables, such as age, social class, background, personality and other factors not associated with their difference in gender.

What is a Strong Hypothesis? A strong hypothesis should fulfill three conditions:

- Adequate for its purpose.
- Testable.
- Better than its rivals.

5.6. Theory

Hypotheses play an important role in the development of theory. How theory differs from hypothesis may cause confusion. We make the general distinction that the difference between theory and hypothesis is one of degree of complexity and abstraction. In general theories tend to be complex, be abstract, and involve multiple variables. Hypotheses, on the other hand, tend to be more simple, limited-variable statements involving concrete instances.

A person not familiar with research uses the term *theory* to express the opposite of *fact*. In this sense, theory is viewed as being speculative or "ivory tower." One hears that mangers need to be less theoretical or that some idea will not work because it is too theoretical. This is an incorrect picture of the relationship between fact and theory to the researcher. In truth, fact and theory are each necessary for the other to be of value. Our ability to make rational decisions, as well as to develop scientific knowledge, is measured by the degree to which we combine fact and theory. We all operate on the basis of theories we hold. In one sense, theories are the generalizations we make about variables and the relationships among them. We use these generalizations to make decisions and predict outcomes. For example, it is midday and you note that the outside natural light is dimming, dark clouds are moving rapidly in from the west, the breeze is freshening, and the air temperature is cooling. Would your understanding of the relationship among these variables (your weather theory) lead you to predict that something decidedly wet will probably occur in short time?

A **theory** is a set of systematically interrelated concepts, definitions, and propositions that are advanced to explain and predict phenomena (facts). In this sense, we have many

theories and use them continually to explain or predict what goes on around us to the degree that our theories are sound and fit the situation, we successful in our explanations and predictions.

In marketing, the product life cycle describes the stages that a product category goes through in the marketplace. The generalized product life cycle has four stages (although the length and sharp of product life cycle differ): introduction, growth, maturity, and decline. In each stage, many concepts, constructs, and hypotheses describe the influences that change revenue and profit. Definitions are also needed for communicating about the claims of the theory and its consistency in testing to reality.

For example, in the growth stage, companies spend heavily on advertising and promotion to create product awareness. In the early period of this stage these expenditures may be made to fuel *primary demand* (construct), improving product class awareness rather than brand awareness. Also, high pricing may reflect *skimming* (concept) to help the company to help the company recover developmental costs. The product manager may alternatively use low pricing, or *penetration pricing* (concept), to build unit volume. In the growth stage sales increase rapidly because many consumers are trying or actually using the product; and those who tried, were satisfied, and bought again-*repeat purchasers* (concept)-are swelling the ranks. If the company is unable to attract repeat purchasers, this usually means death for the product (proposition). The maturity stage is a good time for the company in terms of generating cash (proposition). The costs of developing the product and establishing its position in the marketplace are paid and it tends to be profitable. Firms will often try to use *extension strategies* (constructs). These are attempts to delay the decline stage, "products will consume a disproportionate share of management time and financial resources relative to their potential future worth" (hypothesis). To make this hypothesis fully testable, we would need operational definitions for disproportionate share, time resources, and future worth.

The challenge for the researcher in this example is to build more comprehensive theories to explain and predict how modifying the product and other variables will benefit the firm.

5.7. Models

The term model is used in accounting and finance research and other fields of business to represent phenomena through the use of analogy. A **model** is defined here as a representation of a system that is constructed to study some aspect of that system or the system as a whole. Models differ from theories in that a theory's role is explanation whereas a model's role is representation.

Early models (and even those created as recently as the 1990s for mainframe computers) were enormously expensive and often incomprehensible to all but their developers. Modeling software, such as Excel, has made modeling more inexpensive and accessible. Models allow researchers and mangers to characterize present or future conditions: the effect of advertising on consumer awareness or intention to purchase, a product distribution channel, brand switching behavior, an employee training program, and many aspects of business. A model's purpose is to increase our understanding, prediction, and control of the complexities of the environment.

Descriptive, predictive, and normative models are found in accounting and finance research. *Descriptive* models are used frequently for more complex systems. They allow visualization of numerous variables and relationships. *Predictive* models forecast future events. *Normative* models are used chiefly for control, informing us about what actions should be taken. Models may also be static, representing a system at one point in time, or dynamic, representing the evolution of a system over time.

Models are developed through the use of inductive and deductive reasoning, which we suggested previously, is integral to accurate conclusions about business decisions. A model may originate from empirical observations about behavior based on researched facts and relationships among variables. Inductive reasoning allows the modeler to draw

conclusions from the facts or evidence in planning the dynamics of the model. The modeler may also use existing theory, managerial experience, judgment, or facts deducted from known laws of nature. In this case, deductive reasoning serves to create particular conclusions derived from general premises.

Models are an important means of advancing theories and aiding decision makers. Because the inputs are often unknown, imprecise, or temporal estimates of complex variables, creating and using models in the decision-making process can be a time-consuming endeavor.

Activity 2

1. Distinguish between
    A. Dependent and independent variables.
    B. Moderating, extraneous and intervening variables
2. Differentiate descriptive hypotheses from relational hypotheses.
3. What are the three conditions in which a strong hypothesis should fulfill?

6. What are Research Ethics?

As in other aspects of business, all parties in research should exhibit ethical behavior. Ethics are norms or standards of behavior that guide moral choices about our behavior and our relationships with others. The goal of ethics in research is to ensure that no one is harmed or suffers adverse consequences from research activities. This objective is usually achieved. However, unethical activities are pervasive and include violating nondisclosure agreements, breaking participant confidentiality, misrepresenting results, deceiving people, invoicing irregularities, avoiding legal liability, and more.

The reorganization of ethics as a problem for economic organizations was revealed in a survey where 80 percent of the responding organizations reported the adoption of an ethical code. Surprisingly, the evidence that this effort has improved ethical practices is

questionable. The same study reports limited success for codes of conduct that attempt to restrain improper behavior.

There is no single approach to ethics. Advocating strict adherence to a set of laws is difficult because of the unforeseen constraint put on researchers. Because of Germany's war history, for example, the government forbids many types of medical research. Consequently, the German people do not benefit from many advances in biotechnology and may have restricted access to genetically altered drugs in the future. Alternatively, relying on each individual's personal sense of morality is equally problematic. Consider the clash between those who believe death is deliverance from a life of suffering and those who value life to the point of preserving it indefinitely through mechanical means. Each value system claims superior knowledge of moral correctness.

Clearly, a middle ground between being completely code-governed or relying on ethical relativism is necessary. The foundation for that middle ground is an emerging consensus on ethical standards for researchers. Codes and regulations guide researchers and sponsors. Review boards and peer groups help researchers examine their research proposals for ethical dilemmas. Many design-based ethical problems can be eliminated by carefully planning and constant vigilance. In the end, responsible research anticipates ethical dilemmas and attempts to adjust the design, procedures, and protocols during the planning process rather than treating them as an afterthought ethical research requires personal integrity from the researcher, the project manager, and the research sponsor.

As research is designed, several ethical considerations must be balance:

- Protect the rights of the participant or subject.
- Ensure that the sponsor receives ethically conducted and reported research.
- Follow ethical standards when designing research.
- Protect the safety of the researcher and the team.
- Ensure the research team follows the design.

In general, research must be designed so that a participant does not suffer physical harm, discomfort, pain, embarrassment, or loss of privacy. Begin data collection by explaining

to participants the benefits expected from the research. Explain that their rights and well-being will be adequately protected and say how that will be done. Be certain that interviewers obtain the informed consent of the participant. The use of deception is questionable; when it is used, debrief any participant who has been deceived.

Many sponsors wish to undertake research without veiling themselves. Sponsors have the right to demand and receive confidentiality between themselves and researchers. Ethical researchers provide sponsors with the research design needed to solve the managerial question. The ethical researcher shows the data objectively, despite the sponsor's preferred outcomes.

The research team's safety is the responsibility of the researcher. Researchers should require ethical compliance from team members in following the research design, just as sponsors expect ethical behavior from the researcher.

Many corporations and research firms have adopted codes of ethics. Several professional associations have detailed research provisions. Of interest are the American Association of Public Opinion Research, the American Institute of Certified Public Accountants, the American Marketing Association, the American Political Science Association, and the American Psychological Association. Federal, state, and local governments have laws, policies and procedures in place to regulate research on human beings.

Review Questions

1. What is business research?
2. Why should there be any question about the definition of research?
3. Distinguish among the following sets of items, and suggest the significance of each in a research context:
    a) Concept and Construct
    b) Operational definition and dictionary definition
    c) Concept and variable
    d) Hypothesis and proposition
    e) Theory and model
4. Describe the characteristics of the scientific research method.
5. An automobile manufacturer observes the demand for its brand increasing as per capita income increases. Sales increases also follow low interest rates, which ease credit conditions. Buyer's purchasing behavior is seen to be dependent on age and gender. Other factors influencing sales appear to fluctuate almost randomly (competitor advertising, competitor dealer discounts, introductions of new competitive models).
    - If sales and per capita income are positively related, classify all variables as dependent, independent, moderating, extraneous or intervening.

.

Unit 2: The Research Process

Learning Objectives:

After completion of this unit, you should be able to:

- Describe research is decision-and dilemma-centered;
- Distinguish that the research question is the result of careful exploration and analysis;
- Explain all the stages in the research process; and
- Develop a research design.

1. The Research Process: An Introduction

Writers usually treat the research task as a sequential process involving several clearly defined steps. No one claims that research requires completion of each step before going to the next. Recycling, circumventing, and skipping occur. Some steps are begun out of sequence, some are carried out simultaneously, and some may be omitted. Despite these variations, the idea of a sequence is useful for developing a project and for keeping the project orderly as it unfolds.

The research process begins much as the vignette suggests. A management dilemma triggers the need for a decision. In other situations, a controversy arises, a major commitment of resources is called for, or conditions in the environment signal the need for a decision. Such events cause mangers to reconsider their purposes or objectives, define a problem for solution, or develop strategies for solutions they have identified.

In our view of the research process, the management question-its origin, selection, statement, exploration and refinement-is the critical activity in the sequence. Whether the researcher is involved in basic or applied research, a thorough understanding of the management question is fundamental to success in the research enterprise.

2. The Management-Research question Hierarchy

A useful way to approach the research process is to state the basic dilemma that prompts the research and then try to develop other questions by progressively breaking down the original question into more specific ones. You can think of the outcome of this process as the management-research question hierarchy.

Identifying management dilemmas is rarely difficult (unless the organization fails to track its performance factors-like sales, profits, employee turnover, manufacturing output and defects, on-time deliveries, customer satisfaction etc.). However, choosing one dilemma on which to focus may be difficult. Choosing incorrectly will direct valuable resources (time, manpower, money, and equipment) on a path that may not provide critical decision-making information (the purpose of good research). The choice is like learning to balance a pencil on its point on your finger, a coin on its edge, or a pyramid on its pinnacle: As a manager, only practice makes you proficient. For new managers, or established managers facing new responsibilities, developing several management-research question hierarchies, each staring with a different dilemma, will assist in the choice process.

2.1. The Management Question

The management must move from the **management dilemma** to the management question to proceed with the research process. The management question restates the dilemma in question form:

- What should be done to reduce employee turnover?
- What should be done to increase tenant residency and reduce move-outs?
- What should be done to reduce costs?

Management Question Categories

Management questions are too numerous to list, but we can categorize them:

- Choice of purpose or objectives.
- Generation and evaluation of solutions.
- Troubleshooting or control situation.

The first type concerns the *choice of purposes* or *objectives*. The general question is, "What do we want to achieve?" At the company level the question might be, "Should we at XYZ Corporation reconsider our basic corporate objectives as they concern our public image?" More narrowly, a management question on objectives might ask, "What goals should XYZ try to achieve in its next round of labor negotiations?"

A second category of management questions concerns the *generation and evolution of solutions*. The general question is, "How can we achieve the ends we seek?" research projects in this group usually deal with concrete problems that managers quickly recognize as useful. Projects can involve questions such as:

- "How can we achieve or five-year goal of doubled sales and net profits?"
- "What should be done to reduce post purchase service compliant?

A third class of management questions concerns *troubleshooting* or *control situation.* The problem usually involves monitoring or diagnosing various ways in which an organization is failing to achieve its established goals. This group includes questions such as, "Why does our department incur the highest costs?" and "How well is our program making its goals?"

No matter how the management question is defined, many research directions can be taken. A specific question can lead to many studies. Concern for MetalWork's company image might lead to:

- A survey among various groups to discover their attitudes towards the company.
- Secondary research into what other companies are doing to polish their images.
- A study to forecast expected changes in social attitudes.

The question concerning MetalWork's labor negotiation objective might prompt research into recent settlements in the industry or a survey among workers to find out how well management has met its concerns about the quality of work life. It is the joint responsibility of the researcher and the manager to choose the most productive project.

The Nature of the Management Question

Assume, for example, a researcher is asked to help the new management of a bank. The president is concerned about erosion of the bank's profitability (the management dilemma) and wants turn this situation around. Commercial Bank of Ethiopia is the oldest and the largest of three banks in a city with a population of about 50,000. Profits have stagnated in recent years. The president and the consultant discuss the problem facing the organization and settle on this management question: "How can we improve our profit picture?"

The management question does not specify what kind of research is to be done. This question is strictly the manager's in thrust. It implies that the bank's management faces the task of developing a strategy for increasing profits. The question is broad. Notice that it doesn't indicate whether management should increase profits via increased deposits, downsizing of personnel, outsourcing of the payroll function, or some other means.

Further discussion between the bank president and the researcher shows there are really two questions to be answered. The problem of low deposit growth is liked to concerns of a competitive nature. While lowered deposits directly affect profits, another part of the profit weakness is associated with negative factors within the organization that are increasing costs of operation. The qualified researcher knows that the management question as stated is too broad to guide a definitive research project. As a starting point, the broadly worded question is fine, but CBE will want to refine its management question into these more specific sub questions:

- "How can we increase deposits?"
- "How can we reduce costs?"

This separation of the management question into two sub questions may not have occurred without a discussion between the researcher and the manager.

2.2. Exploration

CBE has done no formal research in the past. It has little specific information about competitors or customers and has not analyzed its internal operations. To move forward

in the management-research question hierarchy and define the research question, the client needs to collect some exploratory information on:

- What factors are contributing to the bank's failure to achieve a stronger growth rates in deposit?
- How well is the bank doing regarding work climate, efficiency operations compared to industry norms, and financial condition compared to industry norms and competitors?

A small focus group is conducted among employees, and trade association data are acquired to compare financial and operating statistics from company annual reports and end-of-year division reports. From the results of these two exploratory activities, it is obvious that CBE's operations are not as progressive as its competitors' but it has its costs well in line. So the revised management question becomes, "What should be done to make the bank more competitive?"

The process of exploration typically begins with a search of published data. In addition, researchers often seek out people who are well informed on the topic, especially those who have clearly stated positions on controversial aspects of the problem.

An unstructured exploration allows the researcher to develop and revise the management question and determine what is needed to secure answers to the proposed question.

2.3. The Research Question

Once the researcher has a clear statement of the management question, s/he and the manager must translate it into a research question. Consider the research question to be a fact-oriented, information-gathering question. There are many different ways to address most management dilemmas. It is at this point of formulating research questions where the insight and the expertise of the manger come into play. Only reasonable alternatives should be considered. If the research is not part of the manager's decision-making environment, the researcher can be of minimal help in this translation. The manager's direction to the researcher is most important. If, however, the researcher is an integral part of the decision-making environment, s/he may assist the manager in evaluating which courses of action should be researched.

Incorrectly defining the research question is a fundamental weakness in the research process. Time and money can be wasted studying an alternative that won't help the manger rectify the dilemma. The researcher's task is to assist the manager in formulating a research question that fits the need to resolve the management dilemma. A research question is the hypothesis of choice that best states the objective of the research study. It is a more specific management question that must be answered. It may be more than one question, or just one. A research process that answers this more specific question provides the manager with the information necessary to make the decision he or she is facing.

Meanwhile at Commercial Bank of Ethiopia, the president has agreed to have the research be guided by the following research question: "Should CBE position itself as a modern, progressive institution (with appropriate changes in services and polices) or maintain its image as the oldest, most reliable institution in town?"\

Fine-Tuning the Research Question

The term *fine-tuning* may seem to be an odd usage for research, but it creates an image that most researchers come to recognize. Fine-tuning the question is precisely what a skillful practitioner must do after the exploration is complete. At this point, a clearer picture of the management and research questions begins to emerge. After a preliminary review of the literature, a brief exploratory study, or both, the project begins to crystallize in one of two ways:

1. It is apparent the question has been answered and the process is finished.
2. A question different from the one originally addressed has appeared.

The research question does not have to be materially different, but it will have evolved in some fashion. This is not cause for discouragement. The refined research question (s) will have better focus and will move the research forward with more clarity than the initial formulated question (s).

In addition to fine-tuning the original question, other research question-related activities should be addressed in this phase to enhance the direction of the project:

1. Examine the concepts and constructs to be used in the study. Are they satisfactorily defined? Have operational definitions been employed where appropriate?
2. Review the research questions with the intent of breaking them down into specific second-and-third level questions.
3. If hypotheses are used, be certain they meet the quality tests mentioned in the preceding chapter.
4. Determine what evidence must be collected to answer the various questions and hypotheses.
5. Set the scope of the study by stating what not a part of the research question is. This will establish boundary to separate contiguous problems from the primary objective.

When the characteristics or plausible causes of the problem are well defined and the research question is clearly stated, it is possible to deduce the essential sub questions that will guide the project planning at this stage of the research process. However, if the research question is somewhat or very poorly defined, the researcher will need further exploration and question revision to refine the original question and generate the martial for constructing investigative questions.

2.4. Investigative Questions

Once the research question (s) has been selected, researcher thinking moves to a more specific level, that of investigative questions. These questions reveal the specific pieces of information the manager fills he or she needs to know to answer the research question. Investigative questions are questions the researcher must answer to satisfactorily arrive at a conclusion about the research question. To formulate them, the researcher takes a general research question and breaks it into more specific questions about which to gather data. This fractioning process can continue down through several levels of increasing specificity. Investigative questions should be included in the research proposal, for they guide the development of the research design. They are the foundation for creating the research data collection instrument.

The researcher working on the CBE project develops two major investigative questions for studying the market with several subquestions under each. The questions provide insight into the lack of deposit growth:

1. What are the public's position regarding financial services and their use?
    a) What specific financial services are used?
    b) How attractive are various services?
    c) What bank-specific and environmental factors influence a person's use of a particular service?
2. What is the bank's competitive position?
    a) What are the geographic patterns of our customers and of our competitors' customers?
    b) What demographic differences are revealed among our customers and those of our competitors?
    c) What words or phrases does the public (both customers and noncustomers) associate with CBE? With CBE competitors?
    d) How aware is the public of the bank's promotional efforts?
    e) What opinion does the public hold of the bank and its competitors?
    f) How does growth in services compare among the competing institutions?

2.5. Measurement Questions

Measurement questions should be outlined by completion of the project-planning activities but usually await pilot testing for refinement. There are two types of measurement questions: predesigned, pretested questions and custom-design questions. Predesigned measurement questions are questions that have been formulated and tested by previous researchers, are recorded in the literature, and may be applied literally or be adapted for the project at hand. Some studies lend themselves to the use of these readily available measurement devices. This provides enhanced validity and can produce the cost of the project. More often, however, the measurement questions should be custom-tailored to the investigative questions. The resources for this task will be the collective insights for all the activities in the research process completed to this point, particularly

insights from exploration. Later, during pilot testing of the data collection instrument (s), these custom-designed questions will be refined.

Measurement questions constitute the fifth level of the hierarchy. In surveys, **measurement question** are the questions we actually ask the respondents. They appear in the questionnaire. In an observation study, measurement questions are the observations researchers must record about each subject studied.

CBE conducts a survey of local residents. The questionnaire contains many measurement questions seeking information that will provide answers to the investigative questions. Two hundred residents complete questionnaires, and information collected is used to guide a reorientation of the bank's image.

The assumptions and facts used to structure the management-research question hierarchy set the direction of the project. Using the hierarchy is a good way to think methodically about the various issues. Think of the hierarchy as six sequential levels moving from the general to the specific. While our approach suggests six discrete levels-concluding with the management decision-the hierarchy is actually more of a continuum. The investigative question stage, in particular, may involve several levels of questioning before it is possible to develop a satisfactory measurement questions.

Activity 1

Dear learner, based on the above discussions

1. Order hierarchically the Management-research question hierarchy (from the management dilemma to the management question) in the research process.
2. Briefly describe the terms management question, research question, investigative question and measurement question.

3. Designing the Study

The **research design** is a blueprint for fulfilling objectives and answering questions. Selecting a design may be complicated by the availability of a large variety methods, techniques, procedures, protocols, and sampling plans. For example, you may decide on a secondary data study, case study, survey, experiment, or simulation. If survey is selected, should it be administered by mail, computer, telephone, the internet, or personal interview? Should all relevant data be collected at one time or at regular intervals? What kind of structure will the questionnaire or interview guide possess? What question wording should be employed? Should the responses be scaled or opened? How will reliability and validity be achieved? Will characteristics of the interviewer influence responses to the management questions? What kind of training should the data collectors receive? Is a sample or a census to be taken? What type of sampling should be considered? Theses questions represent only a few of the decisions that have to be made when just one method is chosen.

The creative researcher actually benefits from its confusing array of options. The numerous combinations spawned by the abundance of tools may be used to construct alternative perspectives on the same problem. By creating a design using diverse methodologies, researchers are able to achieve greater insight than if they followed the most frequent method encountered in the literature or suggested by a disciplinary bias. Although it may be conceded that students or mangers rarely have the resources to pursue a single problem from multimethod, multistudy strategy, the advantages of several competing designs should be considered before settling on a final one.

4. Sampling Design

Another step in planning the design is to identify the target population and select the sample if the census is not desired. The researcher must determine who and how many people to interview, what and how many events to observe, or what and how many record to inspect. A **sample** is a part of the target population, carefully selected to represent that population. When researchers undertake sampling studies, they are interested in

estimating one or more population values and/or testing one or more statistical hypotheses.

If a study's objective is to examine the attitude of Ethiopia automobile assemblers about quality improvement, the population may be defined as the entire adult population of auto assemblers employed by the auto industry in Ethiopia. Definition of the terms *adult* and *assembler* and the relevant job descriptions included under "assembly" and "auto industry" may further limit the population under study. The investigator may also want to restrict the research to readily identifiable companies in the market, vehicle types, or assembly process.

The sampling process must then give every person within the target population a known nonzero chance of selection if probability sampling is used. If there is no feasible alternative, a nonprobability approach may be used.

5. Resources Allocation and Budgets

General notion about research budgets have a tendency to single out data collection as the most costly activity. Data collection requires substantial resources but perhaps less of the budget than clients expect. Employees must be paid, training and travel must be provided, and other expenses incurred must be paid; but this phase of the project often takes no more than one-third of the total research budget. The geographic scope and the number of observations required do affect the cost, but much of the cost is relatively independent of the size of the data-gathering effort. Thus, a guide might be that (1) project planning, (2) data gathering, and (3) analysis, interpretation, and reporting each shares about equally in the budget.

Without budgetary approval, many research efforts are terminated for lack of resources. A budget may require significant development and documentation as in grant and contract research, or it may require less attention as in some in-house projects or investigations funded out of researcher's own resources. The researcher who seeks

funding must be able not only to persuasively justify the costs of the project but also to identify the sources and methods of funding.

6. The Research Proposal

The proposal incorporates the choices the investigator makes in the preliminary steps. A written proposal is often required when a study is being suggested. It ensures that the parties concur on the project's purpose and on the proposed methods of investigation. Time and budgets are often spelled out, as are other responsibilities and obligations. Depending on the needs and desires of the manager, substantial background detail and elaboration of proposed techniques may be included.

The length and complexity of proposal range widely. Accounting and finance research proposals normally range from 1 to 10 pages. Applicants for foundation or government research grants typically file an extensive proposal, often in a standardized format specified by the granting agency. A research proposal may also be oral, where all aspects of the research are discussed but not codified in writing. This is more likely when a manger directs his or her own research or the research activities of subordinates.

Proposal Content

Every proposal regardless of length should include two basic sections:

- Statement of the research question.
- Brief description of research methodology.

In a brief memo type proposal, the research question may be incorporated into a paragraph that also sets put the management dilemma, management questions, and categories of investigative questions. The following statement present the management question facing the manger and point out the nature of the research that will be undertaken:

Commercial Bank of Ethiopia (or CBE), currently the leading bank in the city, has not been growing as fast as its major competitors. Before developing a long-range plan to

enhance the bank's competitive position, it is important to determine the bank's present competitive status, its advantages and opportunities, and its major deficiencies. The primary objective of this proposed research is to develop a body of benchmark information about CBE, its major competitors, and the market for banking services.

Often research proposals are much more detailed and describe specific measurement devices that will be used, time and cost budgets, sampling plans, and many other details.

7. Pilot Testing

The data-gathering phase of the research process typical begins with pilot testing. Pilot testing may be skipped when the researcher tries to condense the project time frame.

A **pilot test** is conducted to detect weaknesses in design and instrumentation and to provide proxy data for selection of a probability sample. It should, therefore, draw subjects from the target population and simulate the procedures and protocols that have been designed for data collection. If the study is a survey to be executed by mail, the pilot questionnaire should be mailed. If the design calls for observation, by unobtrusive researcher, this behavior should be practiced. The size of the pilot group may range from 25 to 100 subjects, depending on the method to be tested, but the respondents do not have to be statistically selected. In very small populations or special applications, pilot testing runs the risk of exhausting the supply of respondents and sensitizing them to the purpose of the study. This risk is generally overshadowed by the improvements made to the design by a trial run.

There are a number of variations on pilot testing. Some of them are intentionally restricted to data collection activities. One form, *pretesting*, may rely on colleagues, respondent surrogates, or actual respondents to refine a measuring instrument. This important activity has saved countless survey studies from disaster by using the suggestions of the respondents to identify and change confusing, awkward, or offensive questions and techniques. One interview study was designed by a group of college professors for EducTV, an educational television consortium. In the pilot test, they discovered that the wording of nearly two-thirds of the questions was unintelligent to the

target group, later found to have a median eighth-grade education. The revised instrument used the respondents' language and was successful. Pretesting may be repeated several times to refine questions, instruments, or procedures.

8. Data Collection

The gathering of data may range from a simple observation at one location to a grandiose survey of multinational corporations at sites in different parts of the world. The method selected will largely determine how the data are collected. Questionnaires, standardized tests, observational forms, laboratory notes, and instrument calibration logs are among the devices used to record raw data.

But what are data? One writer defines **data** as the facts presented to the researcher from the study's environment. First, data may be further characterized by their abstractness, verifiability, elusiveness, and closeness to the phenomenon. As *abstractions*, data are more metaphorical than real. For example, the growth in GDP cannot be observed directly; only the effects of it may be recorded. Second, data are processed by our senses-often limited in comparison to the senses of other living organism. When sensory experiences consistently produce the same result, our data are said to be trustworthy because they may be *verified*. Third, capturing data is *elusive*, complicated by the speed at which events occur and the time-bound nature of observation. Opinions, preferences, and attitudes about spending during the late 1980s differed dramatically one decade later within the same population, due to the sustained prosperity within the final four years of the millennium. Finally, data reflect their truthfulness by *closeness to the phenomena*. *Secondary data* have had at least one level of interpretation inserted between the event and its recording. *Primary data* are sought for their proximity to the truth and control overt error. These cautions remind us to use care in designing data collection procedures and generalizing from results.

Data are edited to ensure consistency across respondents and to locate omissions. In the case of survey methods, editing reduces errors in the recording, improves legibility, and clarifies unclear and inappropriate responses. Edited data are then put into a form that

makes analysis possible. Because it is impractical to place raw data into a report, alphanumeric codes are used to reduce the responses to a more manageable system for storage and future processing. The codes follow various decision rules that the researcher has devised to assist with sorting, tabulating, and analyzing. Personal computers have made it possible to merge editing, coding, and data entry into fewer steps even when the final analysis may be run on a large system.

9. Analysis and Interpretation

Managers need information, not raw data. Researchers generate information by analyzing data after its collection. Data analysis usually involves reducing accumulated data to a manageable size, developing summaries, looking for patterns, and applying statistical techniques. Scaled responses on questionnaires and experimental instruments often require the analyst to derive various functions, as well as to explore relationships among variables. Further, researchers must interpret these findings in light of the client's research question or determine if the results are consistent with their hypotheses and theories. Increasingly, managers are asking research specialists to make recommendations based on their interpretation of the data.

A modest example involves a market research firm that polls 2,000 people from its target population for a new generation of wallet-sized portable telephones. Each respondent will be asked four questions:

1. "Do you prefer the convenience of Pocket-Phone over existing cellular telephone?"
2. "Are there transmission problems with Pocket-Phone?"
3. "Is Pocket-Phone better suited to worldwide transmission than your existing cellular phone?"
4. "Would cost alone persuade you to purchase Pocket-Phone?"

The answers will produce 8,000 pieces of raw data. Reducing the data to a workable size yield eight statistics: the percentage of 'yes' and 'no' answers to each question. When half-dozen demographic questions about the respondents are added, the total amount of data easily triples. If the researcher scaled the four key questions rather than eliciting yes-

no responses, the analysis would likely require more powerful statistical analysis than summarization.

10. Reporting the Results

Finally, it is necessary to prepare a report and transmit the findings and recommendations to the manger for the intended purpose of decision making. The researcher adjusts the style and organization of the report according to the target audience, the occasion, and the purpose of the research. The results of applied research may be communicated via conference call, letter, written report, oral presentation, or some combination of any or all of these methods. Reports should be developed from the manager's or information user's perspective. The sophistication of the design and sampling plan or the software used to analyze the data may help to establish the researcher's credibility, but in the end, the manager's foremost concern is solving the management dilemma. Thus, the researcher must accurately assess the manager's needs throughout the research process and incorporate this understanding into the final product, the research report.

The management decision maker occasionally shelves the research report without taking action. Inferior communication of results is a primary reason for this outcome. With this possibility in mind, a research specialist should strive for:

- Insightful adaptation of the information to the client's needs.
- Careful choice of words in crafting interpretations, conclusions and recommendations.

At a minimum, a research report should contain the following:

- An *executive summary* consisting of a synopsis of the problem, findings, and recommendations.
- An *overview of the research*: the problem's background, literature summary, methods and procedures, and conclusions.
- A section on *implementation strategies* for the recommendations.
- A *technical appendix* with all the materials necessary to replicate the project.

Review Questions

1. Some questions are answerable by research and others are not. Using some management problems of your choice, distinguish between them.
2. A company is experiencing a poor inventory management situation and receives alternative research proposals. Proposal 1 is to use an audit of last year's transactions as a basis for recommendations. Proposal 2 is to study and recommend changes to the procedures and systems used by the materials department. Discuss issues of evaluation in terms of *Ex post facto versus prior evaluation*.
3. Develop the management-research question hierarchy for a management dilemma you face at work or with an organization for which you are volunteer.

Unit 3: Business Research Proposals

Learning Objectives:

After completion of this unit, you should be able to:

- Identify the purpose of the proposal;
- Discuss how proposal is used by the researcher and management decision maker;
- List the content of a research proposal;
- Explain the processes for evaluating the quality of proposals; and
- Produce a complete research proposal.

Proposing Research: An Introduction

Many students and some business researchers view the proposal process as unnecessary work. In actuality, the more inexperienced a researcher is, the more important it is to have a well-planned and adequately documented proposal. When the organization has research specialists on the payroll, the internal research proposal is often all that is needed. Often, however, companies do not have adequate capacity, resources, or the specialized talents in-house to execute a project, so they turn to outside research suppliers (including research specialists, universities, research centers, and consulting firms).

1. The Research Proposal

A proposal is an individual's or company's offer to produce a product or render a service to a particular buyer or sponsor. The purpose of the research proposal is:

1. To present the management question to be researched and relate its importance.
2. To discuss the research efforts of others who have worked on related management questions.
3. To suggest the data necessary for solving the management question and how the data will be gathered, treated, and interpreted.

In addition, a research proposal must present the researcher's plan, services, and credentials in the best possible way to encourage the proposal's selection over competitors. In contract research, survival companies depend on their ability to develop winning proposals. A proposal is also known as a work plan, prospectus, outline, statement of intent, or draft plan. The proposal tells us what, why, how where, and to whom the research will be done. It must also show the benefit of doing the research.

The research proposal is essentially a road map, showing clearly the location from which a journey begins, the destination to be reached, and the method of getting there. Well-prepared proposals include potential problems that may be encountered along the way and methods for avoiding or working around them, much as a road map indicates alternate routes for a detour.

Sponsor Uses

All research has a sponsor in one form or another. The student researcher is responsible to the class instructor. In a corporate setting, whether the research is being done in-house by a research department or under contract to an external research firm, management sponsors the research. University-, government-, or corporate- sponsored (grant) research uses grant committees to evaluate the work.

A research proposal allows the sponsor to assess the sincerity of the researcher's purpose, the clarity of his or her design, the extent of his or her relevant background material, and the researcher's fitness for undertaking the project. Depending on the type research and the sponsor, various aspects of a standard proposal design are emphasized. The proposal displays the researcher's discipline, organization, and logic. It thus allows the research sponsor to assess both the researcher and the proposed design, to compare them against competing proposals on current organizational, scholastic, or scientific needs, and to make the best selection for the project.

Comparison of the research project results with the proposal is also the first step in the process of evaluating the overall research. By comparing the final product with the stated

objectives, it is easy for the sponsor to decide if the research goal-a better decision on the management question-has been achieved.

Another benefit of the proposal is the discipline it brings to the sponsor. Many mangers, requesting research from an in-house, departmental research project, do not adequately define the problem they are addressing. The research proposal acts as a catalyst for discussion between the person conducting the research and the manager. The researcher translates the management question, as described by the manager, into the research question and outlines the objectives of the study. Upon review, the manger may discover that the interpretation of the problem does not encompass all the original symptoms. The proposal, then, serves as the basis for additional discussion between the manager and the researcher until all aspects of the management question are understood. Parts of the management question may not be researchable, or at least not subject to empirical study. An alternate design, such as a qualitative or policy analysis study, may need to be proposed. Upon completion of the discussions, the sponsor and researcher should agree on a carefully worded research question.

2. Structuring the Research Proposal

Each of the following modules is flexible, so its content and length may be adapted to specific needs.

2.1. Executive Summary

The executive summary allows a busy manager or sponsor to understand quickly the thrust of the proposal. It is essentially an informative abstract, giving executives the chance to grasp the essentials of the proposal without having to read the details. The goal of the summary is to secure a positive evaluation by the executive who will pass the proposal on to the staff for a full evaluation. As such, the executive summary should include brief statements of the management dilemma and management question, the research objectives/research question (s), and the benefits of your approach. If the proposal is unsolicited, a brief description of your qualifications is also appropriate.

2.2. Problem Statement

This section needs to convince the sponsor to continue reading the proposal. You should capture the reader's attention by stating the management dilemma, its background, its consequences, and the resulting management question. The importance of answering the management question should emphasize here if a separate module on the importance/benefits of the study is not included later in the proposal. In addition, this section should include any restrictions or areas of the management question that will not be addressed.

Problem statements too broadly defined can not be addressed adequately in one study. It is important that the management question distinguish the primary problem from related problems clearly. Be sure your problem statement is clear without the use of idioms or clichés. After reading this section, the potential sponsor should know the management dilemma and the question, its significance, and why something should be done to change the status quo.

2.3. Research Objectives

This module addresses the purpose of the investigation. It is here that you lay out exactly what is being planned by the proposed research. In a descriptive study, the objectives can be stated as the research question. Recall that the research question can be further broken down into investigative questions. If the proposal is for a causal study, then the objectives can be restated as a hypothesis.

The objectives module flows naturally from the problem statement, giving the sponsor specific, concrete, and achievable goals. It is best to list the objectives either in order of importance or in general terms first, moving to specific terms (i.e., research question followed by underlying investigative questions). The research question (s) (or hypotheses, if appropriate) should be separated from the flow of the text for quick identification.

The research objectives section is the basis for judging the remainder of the proposal and, ultimately, the final report. Verify the consistency of the proposal by checking to see that each objective is discussed in the research design, data analysis, and results section.

2.4. Literature Review

The literature review section examines recent (or historical significant) research studies, company data, or industry reports that act as a basis for the proposed study. Begin your discussion of the related literature and relevant secondary data from a comprehensive perspective, moving to more specific studies that are associated with your problem. If the problem has a historical background, begin with the earliest references.

Avoid the extraneous details of the literature; do a brief review of the information, not a comprehensive report. Always refer to the original source. If you find something of interest in a quotation, find the original publication and ensure you understand it. In this way, you will avoid your errors of interpretation or transcription. Emphasize the important results and conclusions of other studies, the relevant data and trends from previous research, and particular methods or designs that could be duplicated or should be avoided. Discuss how the literature applies to the study you are proposing; show the weakness of faults in the design, discussing how you would avoid similar problems. If your proposal deals solely with secondary data, discuss the relevance of the data and the bias or lack of bias inherent in it.

The literature review may also explain the need for the proposed work to appraise the shortcomings and/or informational gaps in secondary data sources. This analysis may go beyond scrutinizing the availability or conclusions of past studies and their data, to examining the accuracy of secondary sources, the credibility of these sources, and the appropriateness of earlier studies.

Close the literature review section by summarizing the important aspects of the literature and interpreting them in terms of your problem. Refine the problem as necessary in light of your findings.

2.5. Importance/ Benefits of the Study

In this section you describe the explicit benefits that will accrue from your study. The importance of "doing the study now" should be emphasized. Usually this section is not more than a few paragraphs. If you find it difficult to write, then you have probably not adequately clarified the management dilemma. Return to the analysis of the problem and ensure, through additional discussions with your sponsor or your research team or by reexamination of the literature, that you have captured the essence of the problem.

This section also requires you to understand what is most troubling to your sponsor. If it is a potential union activity, you cannot promise that an employee survey will prevent unionization. You can, however, show the importance of this information and its implications. This benefit may allow management to respond to employee concerns and forgo a linkage between those concerns and unionization.

The importance/benefit section is particularly important to the unsolicited external proposal. You must convince the sponsoring organization that you plan will meet its needs.

2.6. Research Design

Up to now, you have told the sponsor what the problem is, what your study goals are, and why it is important for you to do the study. The proposal has presented the study's value and benefits. The design module describes what you are going to do in technical terms. This section should include as many subsections as needed to show the phases of the project. Provide information on your proposed design for tasks such as sample selection and size, data collection method, instrumentation, procedures, and ethical requirements. When more than one way exists to approach the design, discuss the methods you have rejected and why your selected approach is superior.

2.7. Data Analysis

A brief section on the methods used for analyzing the data is appropriate for large-scale contract research projects and doctoral theses. With smaller projects, the proposed data

analysis would be included within the research design section. It is in this section that you proposed handling of the data and the theoretical basis for using the selected techniques. The object of this section is to assure the sponsor you are following correct assumptions and using theoretically sound data analysis procedures.

This module is often an arduous section to write. You can make it easier to write, read, and understand your data analysis by using simple charts and tables featuring "dummy" data.

The data analysis section is so important to evaluating contract research proposals that the researcher should contact an expert to review the latest techniques available for use in the particular research study and compare these to the proposed techniques.

2.8. Nature and Form of Results

Upon finishing this section, the sponsor should be able to go back to the statement of the management question and the research objectives and discover that each goal of the study has been covered. One should also specify the types of data to be obtained and the interpretations that will be made in the analysis. If the data are to be turned over the sponsor for proprietary reasons, make sure this is reflected. Alternatively, if the report will go to more than one sponsor, that should be noted.

This section also contains the contractual statement telling the sponsor exactly what types of information will be received. Statistical conclusions, applied findings, recommendations, action plans, models, strategic plans, and so forth, are examples of the form of results.

2.9. Budget

The budget should be presented in the form the sponsor requests. For example, some organizations require secretarial assistance to be individually budgeted, whereas others insist it be included in the research director's fees or the overhead of the operation. In

addition, limitations on travel, per diem rates, and capital equipment purchases can change the way you prepare a budget.

Typically the budget should be no more than one to two pages. Additional information, back up details, quotes from vendors, and hourly time and payment calculations should be put into an appendix if required or kept in the researcher's file for future reference.
The budget statement in an internal research proposal is based on employee and overhead costs. The budget presented by an external research organization is not just the wages or salaries of its employees but the person-hour price that the contacting firm charges.

It is extremely important that you retain all information you use to generate your budget. If you use quotes from external contractors, get the quotation in writing for your file. If you estimate time for interviews, keep explicit notes on how you make the estimate. When the time comes to do the work, you should know exactly how much money is budgeted for each particular task.

Some costs are more elusive than others. Do not forget to build the cost of proposal writing into your fee. Publication and delivery of final reports can be a last-minute expense that may be easily overlooked in preliminary budgets.

2.10. Schedule

Your schedule should include the major phases of the project, their timetables, and the milestones that signify completion of the phase. Foe example, major phases may be (1) exploratory interviews, (2) final research proposal, (3) questionnaire revision, (4) field interviews, (5) editing and coding,(6) data analysis, and (7) report generation. Each of these phases should have an estimated time schedule and people assigned to the work.

It may be helpful to and to your sponsor if you chart your schedule. Alternatively, if the project is large and complex, a critical path method (CPM) of scheduling may be included.

2.11. Bibliography

For all projects that require a literature review, a bibliography is necessary. Use the bibliographic format required by the sponsor. If none is specified, a standard style manual will provide the details necessary to prepare the bibliography. Many of these sources also offer suggestions for successful proposal writing.

2.12. Appendices

Glossary

The researcher should include a glossary of terms whenever there are many words unique to the research topic and not understood by the general management community. This is a simple section consisting of terms and definitions, similar in format to the glossary in different reference books. Also the researcher should define any acronyms used, even if they are defined within the text.

Activity 1

1. What is research proposal and what is its purpose?
2. Assume that you have asked by someone about proposal, how could you present the structure of a proposal?

3. Evaluating the Research Proposal

Proposals are subject to either formal or informal reviews. Formal reviews are regularly done for solicited proposals. The formal review process varies, but typically includes:

- Development of review criteria.
- Assignment of points to each criterion, using a universal scale.
- Assignment of a weight for each criterion, based on importance of each criterion.
- Generation of a score for each proposal, representing the sum of all weighed criterion scores.

In practice, many factors contribute to a proposal's acceptance and funding. Primarily, the content discussed above must be included to the level of detail required by the

sponsor. Beyond the required modules, other factors can quickly eliminate a proposal from consideration or improve the sponsor's reception of the proposal, among them:

- Neatness.
- Organization, in terms of being both logical and easily understood.
- Completeness in fulfilling the specifications, including budget and schedule.
- Appropriateness of writing style.
- Submission within the timeline.

Although a proposal produced on a word processor and bound with an expensive cover will not overcome design or analysis deficiencies, a poorly presented, unclear, or disorganized proposal will not get a serious attention from the reviewing sponsor.

In terms of the technical writing style of the proposal, the sponsor must be able to understand the problem statement, the research design, and the methodology. The sponsor should clearly understand why the proposed research should be funded and the exact goals and concrete results that will come from the study.

The proposal must also meet specific guidelines set by the sponsoring company or agency, including budgetary restrictions and schedule deadlines. A schedule that does not meet the expected deadlines will disqualify the proposal. A budget that is too high for the allocated funds will be rejected. Conversely, a low budget compared to competing proposals suggests that something is missing or there is something wrong with the researchers.

Finally, a late proposal will not be reviewed. While current project disqualification due to lateness may appear to be the worst result here, there is a possible long-term effect created. Lateness communicates a level of disrespect for the sponsor-that the researcher's schedule is more important than the sponsor's. A late proposal also communicates a weakness in project management, which raises an issue of professional competence. This concern about competence may continue tom plague the researcher during future project proposal reviews.

Review Questions

1. What is the purpose of a research proposal?

2. List the contents of a proposal?

3. Contemplate in your mind the revision of a research proposal. What could be the criteria for reviewing proposal?

4. How could you explain the relationship between problem statement and research objectives?

5. Discuss how proposal is used by the researcher and management decision maker.

Unit 4: The Research Design

Learning Objectives:

After completion of this unit, you should be able to:

- Identify the basic stages of research design;
- Distinguish the major descriptors of research design;
- Discuss the major types of research designs;
- Identify the relationships that exist between variables in research design; and
- List the steps for evaluating the relationships that exist between variables.

1. What is Research Design?

The topics covered by the term research design are wide-ranging. This unit introduces a classification of research designs and provides an overview of the most important design types (exploratory, descriptive, and casual). You can refer to subsequent units for a more thorough coverage of the unique features of qualitative studies and surveys. The objective here is not for you to acquire the details of research design in one reading but for you to understand its scope and to get a glimpse of the available options for tailoring a design to an organization's particular research needs.

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.

These definitions differ in detail, but together they give the essentials of research design:

- An activity – and time –based plan.
- A plan always based on the 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.

2. Classification of Designs

Early in any research study, one faces the task of selecting the specific design to use. A number of different design approaches exist, but, unfortunately, no simple classification system defines all the variations that must be considered. Exhibit 4-1 classifies research design using eight different descriptors. A brief discussion of these descriptors illustrates their nature and contribution to the research.

Exhibit 4-1: **Descriptors of Research Design**

Category	*Options*
The degree to which the research question has been crystallized	• Exploratory study • Formal study
The method of data collection	• Monitoring • Communication study
The power of the researcher to produce effects in the variables under study	• Experimental • Ex post facto
The purpose of the study	• Descriptive • Causal

The time dimension	• Cross-sectional • Longitudinal
The topical scope-breadth and depth-of the study	• Case • Statistical study
The research environment	• Field setting • Laboratory research • Simulation
The participants' perception of research activity	• Actual routine • Modified routine

2.1. Degree of Research Question Crystallization

A study may be viewed as exploratory or formal. The essential distinctions between these two options are the degree of structure and the immediate objective of the study. **Exploratory studies** tend toward loose structures with the objective of discovering future research tasks. The immediate purpose of exploration is usually to develop hypotheses or questions for further research. The formal study begins where the exploration leaves off-it begins with a hypothesis or research question and involves precise procedures and data source specifications. The goal of formal research design is to test the hypotheses or answer the research questions posed.

The exploratory-formal study dichotomy is less precise than some other classifications. All studies have elements of exploration in them, and few studies are completely uncharted. The sequence discussed in unit 2 suggests that more formalized studies contain at least an element of exploration before the final choice of design.

2.2. Method of Data Collection

This classification distinguishes between **monitoring** and communication processes. The former 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. Traffic counts at an intersection, license plates recorded in a restaurant parking lot, a search of

the library collection, an observation of the actions of a group of decision makers-all are examples of monitoring. In each case the researcher notes and records the information available from observations.

In the **communication study**, the researcher questions the subjects and collects their responses by personal means. The collected data may result from (1) interview or telephone conversations, (2) self-administered or self-reported instruments set through the mail, left in convenient locations, or transmitted electronically or by other means, or (3) instruments presented before and/or after a treatment or stimulus condition in an *experiment*.

2.3. Researcher Control of Variables

In terms of the researcher's ability to manipulate variables, we differentiate between experimental and ex post facto designs. In an **experiment**, the researcher attempts to control and/or manipulate the variables in the study. It is enough that we can cause variables to be changed or held constant in keeping with our research objectives. 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. It is important that the researchers using this design not influence the variables; to do so introduce 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.

2.4. The Purpose of the Study

The essential difference between descriptive and causal studies lies in their objectives. If the research is concerned with finding out *who, what, where, when,* or *how much,* then the study is **descriptive**. If it concerned with learning why-that is, how one variable

produces changes in another-it is causal. Research on crime is descriptive when it measures the types of crimes committed, how often, when, where, and by whom. In a **causal study**, we try to explain relationships among variables-for instance, why the crime rate is higher in a city A that in city B.

2.5. 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 studies of the *panel* variety, the researcher may study the same people over time. In marketing, panels are set up to report consumption data on a variety of products. These data, collected from national samples, provide a major data bank on relative market share, consumer response to new products, and new promotional methods. Other longitudinal studies, such as *cohort groups*, use different subjects for each sequenced measurement. The service industry might have looked at the needs of aging baby boomers by sampling 40- to 45-years-olds in 1990 and 50- to- 55-years-olds in 2000. Although each sample would be different, the population of 1945 to 1950 cohort survivors would remain the same.

Some types of information once collected cannot be collected a second time from the same person without the risk of bias. The study of public awareness of an advertising campaign over a six-month period would require different samples for each measurement.

While longitudinal research is important, the constraints of budget and time impose the need for cross-sectional analysis. Some benefits of a longitudinal study can be revealed in a cross-sectional study by adroit questioning about past attitudes, history, and future expectations. Responses to these kinds of questions should be interpreted with care, however.

2.6. The Topical Scope

The statistical study differs from the case study in several ways. **Statistical studies** are designed for breadth rather than depth. They attempt to capture a population's characteristics by making inferences from a sample's characteristics. Hypotheses are tested quantitatively. Generalizations about findings are presented based on the representativeness of the sample and the validity of the design.

Case studies place more emphasis on a full contextual analysis of fewer events and conditions and their interrelations. Although hypotheses are also often used, the reliance on qualitative data makes support or rejection more difficult. An emphasis on detail provides valuable insights 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 case studies have been maligned as "scientifically worthless" because they do not meet minimal design requirements for comparison, they have a significant scientific role. It is known that "important scientific propositions have the form of universals, and a universal can be falsified by a single counterinstance." Thus, a single, well-designed case study can provide a major challenge to a theory and provide a source of new hypotheses and constructs simultaneously.

2.7. The Research Environment

Designs also differ as to whether they occur under actual environmental conditions (**field conditions**) or under staged or manipulated conditions (**laboratory conditions**)

To simulate is to replicate the essence of a system or process. **Simulations** are increasingly used in research, especially in operations research. The major characteristics of various conditions and relationships in actual situations are often represented in mathematical models. Role-playing and other behavioral activities may also be viewed as simulations. A popularly used simulation is the retail service study involving "mystery shoppers."

2.8. Participants' Perceptions

The usefulness of a design may be reduced when a people in a disguised study perceive that research is being conducted. Participants' perceive that research is being conducted. **Participants' perceptions** influence the outcomes of the research in subtle ways or more dramatically as we learned from pivotal Hawthorne studies of the late 1920s. Although there is no wide spread evidence of attempts by participants or respondents to please researchers through successful hypothesis guessing or evidence of the prevalence of sabotage, when participants believe that something out of the ordinary is happening, they may behave less naturally.

There are three levels of perception:

1. Participants perceive no deviations from everyday routines.
2. Participants perceive deviations, but as unrelated to the researcher.
3. Participants perceive deviations as researcher-induced.

The "mystery shoppers" scenario is the perfect example of the final level of perception noted in the above list. If a retail sales associate knows she is being observed and evaluated-with consequences in future compensation, scheduling, or work assignment-she is likely to change her performance. In all research environments and control situations, researchers need to be vigilant to effects that may alter their conclusions. Participants' perceptions serve as a reminder to classify one's study by type, to examine validation strengths and weaknesses, and to be prepared to qualify results accordingly.

Activity 1

1. From the above discussion what is research design?
2. What are the essentials of research design?
3. Distinguish between formal and exploratory studies.
4. Research design is classified into two based on the researcher's control over the variables. Describe them briefly.

3. Exploratory Studies

Exploration is particularly useful when researchers lack a clear idea of the problems they will meet during the study. Through exploration researchers develop concepts more clearly, develop operational definitions, and improve the final research design. Exploration may also save time and money. If the problem is not as important as first thought, more formal studies can be canceled.

Exploration serves other purposes as well. The area of investigation may be so new or so vague that a researcher needs to do an exploration just to learn something about the dilemma facing the manager. Important variables may not be known or thoroughly defined. Hypotheses for the research may be needed. A federal government agency, the Office of Industry Analysis, proposed that research be done on how executives in a given industry made decisions about raw material purchases. Questions were planned asking how (and at what price spreads) one raw material was substituted for another in certain manufactured products. An exploration to discover if industry executives would reveal adequate information about their decision making on this topic was essential for the study's success.

Despite its obvious value, researchers give exploration less attention than it deserves. There are strong pressures for quick answers. Moreover, exploration is sometimes linked to old biases about qualitative research: subjectiveness, nonrepresentativeness, and nonsystematic design. More realistically, exploration saves time and money and should not be slighted.

The objectives of exploration may be accomplished with two different techniques. Both qualitative and quantitative techniques are applicable, although exploration relies more heavily on qualitative techniques. When we consider the scope of qualitative research, several approaches are adaptable for exploratory investigations of management questions:

- Individual depth interviews (usually conversational rather than structured).
- Participant observation (to perceive firsthand what participants in the setting 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 subgroup describes and structures its world at the street level).
- Elite or expert interviewing (for information from influential or well-informed people in an organization or community).
- 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 with wide applicability for the accounting researcher:

1. Secondary data analysis,
2. Experience surveys,
3. Focus groups,
4. Two-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 at a level sufficient for management to make a decision.

Within secondary data exploration, a researcher should start first with an organization's own data archives. Reports of prior research studies often reveal an extensive amount of historical data or decision-making patterns. The second source of secondary data is published documents prepared by authors outside the sponsor organization. There are numerous periodicals and books on all aspects of business. Data from secondary sources help us decide what needs to be done and can be a rich source of hypotheses.

Experience Survey

While published data are a valuable resource, it is seldom that more than a fraction of the existing knowledge in a field is put into writing. A significant portion of what is known on a topic, while in writing, may be proprietary to a given organization and thus unavailable to an outside researcher. Also, internal data archives are rarely well organized, making secondary sources, even when known, difficult to locate. Thus we will profit by seeking information from persons experienced in the area of study, tapping into their collective memories and experiences.

When we interview persons in an experience survey, we should seek their ideas about important issues or aspects of the subject and discover what is important across the subject's range of knowledge. The investigative format we use should be flexible enough so that we can explore various avenues that emerge during the interview.

Focus Groups

A focus group is a group of people (typically 6 to 10 participants), led by a trained moderator, who met for 90 minutes to 2 hours. The facilitator of moderator uses group dynamics principles to focus guide the group in an exchange of ideas, feelings, and experiences on a specific topic. The output of the session is a list of ideas and behavioral observations, with recommendations by the moderator. In exploratory research, the qualitative data that focus groups produce may be used for enriching all levels of research questions and hypotheses and comparing the effectiveness of design options.

Two-Stage Design

A useful way to design a research study is as a **two-stage design**. With this approach, exploration becomes a separate first stage with limited objectives: (1) clearly defining the research question and (2) developing the research design.

In arguing for a two-stage approach, we recognize that much about the problem is not known but should be known before effort and resources are committed. In these circumstances, one is operating in unknown areas, where it is difficult to predict the

problems and costs of the study. Proposals that acknowledge the practicality of this approach are particularly useful when the research budget is inflexible. A limited exploration for a specific, modest cost carries little risk for both sponsor and researcher and often uncovers information that reduces the total research cost.

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

- Established the major dimensions of the research task.
- Defined a set of subsidiary investigative questions that can be used as guides to a detailed research design.
- Developed several hypotheses about possible causes of a management dilemma.
- Learned that certain other hypotheses are remote possibilities that they can be safely ignored in any subsequent study.
- Concluded additional research is not needed or is not feasible.

4. Descriptive Studies

In contrast to exploratory studies, more formalized studies are typically structured with clearly stated hypotheses or investigative questions. Formal studies serve a variety of research objectives:

1. Descriptions of phenomena or characteristics associated with a subject population (the *who*, *what*, *when*, *where*, and *how* of a topic).
2. Estimates of the proportions of a population that have these characteristics.
3. Discovery of associations among different variables.

The thirds study objective is sometimes labeled a *correlational study*, a subset of descriptive studies. A descriptive study may be simple or complex; it may be done in many settings. Whatever the form, a descriptive study can be just as demanding of research skills as the casual study, and we should insist on the same high standards for design and execution.

The simplest descriptive 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 CBE, we might be interested in developing a profile of savers. We first may want to locate them in relation to the main office. The question might be, "What percentage of the savers will live within a 3-Kilometer radius of the office?" Using the hypothesis format, we might predict, "60 percent or more of the savers line within a 3-Kiliometer radius of the office."

We may also be interested in securing information about other variables, such as the relative size of accounts, the number of accounts for the minors, the number of accounts opened within the last six months, and the amount of activity (number of deposits and withdrawals per year) in accounts. Bivariate relationships between these or other variables may be of even greater interest. Cross-tabulations between the distance from the account owner's residence or employment to the branch and account activity may suggest that differential rates of activity are related to account owner location. A cross tabulation of account size and gender of account owner may also show interrelation. Such findings do not imply a causal relationship. In fact, our task is to determine if the variables are independent (or unrelated) and if they are not, then to determine the strength or magnitude of the relationship. Neither procedure tells us which variable is the cause. For example, we might be able to conclude that gender and account size are related but not that gender is a causal factor in account size.

Descriptive studies are often much more complex than this example. One study of savers began as described and then went into much greater depth. Part of the study included an observation of account records that revealed a concentration of nearby savers. Their accounts were typically larger and more active than those whose owners lived at a distance. A sample survey of savers provided information on stages in the family lifecycle, attitudes toward savings, family income levels, and other matters. Correlation of this information with known savings data showed that women owned larger accounts. Further investigation suggested that women with larger accounts were often widowed or working single women who were older than the average account holder.

The correlation between nearness to the office and the probability of having an account at the office suggested the question, "Why would people who live far from the office have an account there?" In this type of question a hypothesis makes its greatest contribution by pointing out directions that the researcher might follow. It might be hypothesized that:

1. Distant savers (operationally defined as those with addresses more than 3 Kilometer from the office) have accounts at the office because they once lived near the office; they were "near" when the account decision was made.
2. Distant savers actually live near the office, but the address on the account is outside the 3-Kilometer radius; they are "near" but the records do not show this.
3. Distant savers work near the office; they are "near" by virtue of their work location.
4. Distant savers are not normally near the office but responded to a promotion that encouraged savers to bank via computer; this is another form of "nearness" in which this concept is transformed into one of "convenience."

When these hypotheses were tested, it was learned that a substantial portion of the distant savers could be accounted for by hypotheses 1 and 3. The conclusion: Location was closely related to saving at a given association. The determination of case is not so simple, however, and these findings still fall within the definition of a descriptive study.

Activity 2

1. Distinguish between exploratory and descriptive studies.
2. Dear Learners, you are to begin reading about causal studies. Before we made the discussion have you heard of anything about these terms 'Symmetric', 'reciprocal', and 'asymmetric' relationship between variables. Any ways try to define each of them in your own words.

5. Causal Studies

The correlation between location and probability of account holding at CBE looks like strong evidence, but the researcher with scientific training will argue that correlation is not causation. Who is right? The essence the disagreement seems to lie in the concept of cause.

The essential element of causation is that A "produces" B or A "forces" B to occur. But that is an artifact of language, not what happens. Empirically, we can never demonstrate an A-B causality with certainty. This is because we do not "demonstrate" such causal linkages deductively or use the form or validation of premises that deduction requires for conclusiveness. Unlike deductive syllogisms, empirical conclusions are inferences-inductive conclusions. As such, they are probabilistic statements based on what we observe and measure. But we cannot observe and measure all the processes that may account for the A-B relationship.

Causal inferences are going to be made. Although they are neither permanent nor universal, they allow us to build knowledge of presumed causes overtime. Such empirical conclusions provide us with successive approximations to the truth. Recognizing this caveat, let's look further at the types of causal relationships of interest in business researchers.

Causal Relationships

Our concern in causal analysis is with how one variable affects, or is "responsible for," changes in another variable. The stricter interpretation of causation, found in experimentation, is that some external factor "produces" a change in the dependent variable. In business research, we often find that the cause-effect relationship is less explicit. We are more interested in understanding, explaining, predicting, and controlling relationships between variables than we are in discerning causes.

If we consider the possible relationships that can occur between two variables, we can conclude 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 variable. 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 company camping club is the result of (dependent) another factor, such as a lifestyle preference.

A **reciprocal relationship** exists when two variables mutually influence or reinforce each other. This could occur if the reading of an advertisement leads to the use of a brand of product. The usage, in turn, sensitizes the person to notice and read more of the advertising of that particular brand.

Most research analysts look for **asymmetric relationships**. With these we postulate that changes in one variable (the independent variable, or IV) are responsible for changes in another variable (the dependent variable, or DV). The identification of the IV and DV is often obvious, but sometimes the choice is not clear. In these later cases we evaluate independence and dependence on the basis of:

1. *The degree to which each variable may be altered*. The relatively unalterable variable is the independent variable (IV) (e.g., age, social status, present manufacturing technology).
2. *The time order between the variables.* The independent variable (IV) precedes the dependent variable (DV).

Finally there are four types of asymmetric relationships: stimulus-response, property-disposition, disposition-behavior, and property-behavior. Experiments usually involve stimulus-response relationships. Property-disposition relationships are often studied in business and social science research. Much of ex post facto research involves relationships between properties, dispositions, and behaviors.

Review Questions

1. Distinguish between the following:
   a) Exploratory and Formal studies
   b) Descriptive and Causal studies.
2. Explain the possible relationships that can occur between two variables.
3. In this unit research design has been classified using eight different descriptors. Discuss each of them.
4. Assume you have been asked to determine how hospitals prepare and train volunteers. Since you know relatively little about this subject, how will you find out? Be as specific as possible.
5. You are the administrative assistant for a division chief in a large holding company that owns several hotels. You and the division chief have just come from the CEO's office, where you were informed that the guest complaints related to house keeping and employee attitude are increasing. Your on-site mangers have mentioned some tension among the workers but have not considered it unusual. The CEO and your division chief instruct you to investigate. Suggest at least three different types of research design that might be appropriate in this situation.

Unit 5: Secondary Data Searches

Learning Objectives:

After completion of this unit, you should be able to:

- Identify the purpose and process of exploring different data sources;
- Discuss the three levels of management decision-related secondary sources.
- Explain the five critical factors for evaluating the value of an information source and its contents.
- Explain what is involved in internal data mining.

1. The Exploratory Phase Search Strategy

Most researchers find a review of secondary sources critical to moving from management question to research question. In moving from management question to research question, the researcher uses both internal and external secondary sources. We address external sources first. Our discussion of data mining of internal sources completes the unit.

In this **exploratory research** phase of your research project, your objective is to accomplish the following:

- Expand your understanding of the management dilemma.
- Look for ways others have addressed and/or solved problems similar to your management dilemma or management question.
- Gather background information on your topic to refine the research question.
- Identify information that should be gathered to formulate investigative questions.
- Identify sources for and actual questions that might be used as measurement questions.
- Identify sources and for actual sample frames that might be used in sample design.

In most cases the exploration phase will be begin with a literature search-a review of books as well as articles in you journals or professional literature that relate your management dilemma. A literature search requires the use of the library's online catalog and one or more bibliographic databases or indexes. For some topics, it may be useful to consult a handbook or specialized encyclopedia first to establish a list of key terms, people, or events that have influenced your topic and also to determine what the major publications are and who the foremost authors are. Other reference materials will be incorporated into your search strategy as needed. In general, this **literature search** has five steps:

1. Define your management dilemma or management question.
2. Consult encyclopedias, dictionaries, handbooks, and textbooks to identify key terms, people, or events relevant to your management dilemma or management question.
3. Apply these key terms, people, or events in searching indexes, bibliographies, and the Web to identify specific secondary sources.
4. Locate and review specific secondary sources for relevance.
5. Evaluate the value of each source and its content.

The result of your literature search may be a solution to the management dilemma. In such a case, no further research is necessary. Often, however, the management question remains unresolved, so the decision to proceed generates a research proposal. The resulting proposal covers at minimum a statement of the research question and a brief description of proposed research methodology. The proposal summarizes the findings of the exploratory phase of the research, usually with a bibliography of secondary sources that have led to the decision to propose as formal research study.

In this unit we will concentrate on the exploration phase of the project and focus on finding, selecting, and evaluating information in both printed and electronic formats. The first step in an exploratory study is a search of the secondary literature. We defined secondary literature as "studies made by others for their own purposes." These studies, representing primary research to their authors, actually represent only a subset of all the information sources available.

1.1. Levels of Information

As you explore your problem or topic, you may consider many different types of information sources, some much more valuable than others. Information sources are generally categorized into three levels: (1) primary sources, (2) secondary sources, and (3) tertiary sources.

Primary sources are original works of research or raw data without interpretation or pronouncements that represent an official opinion or position. Included among the primary sources are memos, letters, complete interviews or speeches (in audio, video, or written transcript formats), laws, regulations, court decisions or standards, and most government data, including census, economic, and labor data. Primary sources are always the most authoritative because the information has not been filtered or interpreted by a second party. Information from all of the above will become your secondary literature supporting *your* original research. Internal sources of primary data would include inventory records, personnel records, purchasing requisition forms, statistical process control charts, and similar data.

Secondary sources are interpretations of primary data. Encyclopedias, textbooks, handbooks, magazine and newspaper articles, and most newscasts are considered secondary information sources. Indeed, nearly all reference materials fall into this category. Internally, sales analysis summaries and inventory annual reports would be examples of secondary sources as they are complied from a variety of primary sources. To an outsider, however, the annual report is viewed as a primary source, as it represents the official position of the corporation. A firm searching for secondary sources can search either internally or externally.

Tertiary sources may be an interpretation of a secondary source but generally are represented by indexes, bibliographies, and other finding aids (e.g., Internet search engines).

From the beginning, it is important to remember that all information is not of equal value. As the source levels indicate, primary sources have more value than secondary sources, and secondary sources have more value than tertiary sources. If the information is essential to solving the management dilemma, it is wise to verify it in a primary source.

1.2. Types of Information Sources

There are dozens of types of information sources, each with a special function. In this section we describe five of the information types used most by business researchers.

A. Indexes and Bibliographies

Indexes and **bibliographies** are the mainstay of any library because they help you identify and locate a single book or journal article from among the millions published. The single most important bibliography in any library is its online catalog. As with all other information types, there are many specialized indexes and bibliographies unique to business topics. These can be very useful in a literature search to find authors and titles of prior works on the topic of interest.

B. Dictionaries

Dictionaries are so ubiquitous they probably need no explanation. We all use them to verify spelling or grammar usage or to define terms. In business, as in every field, there are many specialized dictionaries that define words, terms, or jargon unique to a discipline. Most of these specialized dictionaries include in their word lists in the information to people, events, or organizations that shape the discipline. They are also an excellent place to find acronyms. A growing number of dictionaries and glossaries (terms in a specialized field, area, or topic plus their definitions) are now available on the Web. An example of a printed business dictionary is the *Dictionary of Business and Management*. Information from dictionaries and glossaries may be used to identify key terms for a search of an online or printed database.

C. Encyclopedias

Use an **encyclopedia** to find background or historical information on a topic or to find names or terms that can enhance your search results in other sources. For example, you might use an encyclopedia to find when Microsoft introduced Windows, and then use that date to draw more information from an index to the time period. Encyclopedias are also helpful in identifying the experts in a field and the key writings on a any topic.

D. Handbooks

A **handbook** is a collection of facts unique to a topic. Handbooks often include statistics, directory information, a glossary of terms, and other data such as laws and regulations essential to a field. The best handbooks include source references for the facts they present.

E. Directories

Directories are used for finding names and addresses as well as other data. While many are available and useful imprinted format, directories in digitized format that can be searched by certain characteristics or sorted and then downloaded are far more useful. Many are available through the Web, but the most comprehensive directories are proprietary (that is, must be purchased).

1.3. Evaluating Information Sources

As you begin to collect information about your topic, one thing you will certainly want to do is conduct a **source evaluation**. Librarians evaluate and select information sources based on five factors that can be applied to any type of source, whether printed or electronic. These are:

A. Purpose

The **purpose** of the sources is what the author (or in the case of many Internet sites, the collective authors in an institution) is trying to accomplish. In general, the purpose may be to enlighten or to entertain. Among purposes in the enlighten subset, authors may be attempting to establish credibility, broaden knowledge within a field or

discipline, establish a company image, and for Web sites, even manage inventory or sell merchandise. Often for proprietary sources, the purpose is to make the search process easier for those who buy. Or the purpose may be to define basic computer terms so that non technical mangers can communicate with technical staff. Once you determine the purpose of the source, you will also want to determine whether or how it provides a bias to the information presented. Bias is the absence of a balanced presentation of information. Most researchers expect company Web sites to be biased in favor of the company. However, we expect sources offered by independent organizations to be more balanced, presenting both positive and negative information about relevant organizations without favoring one or the other.

B. Scope

Tied closely to the purpose of the source is its **scope**. What is the date of publication? What time period does this source covers? How much of the topic is covered and to what depth? Is the material covered local, regional, national, or international? If the source is bibliography, what are the criteria for inclusion? If you do not know the scope of your information sources, you may miss essential information by relying on an incomplete source.

This single factor led to tragic results in a university clinical research project in 2001. The medical researcher combed the literature online, found no problems, and proceeds to offer a drug to a health volunteer who then died of complications. The salient reports on the ill effects the drug had been published in the early 1960s and was too old to be included in the online database used by the researcher.

C. Authority

Of major concern to any information user is the **authority** of the source. We have already noted that primary sources are the most authoritative. In any source, both the author and the publisher are indicators of the authority. The author and the author's credentials should be given both imprinted and electronic sources. Footnotes should be provided when appropriate. If credentials are not given, then it is best to check a

bibliographic source. Credential may include the author's educational background, his/her position, or his/her other published and reviewed work.

Authority also applies to Web resources where anyone can post anything. In this environment it is always important to check the credentials of the site. For instance, data and statements about cancer are much more likely to be authoritative if they come from the National Cancer Institute than from a personal page with no information about the author or producer. Any personal page on the Web is suspect unless the author's, or in some cases the institution's, credentials can be verified. Credentials alone are not always enough. Most scholarly articles are validated by a system called *peer review*, in which colleagues from other institutions are asked to comment anonymously on the research presented. In several recent cases authors with very good credentials have bypassed the peer review process and published directly on the Web. In at least one instance the research was seriously flawed. Even the best scholars can make mistakes, so imagine what nonexperts might publish on the Web.

D. Audience

Audience is also an important factor in evaluating an information source, and it, too, is tied to the purpose of the source. The audience for this module is college students. More specifically, its audience is college students who are studying or majoring in accounting and finance, some of whom are practicing accountants. While others, for example educators may benefit from the information, the author take great care to select examples and write in terms that accounting and finance students will easily relate to.

It is often difficult to determine the intended audience for some Web resources, as many Web sites are designed to serve multiple audiences. The Web is available to all, so Web designers need to be especially creative both in capturing audience(s) and in moving users to the appropriate section(s) of their Web site. While an organization may have several different print publications geared to different users, all the users on the Web go through one home page or portal.

When you are evaluating the plausible audience of a source, look for key indicators including vocabulary, types of information, and questions or directions that guide the search.

E. Format

Format factors may vary from source to source but in general relate to how the information is presented and how easy it is to find a specific piece of information. In a printed source, is the arrangement of the information-alphabetical? Hierarchical? Chronological?-nearly always has an impact on the retrieval of information. Indexes are usually essential. Do cross-references link one term to related terms? How are acronyms handled? Is the reference to an item? Table number? Page? How do type fonts or color help you find information?

The format for an electronic source or Web site is generally related to design. Web designers contend with a variety of Internet browsers, individual computer "preferences," and a wide variety of modem or Internet access speeds. On the one hand, the designer wants the page to look great and have the coolest bells and whistles. On the other hand, the designer has to plan for users who are not very patient. People with visual impairments, using text only or specialized software that "reads" the page, must be taken into consideration, too. A page that takes even 30 seconds to load may be abandoned while the user goes on to something else. Users can't flip through pages on the Web nearly so quickly or effectively as they can flip through a reference book. Navigation becomes an issue. If you are holding a book, you know how to get back to the title page. If you are using a Web site, you want to be able to go forward one screen, backward one screen, return to the home page, search the site from any page, and quickly find the home page. On every page within a site, you should be able to see the name of the site owner and the last time the page was update and, preferably, should be able to contact the author or site manager.

Activity 1

1. What could be the possible categorization of information sources?
2. Write at least four types of information sources that you know.
3. Assume that you are evaluating the sources of information in data collection. Hence, what could be your factors (criteria) for selecting a source among the available alternatives?

2. Mining Internal Sources

The term **data mining** describes the process of discovering knowledge from databases stored in data marts or data warehouses. The purpose of data mining is to identify valid, novel, useful, and ultimately understandable patterns in data. Similar to traditional mining, where we search beneath the surface for valuable ore, data mining searches large databases for indispensable information for managing an organization. Both require sifting a large amount of material to discover a profitable vein. Data mining is a useful tool, an approach that combines exploration and discovery with confirmatory analysis.

An organization's own internal historical data are often an underutilized source of information in the exploratory phase. Due to employee turnover, the researcher may lack knowledge that such historical data explicit; or, based on time or budget constraints and the lack of an organized archive, the researcher may choose to ignore such data. While digging through data archives can be as simplistic as sorting through a file containing past patient records or inventory shipping manifests, or rereading company reports and management-authored memos that have grown dusty with age, we will concentrate the remainder of our discussion on more sophisticated structures and techniques.

A **data warehouse** is an electronic repository for databases that organizes large volumes of data into categories to facilitate retrieval, interpretation, and sorting by end users. The data warehouse provides an accessible archive to support dynamic organizational intelligence applications. The key words here are *dynamically accessible*. Data

warehouses that offer archaic methods for data retrieval are seldom used. Data in a data warehouse must be continually updated to ensure that managers have access to data appropriate for real-time decisions. In a data warehouse, the contents of departmental computers are duplicated in a central repository where standard architecture and consistent data definitions are applied. Theses data are available to departments or cross-functional teams for direct analysis or through intermediate storage facilities or **data marts** that compile locally required information. The entire system must be constructed for integration and compatibility among the different data marts.

The more accessible the databases that comprise the data warehouse, the more likely the researcher will use such databases to reveal patterns. Thus researchers are more likely to mine electronic databases than paper ones. It will be useful to remember that data in a data warehouse were once primary data, collected for a specific purpose. When researcher's data mine a company's data warehouse, all the data contained within that databases have become secondary data. The patterns revealed will be used for purposes other than those originally intended. For example, in an archive of sales invoices, we have a wealth of data about what was sold, how much of each item or service, at what price level, to whom, and where and when and how the product were shipped. Initially the company generated the sales invoice to facilitate the process of getting paid for the items shipped. When a researcher mines that sales invoice archive, the search is for patterns of the sales, by product, category, region of the country or world, price level, shipping methods, etc. Therefore, data mining forms a bridge between primary and secondary data.

Traditional databases queries are unidimensional and historical-for example, "How much beer was sold during December in Tourist Hotel?" In contrast, data mining attempts to discover patterns and trends in the data and to infer rules from these patterns. For example, analysis of retail sales by Tourist Hotel identified products that are often purchased together –like beer and diapers-although they may appear to be unrelated. With the rules discovered from the data mining a manager is able to support, review,

and/or examine alternative courses of action for solving a management dilemma, alternatives that may later be studied further in the collection of new primary data.

Activity 2

1. Define the terms data mining and data warehousing.
2. Write what could be done in data mining process.

Data Mining Process

Data mining involves a five step process:

- *Sample*: Decide between census and sample data.
- *Explore*: Identify relationships within the data.
- *Modify*: Modify or transform data.
- *Model*: Develop a model that explains the data relationships.
- *Assess*: Test the model's accuracy.

Sample

The researcher must decide whether to use the entire data set or a sample of the data. If the data set in question is not large, if the processing power is high, or if it is important to understand patterns for every record in the database, sampling should not be done. However, if the data warehouse is very large (terabytes of data), the processing power is limited, or speed is more important than complete analysis, it is wise to draw a sample. In some instances, researchers may use a data mart for their sample-with local data that are appropriate for their geography. Alternatively, the researcher may select an appropriate sampling technique. Since fast turnaround for decisions is often more important than absolute accuracy, sampling is appropriate.

If general patterns exist in the data as a whole, these patterns will be found in the sample. If a niche is so tiny that it is not represented in a sample yet is so important that it influences the big picture, it will be found using explanatory data analysis (EDA).

Explore

After the data are sampled, the next step is to explore them visually or numerically for trends or groups. Both visual and statistical exploration (data visualization) can be used to identify trends. The researcher also looks for outliers to see if the data needed to be cleaned, cases needed to be dropped, or a larger sample needs to be drawn.

Modify

Based on the discoveries in the exploration phase, the data may require modification. Clustering, fractal-based transformation, and the application of fuzzy logic are completed during this phase as appropriate. A data reduction program, such as factor analysis, correspondence analysis, or clustering, may be used. If important constructs are discovered, new factors may be introduced to categorize the data into these groups. In addition, variables based on combinations of existing variables may be added, recorded, transformed, or dropped.

At times, descriptive segmentation of the data is all that is required to answer the investigative question. However, if a complete predictive model is needed, the researcher will move to the next step of the process.

Model

Once the data are prepared, construction of a model begins. Modeling techniques in data mining include neural networks as well as decision tree, sequence-based, classification and estimation, and generic-based models.

Assess

The final step in data mining is to assess the model to estimate how well it performs. A common method of assessment involves applying a portion of data that was not used during the sampling stage. If the model is valid, it will work for this "holdout" sample. Another way to test a model is to run the model against known data. For example, if you know which customers in a file have high loyalty and your model predicts loyalty, you can check to see whether the model has selected these customers accurately.

Review Questions

1. Explain how each of the five evaluation factors for a secondary source influences its management decision-making value:
   a) Purpose
   b) Scope
   c) Authority
   d) Audience
   e) Format
2. Define the distinctions between primary, secondary, and tertiary sources in a secondary search.
3. What is data mining?
4. Some researchers find that their sole sources are secondary data. Why did this happen?
5. What problems of secondary data quality must researchers face? How can they deal with them?
6. Assume you are asked to investigate the use of mathematical programming in accounting applications. You decide to depend on secondary data sources. What type of searching tools might you use? Which tool do you think would be the most fruitful? Sketch a flow diagram of your search sequence.

Unit 6: Qualitative Research

Learning Objectives:

After completion of this unit, you should be able to:

- Identify how qualitative methods differ from quantitative methods;
- Discuss the controversy surrounding qualitative research;
- Define what quantitative research methodologies are.
- Distinguish the types of decisions that use qualitative method; and
- Describe the variety qualitative research methods.

1. What is Qualitative Research?

Researchers basically do research to understand how and why things happen. If the manager needs to know what happened, or how often things happened, quantitative research methodologies would serve the purpose. But to understand the different meanings that people place on their experiences often requires research techniques that delve more deeply into people's hidden interpretations, understandings, and motivations. Qualitative research is designed to tell the researcher how (process) and why (meaning) things happen as they do.

Qualitative research includes an "array of interpretative techniques which seek to describe, decode, translate, and otherwise come to terms with the meaning, not the frequency, of certain more or less naturally occurring phenomena in the social world." Qualitative techniques are used at both the data collection and data analysis stages of a research project. At the data collection stage, the array of techniques includes focus groups; individual depth interviews (IDIs), case studies, ethnography, ground theory, action research, and observation. During analysis, the qualitative researcher uses content analysis of written or recorded materials drawn from personal expressions by

participants, behavioral observations, and debriefing of observers, as well as the study of artifacts and trace evidence from the physical environment.

Qualitative research aims to achieve an in-depth understanding of a situation. Judith Langer, a noted qualitative researcher, indicates that qualitative research is ideal if you want to extract feelings, emotions, motivations, perceptions, consumer "language," or self-described behavior.

Qualitative research draws data from a variety of sources, including the following:

- People (individuals or groups).
- Organizations or institutions.
- Texts (published, including virtual ones).
- Settings and environments (visual/sensory and virtual material).
- Objects, artifacts, media products (textual/visual/sensory and virtual material).
- Events and happenings (textual/visual/sensory and virtual material).

In this unit we will focus on the qualitative methods that draw data from people and organizations.

2. Qualitative versus Quantitative Research

2.1. The controversy

Qualitative research methodologies have roots in a variety of disciplines, including anthropology, sociology, psychology, linguistics, communication, economics, and semiotics. Historically, qualitative research methodologies have been a variable much longer-some as early as the 19th century-than the quantitative tools marketers rely on so heavily. Possibly because of their origins, qualitative methods don't enjoy the unqualified endorsement of upper management. Many senior managers maintain qualitative data are too subjective and susceptible to human error and bias in data collection and interpretation. They believe such research provides an unstable foundation for expensive

and critical business decision. The fact that results cannot be generalized from qualitative study to a larger population is considered a fundamental weakness.

Increasingly, however, managers are returning to these techniques as quantitative techniques fall short of providing the insights needed to make those ever-more-expensive business decisions. Managers deal with the issue of trustworthiness of qualitative data through exacting methodology:

- Carefully using literature searches to build probing questions.
- Thoroughly justifying the methodology or combination of methodologies chosen.
- Executing the chosen methodology in its natural setting (field study) rather than a highly controlled setting (laboratory).
- Choosing sample participants for relevance to the breadth of the issue rather than how well they represent the target population.
- Developing and including questions that reveal the exceptions to a rule or a theory.
- Carefully structuring the data analysis.
- Comparing data across multiple sources and different contexts.
- Conducting peer-researcher debriefing on results for added clarity, additional insights, and reduced bias.

2.2. The Distinction

To understand the distinction between qualitative and quantitative methodologies, let's define the later. **Quantitative research** attempts precise measurement of something. In business research, quantitative methodologies usually measure consumer behavior, knowledge, opinions, or attitudes. Such methodologies answer questions related to how much, how often, how many, when, and who. While the survey is not the only methodology of the quantitative researcher, it is considered a dominant one.

The purpose of qualitative research is based on "researcher immersion in the phenomenon to be studied, gathering data which provide a detailed description of events, situations and interaction between people and things, [thus] provide depth and detail."

Quantitative research is often used for theory testing (Will a Birr-off instant coupon or a Birr 1.50 mail-in rebate generate more sales for National Lottery Administration?), requiring that the researcher maintain a distance from the research so as not to bias the results. Qualitative research-sometimes labeled *interpretative research* because it seeks to develop understanding through detailed description-often builds theory but rarely tests it. Besides the purpose of the research, this process sets up several key distinctions between qualitative and quantitative research, including level of researcher involvement; sampling methodology and size; data collection processes, including participant preparation and researcher and research sponsor involvement; data type and preparation; data analysis and timing; processes for reaching insight and meaning; time frame of insight discovery; and the level of data security.

Unlike the case with quantitative data, both the researcher and the research sponsor often have more significant involvement in collecting and interpreting qualitative data. The researcher may serve as a participant or a catalyst, as a participant observer, or as a group interview moderator. The research sponsor may observe (in some cases via Webcast of interviews directly to the sponsor's desktop computer), influence interview questions, and add interpretations and insights during the process. By contrast, with large quantitative studies, the researcher who interprets the data and draws conclusions from it is rarely the data collector and often has no contact at all with the participant.

Since researchers are immersed in the participant's world, any knowledge they gain can be used to adjust the data extracted from the next participant. In quantitative research, identical data are desired from all participants, so evolution of methodology is not acceptable.

Quantitative data often consists of participant response that are coded, categorized, and reduced to numbers so that these data may be manipulated for statistical analysis. One objective is the quantitative tally of events or opinions, called *frequency of response*. Qualitative data are all about texts. Detailed description of events, situations, and interactions, either verbal or visual, constitute the data. Data may be contained within

transcriptions of interviews or video focus groups, as well as in notes taken during those interactions. But by definition they generate reams of words that need to be coded and analyzed by humans for meaning. While computer software is increasingly used for the coding process in qualitative research, at the heart of the qualitative process is the researcher-and his/her experience-framing and interpreting the data.

Qualitative studies with their smaller sample sizes offer an opportunity for faster turnaround of findings. While speed should never be the primary reason for choosing a methodology, qualitative data may be especially useful to support a low-risk decision that must be made quickly.

Multimillion Birr marketing strategies may lose their market persuasiveness if the competitor reacts too quickly. Data security is therefore of increasing concern. Both group and individual interviewing, the mainstay techniques of qualitative research, can be conducted in highly secure environments. In comparison, once a quantitative survey or field observation or experiment is started, it is quickly common knowledge among a research sponsor's competitors. While the data may not be known, the area of inquiry often can be determined.

3. Qualitative Research Methodologies

The researcher choose a qualitative methodology based on the project's purpose; its schedule; including the speed with insights are needed; its budget; the issue(s) or topic(s) being studied; the types of participants needed; and the researcher's skill, personality, and preferences.

3.1. Sampling

Sample sizes for qualitative research vary by techniques but are generally small. A study might include just two or three focus groups or a few dozed individual depth interviews. Qualitative research involves **nonprobability sampling**-where little attempt is made to generate a representative sample. Several types of nonprobability sampling are common:

- *Purposive sampling*: researchers choose participants arbitrarily for their unique characteristics or their experiences, attitudes, or perceptions; as conceptual or theoretical categories of participants develop during the interviewing process, researchers seek new participants to challenge emerging patterns.
- *Snowball sampling*: participants refer researchers to others who have characteristics, experiences, or attitudes similar to or different from their own.
- *Convenience sampling*: researchers select any readily available individuals as participants.

3.2. Interviews

The **interview** is the primary data collection technique for gathering data in qualitative methodologies. Interviews vary based on the number of people involved during the interview, the level of structure, the proximity of the interviewer to the participant, and the number of interviews conducted during the research.

An interview can be conducted individually (individual depth interview, or IDI) or in groups. Both have a distinct place in qualitative research.

The researcher chooses either an **unstructured interview** (no specific questions or order of topics to be discussed, with each interview customized to each participant; generally starts with a participant narrative) or a **semistructure interview** (generally starts with a few specific questions and follows the individual's tangents of thought with interviewer probes) or a **structured interview** (often uses a detailed interview guide similar to a questionnaire to guide the question order and the specific way the questions are asked, but the questions generally remain open-ended). Structured interviews permit more direct comparability of responses; question variability has been eliminated and thus answer variability is assumed to be real. Also, in the structure interview, the interview's neutrality has been maintained.

Most qualitative research relies on the unstructured or semistructured interview. The unstructured or semistructured interviews used in qualitative research are distinct from the structured interview in several ways. They:

- Rely on developing a dialog between interviewer and participant.
- Require more interviewer creativity.
- Use the skill of the interviewer to extract more and a greater variety of data.
- Use interviewer experience and skill to achieve greater clarity and elaboration of answers.

Many interviews are conducted face-to-face, with the obvious benefit of being able to observe and record nonverbal as well as verbal behavior. An interview, however, can be conducted by phone or online. Phone and online interviews offer the opportunity to conduct more interviews within the same time frame and draw participants from a wider geographic area. These approaches also save the travel expenses of moving trained interviewers to participants, as well as the travel fees associated with bringing participant to a neutral site. Using interviewers who are fresher and more comfortable in conducting an interview-often from their home or office-should increase the quality of the interview. Also, depending on the group from which participants are drawn, there may be insufficient numbers of conduct group interviews in any one market, forcing the use of phone or online techniques.

Interviewer Qualifications

Interviewing requires a trained interviewer (often called a **moderator** for group interviews) or the skill gained from experience. Theses skills including making respondents comfortable, probing for detail without making the respondent feel harassed, remaining neutral while encouraging the participant to talk openly, listening carefully, following a participant's train of thought, and extracting insights from hours of detailed descriptive dialogue. Skilled interviewers learn to use their personal similarities with *or* differences from their interviewee to mine for information; similarities are used to convey sympathy and understanding, while differences are used to demonstrate eagerness to understand and emphasize.

In quantitative research we are more interested in the data collector's following a prescribed procedure. Whereas in qualitative research, the individual conducting the interview needs a fuller understanding of the dilemma and how the insights will be used. So a skilled interviewer must be a "quick-study," someone who can grasp an understanding of an issue without necessarily having prior experience with the product or service or being a technical expert.

The interviewer needs to be able to extract information from a willing participant who often is not consciously aware that he/she possesses the information desired. The actual interviewer is usually responsible for generating the **interview** or **discussion guide**, the list of topics to be discussed (unstructured interview) or the questions to be asked (semistructured) and in what order (structured). In building this guide, many interviewers employ a hierarchical questioning structure. Broader questions starts the interview, designed to put participants at ease and give them a sense that they have a lot to contribute, followed by increasingly more specific questions to draw out detail.

The interviewer is often responsible for generating the screening questions used to recruit participants for the qualitative research. This preinterview uses a device similar to a questionnaire, called a recruitment screener. Each question is designed to reassure the researcher that the person, who has the necessary information and experiences, as well as the social and language skills to relate the desired information, is invited to participate. Data gathered during the recruitment process are incorporated into the adapt analysis phase of the research, as recruitment data provide additional context for participants' expressions.

In general, then, the interviewer is a consultant with wide-ranging responsibilities:

- Recommends the topics and questions.
- Controls the interview, but also plans-and may manage-the locations and facilities for the study.
- Proposes the criteria for drawing the sample participants.
- Writes the recruitment screener and may recruit participants.

- Develops the various pretasking exercises.
- Prepares any research tools (e.g., picture sorts or written exercises) to be used during the interview.
- Supervises the transcription process.
- Helps analyze the data and draw insights.
- Writes or directs the writing of the client report, including extracting video clips for the oral report.

Individual Depth Interviews

An **individual depth interview** (**IDI**) is an interaction between an individual interviewer and a single participant. Individual depth interviews generally take between 20 minutes (telephone interviews) and 2 hours (prescheduled, face-to-face interviews) to complete, depending on the issues or topics of interest and the contact method used. Some techniques such as *life histories* may take as long as five hours. Participants are usually paid to share their insights and ideas.

Group Interviews

A **group interview** is a data collection method using a single interviewer with more than one research participant. Group interviews can be described by the group's size or its composition.

Group interviews vary widely in size: *dyads* (2 people), *triads* (3 people), *mini-groups* (2 to 6 people), *small groups* (focus groups-6 to 10 people-unarguably the most well known of group interview techniques), or *supergroups* (up to 20 people). The smaller groups are usually used when the overall population from which the participants are drawn is small, when the topic or concept list is extensive or technical, or when the research calls for greater intimacy. Dyads also are used when the special nature of a friendship or other relationship (e.g., spouses, superior-subordinate, and siblings) is needed to stimulate frank discussion on a sensitive topic. Dyads and triads are also used frequently with young children who have lower levels of articulation or more limited attention spans and are thus more difficult to control in large groups. A supergroup is used when a wide

range of ideas is needed in a short period of time and when the researcher is willing to sacrifice a significant amount of participant interaction for speed.

In terms of composition, groups can be **heterogeneous** (consisting of different individuals; commonality of opinions, backgrounds, actions) or **homogeneous** (consisting of similar individuals; commonality of opinions, backgrounds, actions). Groups also can comprise **experts** (individuals exceptionally knowledgeable about the issues to be discussed) or **nonexperts** (those who have at least some desired information but at an unknown level).

Driven by the belief that the data extracted will be richer because of the interaction, group interviews are one of the few research techniques in which the participants are encouraged to interact. However, given time constraints, group interviews permit spending only limited time extracting detail from each participant. This problem is magnified when a group interview is structured to cover numerous questions or topics. Another drawback of the group interview is the increased difficulty recruiting, arranging, and coordinating group discussions. But this aggravation-which can be subcontracted to a specialist research supplier-is deemed a small price to pay for the insights that often are revealed by group interaction.

Activity 1

1. Distinguish between qualitative and quantitative research.
2. While conducting an interview what could be the responsibilities of the interviewer?
3. Classify group interviews in terms of composition of the group?

4. Combining Qualitative Methodologies

4.1. Case Study

The **case study**, also referred to as the *case history*, is a powerful research methodology that combines individual and (sometimes) group interviews with record analysis and

observation. Researchers extract information from company brochures, annual reports, sales receipts, and newspaper and magazine articles, along with direct observation (usually done in the participant's "natural" setting), and combine with it with interview data from participants. The objective is to obtain multiple perspectives of a single organization, situation, event, or process at a point in time or over a period of time. Case study methodology-or the written report from such a research project, often called a *case analysis* or *case write-up*- can be used to understand particular marketing processes. For example, one study might evaluate new product development processes for similarities, especially the use of outside consultants, ideation techniques, and computer simulation. Another study might examine in detail the purchaser's response to a marketing stimulus like a display. The results of the research could be used to experiment with modifications of the new product development process or with display selection and placement processes to generate higher-value transactions. The research problem is usually a how and why problem, resulting in a descriptive or explanatory study.

Researchers select the specific organizations or situations to profile because these examples or subjects offer critical, extreme, or unusual cases. Researchers most often choose multiple subjects, rather than a single subject, to study because of the opportunity for cross-case analysis. In studying multiple subjects, a deeper understanding of the subject emerges. When multiple units are chosen, it is because they offer similar results for predictable reasons (literal replication) or contrary results for predictable reasons (theoretical replication). While theoretical sampling seems to be common, a minimum of 4 cases with a maximum of 15 seems to be favored.

In the case study, interview participants are invited to tell the story of their experience, with those chosen representing different levels within the same organization or different perspectives of the same situation or process to permit depth of perspective. The flexibility of the case study approach and the emphasis on understanding the context of the subject being studied allow for a richness of understanding sometimes labeled *thick description*.

During analysis, a single case analysis is always performed before any cross-case analysis is conducted. The emphasis is on what differences occur, why, and with what effect. Perspective inferences about best practices are concluded after completing case studies on several organizations or situations and are speculative in nature.

Students, you are quite familiar with studying cases as a means of learning accounting principles.

4.2. Action Research

Accountants conduct research in order to provide insights for managers to make decisions in specific scenarios. **Action research** is designed to address complex, practical problems about which little is known-thus no known heuristics exist. So the scenario is studied; a corrective action is determined, planned and implemented; the results of the action are observed and recorded; and the action is assessed as effective or not. The process is repeated until a desired outcome is reached, but along the way much is learned about the processes and about the prescriptive actions being studied. Action researches investigate the effects of applied solutions. Whatever theories are developed is validated through practical application.

Suppose a restaurant that had never received a customer complaint earns its first challenge by a disgruntled dinner. If no general rule existed about how to treat unhappy patrons, the organization could study the situation and come up with alternative actions. It might:

- Ignore the problem. (Its lack of experience would prevent it from knowing that negative word of mouth-negative buzz-would be the likely result.)
- Do whatever is necessary to replace the unsatisfactory meal within the shortest period of time.
- Accept the current circumstance as uncorrectable; apologize to the customer a free meal to get him/her back in the restaurant another day.

In action research, one of those alternatives would be chosen and implemented, and then the results recorded. Was the customer happy when he/she left? Did the customer return to dine another evening or never return again? Over the next three months, what was the

customer's full revenue value? If the customer didn't return, the next time disgruntled customer voice dissatisfaction a different action would be chosen, implemented, and then assessed in comparison to the first option's results.

5. Quantitative Research Methodologies

Quantitative research is the systematic and scientific investigation of quantitative properties and phenomena and their relationships. The objective of quantitative research is to develop and employ mathematical models, theories and hypotheses pertaining to natural phenomena. The process of measurement is central to quantitative research because it provides the fundamental connection between empirical observation and mathematical expression of an attribute.

It usually starts with a theory or a general statement proposing a general relationship between variables. With this approach it is likely that the researchers will take an objective position and their approach will be to treat phenomena as hard and real. As a result proponents of such studies claim that quantitative research is undertaken in a value free framework. Quantitative researchers favor methods such as surveys and experiments, and will attempt to test hypotheses or statements with a view to infer from the particular to the general. This approach typically concentrates on measuring or counting and involves collecting and analyzing numerical data and applying statistical tests.

6. Merging Qualitative and Quantitative Methodologies

Triangulation is the term used to describe the combing of several qualitative methods or combining qualitative with quantitative methods. Because of the controversy described earlier, qualitative studies may be combined with quantitative ones to increase the perceived quality of the research, especially when a quantitative study follows a qualitative one and provides validation for the qualitative findings. Four strategies for combining methodologies are common in accounting research:

1. Qualitative and quantitative studies can be conducted simultaneously.

2. A qualitative study can be ongoing while multiple waves of quantitative studies are done, measuring changes in behavior and attitudes over time.
3. A qualitative study can precede a quantitative study, and a second qualitative study them might follow the quantitative study, seeking more clarification.
4. A quantitative study can precede a qualitative study.

An example of the first strategy would be the combination of a public opinion poll at the time focus groups are being held to discover ways to sway a particular public's opinion. For the second strategy, we might collect life histories while multiple waves of questionnaires are measuring the response to differing promotional tactics. For the third, we could perform a qualitative study to identify people's behaviors and perceptions with respect to furniture shopping processes and interior decorating; then we could use that information to develop a quantitative study to measure the actual frequency of behaviors and attitudes. And, fourth, we might survey people's behavior and attitudes toward a brand and find we have some IDIs to explain findings that are unclear. Many researchers recognize that qualitative research compensates for the weaknesses of quantitative research and vice versa. These forward thinkers believe that the methodologies complement rather than rival each other.

Review Questions

1. How does qualitative research differ from quantitative research?
2. How do data from qualitative research differ from data in quantitative research?
3. Discuss the main theme of quantitative research methodology.
4. Why do senior executives feel more comfortable relying on quantitative data than qualitative data? How might a qualitative research company lessen the senior-level executive's skepticism?
5. Distinguish between structured, semistructured, and unstructured interviews.
6. Discuss what triangulation means.

Unit 7: Surveys

Learning Objectives:

After completion of this unit, you should be able to:

- Describe the process for selecting the appropriate and optimal communication approach;
- Identify what factors affect participation in communication studies; and
- Differentiate the major sources of error in communication studies and how to minimize them.

1. Characteristics of the Communication Approach

Research designs can be classified by the approach used to gather primary data. There are two alternatives. We can observe conditions, behavior, events, people, or processes. Or we can communicate with people about various topics, including participant's attitudes, motivations, intentions, and expectations. The researcher determines the appropriate data collection approach largely by identifying the types of information needed-investigative questions the researcher must answer. Information about past events is often available only through surveying or interviewing people who remember the events. Thus, the choice of a communication versus an observation approach may seem an obvious one, given the directions in which investigative questions may lead. The characteristics of the sample unit-specifically, whether a participant can articulate his/her ideas, thoughts, and experiences-also play a role in the decision. The researcher's choice of a communication approach affects the following:

- The creation and selection of the measurement questions.
- Instrument design, which incorporates attempts to reduce error and create participant-screening procedure.
- Sampling issues, which drive contact and callback procedures.

- Data collections procedures, which create the need for follow-up procedures (when self-administered instruments are used) and possible interviewer training (when personal or telephone survey methods are used).

In this unit we focus on the choices the researcher must make once the communication approach has been chosen. We discuss the characteristics and applications of the various communication approaches as well as their individual strengths and weaknesses.

The **communication approach** involves surveying or interviewing people and recording their responses for analysis. **A survey** is a measurement process used to collect information during a highly structured interview-sometimes with a human interviewer and other times without. Questions are carefully chosen or crafted, sequenced, and precisely asked of each participant. The goal of the survey is to derive comparable data across subsets of the chosen sample so that similarities and differences can be found. When combined with statistical probability sampling for selecting participants, survey findings and conclusions are projectable to large and diverse populations.

The great strength of the survey as a primary data-collecting approach is its versatility. Abstract information of all types can be gathered by questioning others. Additionally, a few well-chosen questions can yield information that would take much more time and effort to gather by observation. A survey that uses the telephone, mail, a computer, e-mail, or the Internet as the medium of communication can expend geographic coverage at a fraction of the cost and time required by observation. The bad news for communication research is all communication research has some error. Understanding the various sources of error helps researchers avoid or diminish such error.

1.1. Error in Communication Research

There are three major sources of error in communication research: measurement questions and survey instruments, interviewers, and participants. Researchers cannot help a business decision maker answer a research question if they (1) select or craft inappropriate questions, (2) ask them in an appropriate order, or(3) use inappropriate transitions to elicit information.

A. Measurement Questions and Survey Instruments error

The design of survey questions is influenced by the need to relate each question to the others in the instrument. Often the content of one question assumes other questions have been asked and answered. The psychological order of the questions is also important; question sequence can encourage or discourage commitment and promote or hinder the development of research-participant rapport. The basic principle used to guide sequence decisions is this: The nature and needs of the participant must determine the sequence of questions and the organization of the interview schedule.

A defective instrument can cause distortion in two major ways. First, it can be too confusing and ambiguous. The use of complex words and syntax beyond particular comprehension is typical. Leading questions, ambiguous meanings, mechanical defects (inadequate space for replies, response-choice omissions, and poor printing), and multiple questions suggest the range of problems. Many of these problems are the direct result of operational definitions that are insufficient, resulting in an inappropriate scale being chosen or developed.

B. Interviewer error

From the introduction to the conclusion of the interview, there are many points where the interviewer's control of the process can affect the quality of the data. **Interviewer error**, a major source of sampling error and response bias, is caused by numerous actions:

- *Failure to secure full participant cooperation (sampling error)*: the sample is likely to be biased if interviewers do not do a good job of enlisting participant cooperation.
- *Failure to record answers accurately and completely (data entry error)*: error may result from an interview recording procedure that forces the interviewer to summarize or interpret participant answers or that provides insufficient space to record verbatim answers as provided by the participant.
- *Failure to consistently execute interview procedures*: the precision of the survey estimates will be reduced and there will be more error around estimates to the extent that interviewers are inconsistent in ways that influence the data.

- *Failure to establish appropriate interview environment*: answers may be systematically inaccurate or biased when interviewers fail to appropriately train and motivate participants or fail to establish a suitable interpersonal setting.
- *Falsification of individual answers or whole interviews*: perhaps the most insidious form of interviewer error is cheating. Surveying is difficult work, often done by part-time employees, usually with only limited training and under little direct supervision. At times, falsification of an answer to an overlooked question is perceived as an easy solution to counterbalance the incomplete data. This easy, seemingly harmless first step can be followed by more pervasive forgery. It is not known how much of this occurs, but it should be of constant concern to research directors as they develop their data collection design and to those organizations that outsource survey projects.
- *Inappropriate influencing behavior*: it is also obvious that an interviewer can distort the results of any survey by inappropriate suggestions, directions, or verbal probes; by word emphasis and question rephrasing; by tone of voice; or by body language, facial interaction to an answer, or other nonverbal signals. These activities, whether intentional or merely due to carelessness, are widespread. This problem was investigated using a simple questionnaire and participants who then reported on the interviewers. The conclusion was, "The high frequency of deviations from instructed behavior is alarming."
- *Physical presence bias*: interviewers can influence participants in unperceived subtle ways. Older interviewers are often seen as authority figures by young participants, who modify their responses accordingly. Some research indicates that perceived social distance between interviewer and participant has a distorting effect, although the studies do not fully agree on just what their relationship is.

In light of the numerous studies on the various aspects of interview bias, the safest course for a researcher is to recognize the constant potential for response error.

C. Participant Error

Three broad conditions must be met by participants to have a successful survey:

- The participant must posses the information being targeted by the investigative questions.
- The participant must understand his/her role in the interview as the provider of accurate information.
- The participant must have adequate motivation to cooperate.

Thus participants cause error in two ways: whether they respond (willingness) and how they respond.

Participation-Based Errors Three factors influence participation:

- The participant must believe that the experience will be pleasant and satisfying.
- The participant must believe that answering the survey is an important and worthwhile use of his /her time.
- The participant must dismiss any mental reservations that he/she might have about participation.

Whether the experience will be pleasant and satisfying depends heavily on the interviewer in personal and telephone surveys. Typically, participants will cooperate with an interviewer whose behavior reveals confidence and who engages people on personal level. Effective interviewers are differentiated not by demographic characteristics but by these interpersonal skills. By confidence, we mean that most participants are immediately convinced they will want to participate in the study and cooperate fully with the interviewer. An engaging personal style is one where the interviewer instantly establishes credibility by adapting to the individual needs of the participant. For the survey that does not employ human interpersonal influence, convincing the participant that the experience will be enjoyable is the task of a prior notification device or the study's written introduction.

For the participant to think that answering the survey is important and worthwhile, some explanation of the study's purpose is necessary, although the amount of disclosure will vary based on the sponsor's objectives. In personal or phone surveys the researcher will provide the interviewer with instructions for discovering what explanation is needed and supplying it. Usually, the interviewer states the purpose of the study, tells how the

information will be used, and suggest what is expected of the participant. Participants should feel that their cooperation will be meaningful to themselves and to the survey results. When this is achieved, more participants will express their views willingly.
The quality and the quantity of information secured depend heavily on the ability and willingness of participants to cooperate. Potential participant often have reservations about being interviewed that must be overcome. They may suspect the interviewer has an illegitimate purpose. They may view the topic as too sensitive and thus the interview as potentially embarrassing or intrusive. Or they may feel inadequate or fear the questioning will belittle them. Previous encounters with businesses that have attempted to disguise their sales pitch or fund-raising activities as a research survey can also erode participant's willingness to cooperate. In personal and phone interviews, participants often react more to their feelings about the interviewer than to the content of the questions.

At the core of a survey or interview is an interaction between two people or between a person and a questionnaire. In the interaction of the participant is asked to provide information. While he/she ahs hope of some minimal personal reward-in the form of compensation for participation or enhanced status or knowledge-he/she has little hope of receiving immediate or direct benefit from the data extracted. Thus participant motivation is a responsibility of the researcher and the interviewer. Studies of reactions to many surveys show that participants can be motivated to participate in personal and phone interviews and, in fact, can even enjoy the experience. In one study, more 90 percent of participants said the interview experience was interesting, and 3/4th reported they were willing to be interviewed again. In intercept/self-administered studies, the interviewer's primary role is to encourage participation as the participant completes the questionnaire on his/her own.

By failing to respond or refusing to respond, participants create a nonrepresentative sample for the study overall or for a particular item or question in the study. In surveys, **nonresponse error** occurs when the response of participants differ in some systematic way from the responses of nonparticipants. This occurs when the researcher (1) cannot locate the person (the predesignated sample element) to be studied or (2) is successful in

encouraging the person to participate. This is an especially difficult problem when you are using a probability sample of subjects. Many studies have shown that better-educated individuals and those more interested in the topic participate in the surveys. A high percentage of those who reply to a given survey have usually replied to others, while large shares of those who don't respond are habitual nonparticipants.

Response-Based Errors **Response error** is generated in two ways: when the participant fails to give a correct answer or fails to give the complete answer. The interviewer can do little about the participant's information level. Screening questions qualify participants when there is doubt about their ability to answer. The most appropriate applications for communication research are those where participants are uniquely qualified to provide the desired information. Questions can be used to inquire about characteristics of a participant, such as his/her household income, age, sexual preference, ethnicity, or family lifecycle stage. Questions can also be asked that reveal information exclusively internal to the participant.

If we ask participants to report on events that they have not personally experienced, we need to assess the replies carefully. If our purpose is to learn what the participant understands to be the case, it is legitimate to accept the answers given. But if our intent is to learn what the event or situation actually was, we must recognize that the participant is reporting secondhand data and the accuracy of the information declines.

Participants also cause error by responding in such a way as to unconsciously or consciously misrepresent their actual behavior, attitudes preferences, motivation, or intentions (*response bias*). Participants create response bias when they modify their responses to be socially acceptable or to save face or reputation with the interviewer (*social desirability bias*), and sometimes even in an attempt to appear rational and logical.

One major cause of response bias is *acquiescence*-the tendency to be agreeable. On the participant's part, acquiescence may be a result of lower cognitive skill or knowledge related to a concept or construct, language difficulties, or perceived level of anonymity.

However, researchers can contribute to acquiescence by the speed with which they ask questions (the faster questions are asked, the more acquiescence) and the placement of questions in an interview (the later the question, the more acquiescence.)

Sometimes participants may not have an opinion on the topic of concern. Under this circumstance, their proper response should be "don't know" or "have no opinion." Some research suggests that most participants who choose the don't-know response option actually possess the knowledge or opinion that the researcher seeks. Participant may choose the option because they may want to shorten the time spent in the participation process, may be ambivalent or have conflicting opinions on the topic, may feel they have insufficient information to form a judgment-even though they actually have taken a position- don't believe that the response choices match their position, or don't possess the cognitive skills to understand the response options. If they choose the don't know option for any of these reasons, studies suggest that probing for their true position will increase both reliability and validity of the data. However, forcing a participant to express some opinion he/she does not hold by withholding a don't-know option makes it difficult for researchers to know the reliability of the answers.

Participants may also interpret a question or concept differently from what was intended by the researcher. This occurs when the researcher uses words that are unfamiliar to the participant. Thus, the individual answers a question that is different from the one the researcher intended to ask.

Regardless of the reasons, each source of participant-initiated error diminishes the value of the data collected. It is difficult for a researcher to identify such occasions. Thus, communicated responses should be accepted for what they are-statements by individuals that reflect varying degrees of truth and accuracy.

1.2. Choosing a Communication Method

Once the researcher had determined that surveying or interviewing is the appropriate data collection approach, various means may be used to secure information from individuals. A researcher can conduct a semistructured interview or survey by personal interview or

telephone or can distribute a self-administered survey by mail, fax, computer, e-mail, the Internet, or a combination of these. While there are commonalities among these approaches, several considerations are unique to each.

2. Self-Administered Surveys

The self-administered questionnaire is ubiquitous in modern living. Often a short questionnaire is left to be completed by the participant in a convenient location or is packaged with a product. Self-administered **mail surveys** are delivered not only by Ethiopian Postal Service but also via fax and courier service. Other delivery modalities include *computer-delivered* and *intercept* studies.

3. Survey via Telephone Interview

The telephone survey is still the workhorse of the survey research. Numerous firms field phone omnibus studies each week. Individual questions in these studies are used to capture everything from people's feeling about the rise in gasoline prices to the power of a celebrity spokesperson in an advertising campaign or the latest fashion trend. Of the advantages that the telephone interviewing offers, probably none ranks higher than its moderate cost. One study reports that sampling and data collection costs for telephone surveys can run from 45 to 64 percent lower than costs for comparable personal interviews. Much of savings comes from cuts in travel costs and administrative savings from training and supervision. When calls are made from a single location, the researcher may use fewer, yet more skilled, interviewers. Telephones are especially economical when callbacks to maintain precise sampling requirements are necessary and participants are widely scattered. Long distance service options make it possible to interview nationally at a reasonable cost.

4. Survey via Personal Interview

A survey via personal interview is a two-way conversation between a trained interviewer and a participant. There are real advantages as well as clear limitations to surveys via personal interview. The greatest value lies in the depth of information and detail that can

be secured. It far exceeds the information secured from telephone and self-administered studies via mail or computer (both intranet and Internet). The interviewer can also do more things to improve the quality of the information received than is possible with another method.

Activity 1

1. What do we mean when we say survey in research studies?
2. Write the three major sources of error in communication research.
3. Compare and contrast Survey via telephone interview with and survey via personal interview.

5. Selecting an Optimal Survey Method

The choice of a communication method is not as complicated as it might first appear. By comparing your research objectives with the strengths and weaknesses of each method, you will be able to choose one that is suited to your needs.

When your investigative questions call for information from hard-to-reach or inaccessible participants, the telephone interview, mail survey, or computer-delivered survey should be considered. However, if data must be collected very quickly, the mail survey would likely be ruled out because of lack of control over the returns. Alternatively, you may decide your objective requires extensive questioning and probing; then the survey via personal interview should be considered.

If none of the choices turns out to be a particularly good fit, it is possible to combine the best characteristics of two or more alternatives into a *hybrid* survey. Although this decision will incur the costs of the combined modes, the flexibility of tailoring a method to your unique needs is often an acceptable trade-off.

Ultimately all researchers are confronted by the practical realities of cost and deadlines. On the average, surveys via personal interview are the most expensive communication

method and take the most field time unless a large field team is used. Telephone surveys are moderate in cost and offer the quickest option. Questionnaires administered by e-mail or the Internet are the least expensive. When your desired sample is available via the Internet, the Internet survey may prove to be the least expensive communication method with the most rapid (simultaneously) data availability. The use of the computer to select participants and reduce coding and processing time will continue to improve the cost-to-performance profiles of this method in the future.

Most of the time, an optimal method will be apparent. However, managers' needs for information often exceed their internal resources. Such factors as specialized expertise, large field team, unique facilities, or a rapid turnaround prompt managers to seek assistance from research vendors of survey-related services.

Review Questions

1. Distinguish among response error, interviewer error, and nonresponse error.
2. How do environmental factors affect response rates in personal interviews? How can we overcome these environmental problems?
3. Assume you are planning to interview shoppers in a shopping mall about their views on increased food prices and what the federal government should do about them. In what different ways might you try to motivate shoppers to cooperate in your survey?
4. You decide to take a telephone survey of 40 families in the 881-exchange area. You want an excellent representation of all subscribers in the exchange area. Explain how you will carry out this study.

Unit 8: Sampling

Learning Objectives:

After completion of this unit, you should be able to:

- Differentiate census inquiry and sample survey;
- Identify the steps in sampling design;
- List the criteria for selecting a sampling procedure;
- Describe the characteristics of a good sample design; and
- Discuss the different types of sample designs.

1. Census and Sample Survey

All items in any field of inquiry constitute a 'Universe' or 'Population'. A complete enumeration of all items in the 'population' is known as a census inquiry. It can be presumed that in such an inquiry, when all items are covered, no element of chance is left and highest accuracy is obtained. But in practice this may not be true. Even the slightest element of bias in such an inquiry will get larger and larger as the number of observation increases. Moreover, there is no way of checking the element of bias or its extent except through a resurvey or use of sample checks. Besides, this type of inquiry involves a great deal of time money and energy. Therefore, when the field of inquiry is large, this method becomes difficult to adopt because of the resources involved. At times, this method is practically beyond the reach of ordinary researchers. Perhaps, government is the only institution which can get the complete enumeration carried out. Even the government adopts this in vary rare cases such as population census conducted once in a decade. Further, many a time it is not possible to examine every item in the population, and sometimes it possible to obtain sufficiently accurate results by studying only a part of total population. In such cases there is no utility of census surveys.

However, it needs to be emphasized that when the universe is a small one, it is no use resorting to a sample survey. When field studies are undertaken in a practical life,

considerations of time and cost almost invariably lead to a selection of respondents i.e., selection of only a few items. The respondents selected should be a representative of the total population as possible in order to produce a miniature cross-section. The selected respondents constitute what is technically called a 'sample' and the selection process is called 'sampling technique'. The survey so conducted is known as 'sample survey'. Algebraically, let the population size be ***N*** and if a part of the size ***n*** (which is <***N***) of this population is selected according to some rule for studying some characteristic of the population, the group consisting of these ***n*** units is known as 'sample'. Researcher must prepare a sample design for his/her study i.e., he/she must plan how a sample should be selected and of what size such a sample would be.

Implications of a sample design

A sample design is a definite plan for obtaining a sample from a given population. It refers to the technique or the procedure the researcher would adopt in selecting items for the sample. Sample design may as well lay down the number of items to be included in the sample i.e., the size of the sample. Sample design is determined before data are collected. There are many sample designs from which a researcher can choose. Some designs are relatively more precise and easier to apply than others. Researcher must select/ prepare a sample design which should be reliable and appropriate for his/her research study.

2. Steps in a Sampling Design

While developing sample design, the researcher must pay attention to the following points:

i. **Type of universe**: the first step in developing any sample design is to clearly define the set of objects, technically called the Universe, to be studied. The universe can be finite or infinite. In finite universe the number of items is certain, but in case of an infinite universe the number of items is infinite, i.e., we can not have any idea about the total number of items. The population of a city, the number of workers in a factory and the like are examples of finite universes,

whereas the number of stars in the sky, listeners of a specific radio program, throwing of a dice etc. are examples of infinite universes.

ii. **Sampling unit**: a decision has to be taken concerning a sampling unit before selecting sample. Sampling unit may be a geographical one such as state, district, village, etc., or a construction unit such as house, flat etc., or it may be a social unit such as family, club, school etc., or it may be an individual. The researcher will have to decide one or more of such units that he/she has to select for the study.

iii. **Source list**: it is also known as 'sampling frame' from which sample is to be drawn. It contains the names of all items of a universe (in case of finite universe only). If source list is not available, researcher has to prepare it. Such a list should be comprehensive, correct, reliable and appropriate. It is extremely important for the source list to be as representative of the population as possible.

iv. **Size of sample**: this refers to the number of items to be selected from the universe to constitute a sample. This is a major problem before a researcher. The size of sample should neither be excessively large, nor too small. It should be optimum. An optimum sample is one which fulfills the requirements of efficiency, representativeness, reliability and flexibility. While deciding the size of sample, researcher must determine the desired precision as also an acceptable confidence level for the estimate. The size of population variance needs to be considered as in case of larger variance usually a bigger sample is needed. The size of population must be kept in view for this also limits the sample size. The parameters of interest in a research study must be kept in view, while deciding the size of the sample. Costs too dictate the size of the sample that we can draw. As such, budgetary constraint must invariably be taken into consideration when we decide the sample size.

v. **Parameters of interest**: in determining the sample design, one must consider the question of the specific population parameters which are of interest. For instance, we may be interested in estimating the proportion of persons with some characteristic in the population, or we may be interested in knowing some average or the other measure concerning the population. There may also be important sub-

groups in the population about whom we would like to make estimates. All this has a strong impact upon the sample design we would accept.

vi. **Budgetary constraint**: cost considerations, from practical point of view, have a major impact upon decisions relating to not only the size of the sample but also to the type of sample. This fact can even lead to the use of a non-probability sample.

vii. **Sampling procedure**: finally, the researcher must decide the type of sample he/she will use i.e., he/she must decide about the techniques to be used in selecting the items for the sample. In fact, this technique or procedure stands for the sample design itself. There are several sample deigns (explained in the pages that follow) out of which the researcher must choose one for his/her study. Obviously, he must select that design which, for a given sample size and for a given cost, has a smaller sampling error.

3. Criteria for Selecting a Sampling Procedure

In this context one must remember that two costs are involved in a sampling analysis viz., the cost of collecting the data and the cost of an incorrect inference results from the data. Researcher must keep in view the two causes of incorrect inferences viz., systematic bias and sampling error. Systematic bias results from errors in the sampling procedures, and it cannot be reduced or eliminated by increasing the sample size. At best the causes responsible for these errors can be detected and corrected. Usually a systematic bias is the result of one or more of the following factors:

1. **Inappropriate sampling frame**: if the sampling frame is inappropriate i.e., a biased representation of the universe, it will result in a systematic bias.
2. **Defective measuring device**: if the measuring device is constantly in error, it will result in systematic bias. In survey work, systematic bias can result if the questionnaire or the interviewer is biased. Similarly, if the physical measuring device is defective there will be systematic bias in the data collected through such a measuring device.
3. **Non-respondents**: if we are unable to sample all the individuals initially include in the sample, there may arise a systematic bias. The reason is that in such a situation the likelihood of establishing contact or receiving a response

from an individual is often correlated with the measure of what is to be estimated.

4. **Indeterminacy principle**: sometimes we find that individuals act differently when kept under observation than what they do when kept in non-observed situations. For instance, if workers are aware that somebody is observing them in course of a work study on the basis of which the average length of time to complete a task will be determined and accordingly the quota will be set for piece work, they generally tend to work slowly in comparison to the speed with which they work if kept unobserved. Thus, the indeterminacy principle may also be a cause of a systematic bias.
5. **Natural bias in the reporting of data:** Natural bias of respondents in the reporting of data is often the cause of a systematic bias in many inquiries. There is usually a downward bias in the income data collected by government taxation department, where as we find an upward bias in the income data collected by some social organization. People generally understate their incomes if asked about it for tax purpose, but they overstate the same if asked for social status or their affluence. Generally in physiological surveys, people tend to give what they think is the 'correct' answer rather than revealing their true feeling.

Sampling errors are the random variations in the sample estimates around the true population parameters. Since they occur randomly and are equally likely to be in either direction, their nature happens to be of compensatory type and the expected value of such errors happens to be equal to zero. Sampling error decreases with the increase in the size of the sample, and it happens to be of a smaller magnitude in case of homogeneous population.

Sampling error can be measured for a given sample design and size. The measurement of sampling error is usually called the 'precision of the sampling plan'. If we increase the sample size, the precision can be improved. But increasing the size of the sample has its own limitations viz., a large sized sample increases the cost of collecting data and also enhances the systematic bias. Thus the effective way to increase precision is usually to select a better sampling design which has a smaller sampling error fro a given sample

size at a given cost. In practice, however, people prefer a less precise design because it is easier to adopt the same and also because of the fact that systematic bias can be controlled in a better way in such a design.

In brief, while selecting a sampling procedure, researcher must ensure that the procedure causes a relatively small sampling error and helps to control the systematic bias in a better way.

4. Characteristics of Good Sample Design

From what has been stated above, we can list down the characteristics of a good sample design as under:

a) Sample design must result in a truly representative sample.
b) Sample design must be such which results in a small sampling error.
c) Sampling design must be viable in the context of funds available for the research study
d) Sample design must be such so that systematic bias can be controlled in a better way
e) Sample should be such that the results of the sample study can be applied, in general, fro the universe with a rescannable level of confidence.

Activity 1

1. Differentiate between census and sample survey.
2. Could you rehearse the criteria for selecting a sampling procedure?
3. What are the characteristics of good sampling design?

5. Different Type of Sample Designs

There are different types of sample designs based on two factors viz., the representation basis and the element selection technique. On the **representation basis**, the sample may be probability sampling or it may be non-probability sampling. Probability sampling is based on the concept of random selection, whereas non-probability sampling is 'non-random' sampling. On **element selection basis**, the sample may be either unrestricted or

restricted. When each sample element is drawn individually from the population at large, then the sample so drawn is known as 'unrestricted sample', whereas all other forms of sampling are covered under the term 'restricted sampling'. The following chart exhibits the sample designs as explained above.

CHART SHOWING BASIC SAMPLING DESIGNS

Element selection technique ↓	Representation basis: Probability sampling	Representation basis: Non-probability sampling
Unrestricted sampling	Simple random sampling	Haphazard sampling or convenience sampling
Restricted sampling	Complex random sampling (such as cluster sampling, systematic sampling, stratified sampling etc.)	Purposive sampling (such as quota sampling, judgment sampling)

Thus, sampling designs are basically of two types' viz., non-probability sampling and probability sampling. We take up these two designs separately.

5.1. Non-probability Sampling

Non-probability sampling is that sampling procedure which does not afford any basis for estimating the probability that each item in the population has of being included in the sample. Non-probability sampling is also known by different names such as deliberate sampling, purposive sampling and judgment sampling. In this type of sampling, items for the sample are selected deliberately by the researcher; his choice concerning the items remains supreme. In other words, under non-probability sampling the organizers of the inquiry purposively choose the particular units of the universe for constituting a sample on the basis that the small mass that they so select out of a huge one will be typically or representative the whole. For instance, if economic conditions of people living in a state

are to be studied, few towns and villages may be purposively selected for intensive study on the principle that they can be representative of the entire state. Thus, the judgment of the organizers of the study plays an important part in this sampling design.

In such a design, personal element has a great chance of entering into the selection of the sample. The investigator may select a sample which shall yield results favorable to his point of view and if that happens, the entire inquiry may get vitiated. Thus, there is always the danger of bias entering into this type of sampling technique. But if the investigators are impartial, work without bias and have the necessary experience so as to take sound judgment, the results obtained from an analysis of deliberately selected sample may be tolerably reliable. However, in such a sampling, there is no assurance that every element has some specifiable chance of being included. Sampling error in this type of sampling cannot be estimated and the element of bias, great or small is always there. As such this sampling design is rarely adopted in large inquiries of importance. However, in small inquiries and researches by individuals, this design may be adopted because of the relative advantage of time and money inherent in this method of sampling.

Quota sampling is also an example of non-probability sampling. Under quota sampling the interviews are simply given quotas to be filled from the different strata, which some restrictions on how they are to be filled. In other words, the actual selection of the items for the sample is left to the interview's discretion. This type of sampling is very convenient and relatively inexpensive. But the samples so selected certainly do not possess the characteristic of random samples. Quota samples are essentially judgment samples and inferences drawn on their basis are not amenable top statistical treatment in a formal way.

5.2. Probability Sampling

Probability sampling is also known as 'random sampling' or 'chance sampling'. Under this sampling design, every item of the universe has an equal chance of inclusion in the sample. It is, so to say, a lottery method in which individual units are picked up from the whole group not deliberately but by some mechanical process. Here it is blind chance

alone that determines whether one item or the other is selected. The results obtained from probability or random sampling can be assured in terms of probability i.e., we can measure the errors of estimation or the significance of results obtained from a random sample, and this fact brings out the superiority of random sampling design. Random sampling ensures the law of Statistical Regularity which states that if on an average the sample chosen is a random one, the sample will have the same composition and characteristics as the universe. This is the reason why random sampling is considered as the best technique of selecting a representative sample.

Random sampling from a finite population refers to that method of sample selection which gives each possible sample combination an equal probability of being picked up and each item in the entire population to have an equal chance being included in the sample. This applies to sampling without replacement i.e., once an item is selected for the sample, it can not appear in the sample again (Sampling with replacement is used less frequently in which procedure the element selected for the sample is returned to the population before the next element is selected. In such a situation the same element could appear twice in the same sample before the second element is chosen). in brief, the implications of random sampling (or simple random sampling) are:

a) It gives each element in the population an equal probability of getting into the sample; and all choices are independent of one another.
b) It gives each possible sample combination an equal probability of being chosen.

Keeping this in view we can define a simple random sample (or simply a random sample) from a finite population as a sample which is chosen in such a way that each of the ${}^{N}C_{n}$ possible samples has the same probability, $1/ {}^{N}C_{n,}$ of being selected. To make it more clear we take a certain finite population consisting of six elements (say a, b, c, d, e, f) i.e., N= 6. Suppose that we want to take a sample of size n=3 from it. Then there are ${}^{6}C_{3}=20$ possible distinct samples of the required size, and they consist of the elements *abc, abd, abe, abf, acd, ace, acf, ade, adf, aef, bcd, bce, bcf, bde, bdf, bef, cde, cdf, cef and def*. if we choose one of these samples in such a way that each has the probability 1/20 of being chosen, we will then call this a random sample.

5.3. How to select a Random Sample?

With regard to the question of how to take a random sample in actual practice, we could, in simple cases like the one above, write each of the possible samples on a slip of paper, mix these slips thoroughly in a container and then draw as a lottery either blindfolded or by rotating a drum or by any other similar device. Such a procedure is obviously impractical, if not altogether impossible complex problems of sampling. In fact, the practical utility of such a method is very much limited.

Fortunately, we can take a random sample in a relatively easier way without taking the trouble of enlisting all possible samples on paper-slips as explained above. In stead of this, we can write the name of each element of a finite population a slip of paper, put the slips of paper so prepared into a box or a bag and mix them thoroughly and then draw (without looking) the require number of slips for the sample one after the other without replacement. In doing so we must make sure that in successive drawings each of the remaining elements of the population has the same chance of being selected. This procedure will also result in the same probability for each possible sample. We can verify this by taking the above example. Since we have a finite population of 6 elements and we want to select a sample of size 3, the probability of drawing any one element for our sample in the first draw is 3/6, the probability of drawing one more element in the second draw is 2/5, (the first element drawn is not replaced) and similarly the probability of drawing one more element in the third draw is ¼. Since these draws are independent, the joint probability of the three elements which constitute our sample is the product of their individual probabilities and this works out to 3/6 x 2/5 x ¼ = 1/20. This verifies our earlier calculation.

Even this relatively easy method of obtaining a random sample can be simplified in actual practice by the use of random number tables. Various statisticians like Tippet, Yates, and Fisher have prepared tables of random numbers which can be used for selecting a random sample.

5.4. Random Sample from an Infinite Universe

So far we have talked about random sampling, keeping in view only the finite populations, but what about random sampling in the context of infinite populations? It is relatively difficult to explain the concept of random sample from an infinite population. However, a few examples will show the basic characteristics of such a sample. Suppose we consider the 20 throws of a fair dice as a sample from the hypothetical infinite population which consists of the results of all possible throws of the dice. If the probability of getting a particular number, say 1, is the same for each throw and the 20 throws are all independent, then we say that the sample is random. Similarly, it would be said to be sampling from an infinite population if we sample with replacement from a finite population and our sample would be considered as a random sample if in each draw all elements of the population have the same probability of being selected and successive draws happen to be independent. In brief, one can say that the selection of each item in a random sample from an infinite population is controlled by the same probabilities and that successive selections are independent of one another.

Activity 2

1. Distinguish between probability and non-probability sampling designs.
2. On the basis of element selection technique, sample designs are categorized into _____and_____.
3. What do we mean when we say simple random sampling?

5.5. Complex Random Sampling Designs

Probability sampling under restricted sampling technique, as stated above, may result in complex random sampling designs. Such designs may as well be called 'mixed sampling designs' for many of such designs may represent a combination of probability and non-probability sampling procedures in selecting a sample. Some of the popular complex random sampling designs are as follows:

A) Systematic sampling: in some instances, the most practical way of sampling is to select every i^{th} item on a list. Sampling of this type is known as systematic sampling. An element of randomness is introduced into this kind of sampling by using random numbers to pick up the unit with which to start. For instance, if a 4 percent sample is desired, the first item would be selected randomly from the first twenty-five and thereafter every 25^{th} item would automatically be included in the sample. Thus, in systematic sampling only the first unit is selected randomly and the remaining units of the sample are selected at fixed intervals. Although a systematic sample is not a random sample in the strict sense of the term, but it is often considered reasonable to treat systematic sample as if it were a random sample.

Systematic sampling has certain plus points. It can be taken as an improvement over a simple random sample in as much as the systematic sample is spread more evenly over the entire population. It is an easier and less costly method of sampling and can be conveniently used even in case of large populations. But there are certain dangers too in using this type of sampling. If there is a hidden periodicity in the population, systematic sampling will prove to be an inefficient method of sampling. For instance, every 25^{th} item produced by a certain production process is defective. If we are to select a 4% sample of the items of this process in a systematic manner, we would either get all defective items in our sample depending upon the random starting position. If all elements of the universe are ordered in a manner representative of the total population, i.e., the population list is in random order, systematic sampling is considered equivalent to random sampling. But if this is not so, then the results of such sampling may, at times, not be very reliable. In practice, systematic sampling is used when lists of population are available and they are of considerable length.

B) Stratified sampling: if a population from which a sample is to be drawn does not constitute a homogeneous group, stratified sampling technique is generally applied in order to obtain a representative sample. Under stratified sampling the population is divided into several subpopulations that are individually more homogenous than the total population (the different sub-populations are called 'strata') and then we select items

from each stratum to constitute a sample. Since each stratum is more homogenous than the total population, we are able to get more precise estimates from each stratum and by estimating more accurately each of the component parts; we get a better estimate of the whole. In brief, we can say that stratified sampling results in more reliable and detailed information.

The following three questions are highly relevant in the context of stratified sampling:

a) How to form strata?
b) How should items be selected from each stratum?
c) How may items be selected from each stratum or how to allocate the sample size to each stratum?

Regarding the first question, we can say that the strata be formed on the basis of common characteristic(s) of the items to be put in each stratum. This means that various strata be formed in such a way as to ensure elements being most homogenous within each stratum and most heterogamous between the different strata. Thus, strata are purposively formed and are usually based on past experience and personal judgment of the researcher. One should always remember that careful consideration of the relationship between the characteristics of the population and the characteristics to be estimated are normal used to define the strata. At times, pilot study may be conducted for determining a more appropriate and efficient stratification plan. We can do so by taking small samples of equal size from each of the proposed strata and then examining the variances within and among the possible stratifications, we can decide an appropriate stratification plan for our inquiry.

In respect of the second question, we can say that the usual method, for selection of items for the sample from each stratum, resorted to is that of simple random sampling. Systematic sampling can be used if it is considered more appropriate in certain situations.

Regarding the third question, we usually follow the method of proportional allocation under which the sizes of the samples from the different strata are kept proportional to the sizes of the strata. That is, if P_i represents the proportion of population included in

stratum *i*, and *n* represents the total sample size, the number of elements selected from stratum *i* is $n * P_i$. To illustrate it, let us suppose that we want a sample of size $n = 30$ to be drawn from a population size of $N = 8000$ which is divided into three strata of size $N_1 = 4000$, $N_2 = 2400$, and $N_3 = 1600$. Adopting proportional allocation, we shall get the same samples as under for the different strata:

For strata with $N_1 = 4000$, we have $P_1 = 4000/8000$

And hence $n_1 = n * P_1 = 30\ (4000/8000) = 15$

Similarly, for strata with $N_2 = 2400$, we have

$n_2 = n * P_2 = 30\ (2400/8000) = 9$, and

For strata with $N_3 = 1600$, we have

$n_3 = n * P_3 = 30\ (1600/8000) = 6$.

Thus, using proportional allocation, the sample sizes for different strata are 15, 9, and 6 respectively which is in proportion to the sizes of the strata viz., 4000: 2400: 1600. Proportional allocation is considered most efficient and an optimal design when the cost of selecting an item is equal for each stratum, there is no difference in within-stratum variances, and the purpose of sampling happens to be to estimate the population value of some characteristic. But in case the purpose happens to be to compare the differences among the strata, then equal sample selection from each stratum would be more efficient even if the strata differ in sizes. In cases where strata differ not only in size but also in variability and it is considered reasonable to take large samples from the more variable strata and smaller samples from the less variable strata, we can then account for both (differences in stratum size and differences in stratum variability) by using disproportionate sampling design by requiring:

$$n_1 / N_1\sigma_1 = n_2 / N_2 Q = \square = n_k / N_k\sigma_k$$

Where σ_1, $\sigma 2$, ….. and σ_k denote the standard deviations of the k strata, N_1, N_2, ….. N_k denote the sizes of the k strata and n_1, n_2, …., n_k denote the sample sizes of k strata. This is called '*optimum allocation*' in the context of disproportionate sampling. The allocation in such a situation results in the following formula for determining the sample sizes for different strata:

$$n_i = \frac{n.N_i\sigma_i}{N_1 o_1 + N_2\sigma_2 + + N_k\sigma_k} \qquad for \;\; i = 1,2,...and\ k.$$

We may illustrate the use of this by an example.

Illustration 1

A population is divided into three so that N_1=5000, N_2 = 2000 and N_3 = 3000. Respective standard deviations are:

σ_1=15, σ_2 = 18 and σ_3 = 5

How should a sample of size n = 84 be allocated to the three strata, if we want optimum allocation using disproportionate sampling design?

Solution: Using the disproportionate sampling design for optimum allocation, the sample sizes for different strata will be determined as under:

Sample size for strata with N1=5000

$$n_1 = \frac{84(5000)(15)}{(5000)(15)+(2000)(18)+(3000)(5)}$$
$$= 6300000 / 126000$$
$$= 50$$

Sample size for strata with N_2 = 2000

$$n_2 = \frac{84(2000)(18)}{(5000)(15)+(2000)(18)+(3000)(5)}$$
$$= 3024000 / 126000$$
$$= 24$$

Sample size for strata with N_3 = 3000

$$n_3 = \frac{84(3000)(5)}{(5000)(15)+(2000)(18)+(3000)(5)}$$
$$= 1260000 / 126000$$
$$= 10$$

In addition to differences in stratum size and differences in stratum variability, we may have differences in stratum sampling cost, and then we can have cost optimal disproportionate sampling design by requiring

$$\frac{n_1}{N_1\sigma_1\sqrt{C_1}} = \frac{n_2}{N_2\sigma_2\sqrt{C_2}} = \ldots = \frac{n_k}{N_k\sigma_k\sqrt{C_k}}$$

Where

C_1 = Cost of sampling in stratum 1

C_2 = Cost of sampling in stratum 2

C_k = Cost of sampling in stratum k

And all other terms remain the same as explained earlier. The allocation in such a situation results in the following formula for determining the sample sizes for different strata:

$$n_i = \frac{n.N_i\sigma_i/\sqrt{C_i}}{N_1\sigma_1\sqrt{C_1} + N_2\sigma_2\sqrt{C_2} + \ldots + N_k\sigma_k\sqrt{C_k}} \quad for\ i = 1,2,\ldots,k$$

It is not necessary that stratification be done keeping in view a single characteristic. Populations are often stratified according to several characteristics. For example, a system-wide survey designed to determine the students toward a new teaching plan, a state college system with 20 colleges might stratify the students with respect to class, sex and college. Stratification of this type is known as *cross-stratification*, and up to a point such stratification increases the reliability of estimates and is much used in opinion surveys

From what has been stated above in respect of stratified sampling, we can say that the sample so constituted is the result of successive application of purposive (involved in stratification of items) and random sampling methods. As such it is an example of mixed sampling. The procedure wherein we first have stratification and then simple random sampling is known as stratified random sampling.

C) Cluster sampling: if the total area of interest happens to be a big one, a convenient way in which a sample can be taken is to divide the area into a number of smaller non-

overlapping area and then to randomly select a number of these smaller areas (usually called clusters), with the ultimate sampling consisting of all (or samples of) units in these small areas or clusters.

Thus in cluster sampling the total population is divided a number of relatively small subdivisions which are themselves clusters of still smaller units and then some of these clusters are randomly selected for inclusion in the overall sample. Suppose we want to estimate the proportion of machine-parts in an inventory which are defective. Also assume that there are 20000 machine parts in the inventory at a given point of time, stored in 400 cases of 50 each. Now using a cluster sampling, we would consider the 400 cases as clusters and randomly select 'n' cases examine all the machine-parts in each randomly selected case.

Cluster sampling, no doubt, reduces cost by concentrating surveys in selected clusters. But certainly it is less precise than random sampling. There is also not as much information in 'n' observations within a cluster as there happens to be in 'n' randomly drawn observations. Cluster sampling is used only because of the economic advantage it possesses; estimates based on cluster samples are usually more reliable per unit cost.

D) Area sampling: If cluster happen to be some geographic subdivision, in that case cluster sampling is better known as area sampling. In other words, cluster designs, where the primary sampling unit represents a cluster of units based on geographic area, are distinguished as area sampling. The plus and minus point of cluster sampling are also applicable to area sampling.

E) Multi-stage sampling: multi-stage sampling is a further development of the principle of cluster sampling. Suppose we want to investigate the working efficiency of commercial banks in Ethiopia and we want to take a sample of few banks for this purpose. The first stage is to select large primary sampling unit such as states in a country. Then we may select certain districts and interview all banks in the chosen districts. This would represent a two stage sampling design with the ultimate sampling units being clusters of districts.

If instead of taking a census of all banks within the selected districts, we select certain towns and interview all banks in the chosen towns. This would represent a three-stage sampling design. If instead of taking a census of all banks within the selected towns, we randomly sample banks from each selected town, then it is a case of using a four-stage sampling plan. If we select randomly at all stages, we will have what is known as 'multi-stage random sampling design'.

Ordinary multi stage sampling is applied in big inquires extending to a considerable large geographical area, say, the entire country. There are two advantages of this sampling design viz.,

a) It is easier to administer than most single stage design mainly because of the fact that sampling frame under multi stage sampling is developed in partial units.
b) A large number of units can be sampled for a given cost under multistage sampling because of sequential clustering, whereas this is not possible in most of simple design.

F) Sampling with probability proportional to size: In case the cluster sampling units do not have the same number of approximately the same number of elements, it is considered appropriate to use a random selection process where the probability of each cluster being included in the sample is proportional to the size of the cluster. For this purpose, we have to list the number of elements in each cluster irrespective of the method of ordering the cluster. Then we must sample systematically the appropriate number of elements from the cumulative totals. The actual numbers selected in this way do not refer to individual elements, but indicate which clusters and how many from the cluster are to be selected by simple random sampling or by systematic sampling. The results of this type of sampling are equivalent with those of a simple random sample and the method is less cumbersome and is also relatively less expensive. We can illustrate this with the help of an example.

Illustration 2

The following are the number of departmental stores in 15 cities: 35, 17, 10, 32, 70, 28, 26, 19, 26, 66, 37, 44, 33, 29 and 28. If we want to select a sample of 10 stores, using cities as clusters and selecting within clusters proportional to size, how many stores from each city should be chosen? (Use starting point of 10).

Solution: let us put the information as under:

City Number	No. of departmental stores	Cumulative total	sample
1	35	35	10
2	17	52	
3	10	62	60
4	32	94	
5	70	164	110 160
6	28	192	
7	26	218	210
8	19	237	
9	26	263	260
10	66	329	310
11	37	366	360
12	44	410	410
13	33	443	
14	29	472	460
15	28	500	

Since in the given problem, we have 500 departmental stores from which we have to select a sample of 10 stores, the appropriate sampling interval is 50. As we have to use the starting point of 10^{*}, so we add successively increments of 50 till 10 numbers have been selected. The numbers, thus obtained are: 10, 60, 110, 160, 210, 260, 310, 360, 410, and 460 which have been shown in the last column of the table against the concerning cumulative totals. From this we can say that two stores should be selected randomly from

city number five and one each from city number 1, 3, 7, 9, 10, 11, 12, and 14. This sample of 10 stores is the sample with probability proportional to size.

G) Sequential Sampling: this sampling design is some what complex sample design. The ultimate size of the sample under this technique is not fixed in advance, but is determined according to mathematical decision rules on the basis of information yielded as survey progresses. This is usually adopted in case of two acceptance sampling plan in context of statistical quality control. When a particular lot is to be accepted or rejected on the basis of a single sample, it is known as single sampling: when the decision is to be taken on the basis of two samples, it is known as double sampling and in case the decision rests on the basis of more than two samples but the number of samples is certain and decided in advance, the sampling is known as multiple sampling. But when the number of samples more than two but it is neither certain nor decided in advance, this type of system is often referred to as sequential sampling. Thus, in brief, we can say that in sequential sampling, one can go on taking samples one after another as long as on desires to do so.

* If the starting point is not mentioned, then the same can randomly be selected.

Review Questions

1. What do you mean by 'Sample Design'? What points should be taken into consideration by a researcher in developing a sample design for his research project.
2. How would you differentiate between simple random sampling and complex random sampling designs? Explain clearly by giving examples.
3. Why probability sampling is generally preferred in comparison to non-probability sampling?
4. Explain the procedure of selecting a simple random sample.
5. Under what circumstances stratified random sampling design is considered appropriate? How would you select such sample? Explain by means of an example.
6. Distinguish between:
   a) Restricted and unrestricted sampling;
   b) Convenience and purposive sampling;
   c) Systematic and stratified sampling;
   d) Cluster and area sampling.
7. Explain and illustrate the procedure of selecting a random sampling.
8. "A systematic bias results from errors in the sampling procedures". What do you mean by such a systematic bias? Describe the important causes responsible for such a bias.

Unit 9: Analysis and Presentation of Data

Learning Objectives:

After completion of this unit, you should be able to:

- Describe the importance of editing the collected raw data;
- Explain how coding is used to assign numbers and other symbols to answer and categorize responses;
- Identify the options for data entry and manipulation;
- Use the exploratory data analysis techniques;
- Describe the contents and technical specifications of research reports;
- Organize your presentation to be informative.

1. Data Preparation and Description

Introduction

Once the data begin to flow, a researcher's attention turns to data analysis. **Data preparation** includes editing, coding, and data entry and is the activity that ensures the accuracy of the data and their conversion from raw form to reduced and classified forms that are more appropriate for the analysis. Preparing a descriptive statistical summary is another preliminary step leading to an understanding of the data collected. It is during this step that data entry errors may be revealed and corrected.

1.1. Editing

The customary first step in analysis is to edit the raw data. Editing detects errors and omissions, corrects them when possible, and certifies that maximum data quality standards are achieved. The editor's purpose is to guarantee that data are:

- Accurate.
- Consistent with the intent of the question and other information in the survey.
- Uniformly entered.

- Complete.
- Arranged to simplify coding and tabulation.

In the following question asked of adults 18 or older, one respondent checked two categories, indicating that he was a retired officer and currently serving on active duty.

Please indicate your current military status:

☒ Active Duty	☒ Reserve	☐ Retired
☐ National Guard	☐ Separated	☐ Never served in the military

The editor's responsibility is to decide which of the responses is consistent with the intent of the question or other information in the survey and most accurate for this individual participant.

A. Field Editing

In large projects, field editing review is a responsibility of the supervisor. It, too, should be done soon after the data have been gathered. During the stress of data collection in a personal interview and paper-and-pencil recording in an observation, the researcher often uses ad hoc abbreviations and special symbols. Soon after the interview, experiment, or observation, the investigator should review the reporting forms. It is difficult to complete what was abbreviated or written in shorthand or noted illegibly if the entry is not caught that day. When entry gaps are present from interviews, a callback should be made rather than guessing what the respondent "probably would have said." Self-interviewing has no place in quality research.

A second important control function of the field supervisor is to validate the field results. This normally means he/she will reinterview some percentage of the respondents, at least on some questions, verifying that they have participated and that the interviewer performed adequately. Many research firms will contact about 10 percent of the respondents in this process of data validation.

B. Central Editing

At this point, the data should get a thorough editing. For a small study, the use of a single editor produces maximum consistency. In large studies, editing tasks should be allocated

so that each editor deals with one entire section. Although the later approach will not identify inconsistencies between answers in different sections, the problem can be handled by identifying questions in different sections that might point to possible inconsistency and having one editor check the data generated by these questions.

Sometimes it is obvious that an entry is incorrect-for example, when data clearly specify time in days (e.g., 13) when it was requested in weeks (you expect a number of 4 or less)-or is entered in the wrong place. When replies are inappropriate or missing, the editor can sometimes detect the proper answer by reviewing the other information in the data set. This practice, however, should be limited to the few cases where it is obvious what the correct answer is. It may be better to contact the respondent for correct information, if time and budget allow. Another alternative is for the editor to strike out the answer if it is in appropriate. Here an editing entry of "no answer" or "unknown" is called for.

Another problem that editing can detect concerns faking an interview that never took place. This "armchair interviewing" is difficult to spot, but the editor is in the best position to do so. One approach is to check responses to open-ended questions. These are most difficult to fake. Distinctive response patterns in other questions will often emerge if data falsification is occurring. To uncover this, the editor must analyze as a set the instruments used by each interviewer.

Here are some useful rules to guide editors in their work:

- Be familiar with instructions given to interviewers and coders.
- Do not destroy, erase, or make illegible the original entry by the interviewer; original entries should remain legible.
- Make all editing entries on an instrument in some distinctive color and in a standardized form.
- Initial all answers changed or supplied.
- Place initials and date of editing on each instrument completed.

1.2. Coding

Coding involves assigning numbers or other symbols to answers so that the responses can be grouped into a limited number of categories. In coding the *categories* are the partitions of data sets of a given variable (for example, if the variable is *gender*, the population are *male* and *female*). *Categorization* is the process of using rules to partition a body of data. Both closed and free-response questions must be coded.

The categorization of data sacrifices some data detail but is necessary foe efficient analysis. Most statistical and banner/table software programs work more efficiently in the *numeric* mode. Instead of entering the word *male* or *female* in response to a question that asks for the identification of one's gender, we would use numeric codes (for example, 0 for male and 1 for female). Numeric coding simplifies the researcher's task in converting a nominal variable, like gender, to a "dummy variable". Statistical software also can use alphanumeric codes, as when we use M and F, or other letters, in combination with numbers and symbols for gender.

A. Coding Rules

Four rules guide the pre-and post coding and categorization of data set. The categories within a single variable should be:

- Appropriate to the research problem and purpose.
- Exhaustive.
- Mutually exclusive.
- Derived from one classification principle.

Researchers address these issues when developing or choosing each specific measurement question. One of the purposes of pilot testing of any measurement instrument is to identify and anticipate categorization issues.

I. Appropriateness

Appropriateness is determined at two levels: (1) the best partitioning of the data for testing hypotheses and showing relationships and (2) the availability of comparison data.

For example, when actual age is obtained (ratio scale), the editor may decide to group data by age ranges to simplify pattern discovery within the data. The number of age groups and breadth of each range, as well as the endpoints in each range, should be determined by comparison data-for example, Ethiopia census age ranges, a customer database that includes age ranges, or the age data available from Ethiopian Herald for making an advertising media buy.

II. Exhaustiveness

Researchers often add an "other" option to a measurement question because they know they cannot anticipate all possible answers. A large number of "other" responses, however, suggest the measurement scale the researcher designed did not anticipate the full range of information. The editor must determine if "other" responses appropriately fit into established categories, if new categories must be added, if "other" data will be ignored, or if some combination of these actions will be taken.

While the exhaustiveness requirement for a single variable may be obvious, a second aspect is less apparent. Does one set of categories-often determined before the data are collected-fully capture all the information in the data? For example, responses to an open-ended question about family economic prospects for the next year may originally be categorized only in terms of being "optimistic" or "pessimistic." It may also be enlightening to classify responses in terms of other concepts such as the precise focus of these expectations (income or jobs) and variations in responses between family heads and others in the family.

III. Mutually exclusive

Another important rule when adding categories or realigning categories is that category components should be mutually exclusive. This standard is met when a specific answer can be placed in one and only one cell in a category set. For example, in a survey, assume that you asked participants for their occupation. One editor's categorization scheme might include (1) professional, (2) managerial, (3) sales, (4) clerical, (5) crafts, (6) operatives, and (7) unemployed. As an editor, how would you code a participant's answer

that specified "sales-person at Gap and full-time student" or may be "elementary teacher and tax preparer"? According to census data, it is becoming common for adults in our society to have more than one job. Here, operational definitions of the occupations categorized as "professional," "managerial," and "sales" should help clarify the situation. But the editor facing this situation also would need to determine how the second-occupation data are handled. One option should be to add a second-occupation field to the data set; another would be to develop distinct codes for each unique multiple-occupation combination.

IV. Single Dimension

The problem of how to handle an occupancy entry like "unemployed sales person" brings up a fourth rule of category design. The need for a category set to follow a single classificatory principle means every option in the category set is defined in terms of one concept or construct. Returning to the occupation example, the person in the study might be both a salesperson and unemployed. The "salesperson" label expresses the concept *occupation type*; the response "unemployed" is another dimension concerned with *current employment status* without regard to the respondent's normal occupation. When a category set encompasses more than one dimension, the editor may choose to split the dimensions and develop an additional data filed; "occupation" now becomes two variables: "occupation type" and "employment status."

B. Missing Data

Missing data are information from a participant or case that is not available for one or more variables of interest. In survey studies, missing data typically occur when participants accidentally skip, refuse to answer, or do not know the answer to an item on the questionnaire. In longitudinal studies, missing data may result from participants dropping out of the study, or being absent for one or more data collection periods. Missing data also occur due to researcher error, corrupted data files, and changes in the research or instrument design after the data were collected from some participants, such as when variables are dropped or added. The strategy for handling missing data consists of a two-step process: the researcher first explores the pattern of missing data to

determine the mechanism for *missingness* (the probability that a value is missing rather than observed) and then selects a missing-data technique.

I. Mechanisms for missing data

In order to select a missing-data technique, the researcher must first determine what caused the data to be missing. There are three basic mechanisms for this: data missing completely at random (MCAR); data missing at random (MAR); and data not missing at random (NMAR). If then probability of missingness for a particular variable is dependent on neither the variable itself nor any other variable in the data set, then data are MCAR. Data are considered MAR if the probability of missingness for a particular variable is dependent on another variable but not itself when other variables are held constant. The practical significance of this distinction is that the proper missing-data technique can be selected that will minimize bias in subsequent analyses. The third type of mechanism, NMAR, occurs when data are not missing completely at random and they are not predictable from other variables in the data set. Data NMAR are considered *nonignorable* and must be treated on an improvised bias.

II. Missing Data Techniques

Three basic types of techniques can be used to salvage data sets with missing data: (1) listwise deletion, (2) pairwise deletion, and (3) replacement of missing values with estimated scores. *Listwise deletion*, or complete case analysis, is perhaps the simplest approach, and is the default option in most statistical packages like SPPSS and SAS. With this method, cases are deleted from the sample if they have missing values on any of the variables in the analysis. Likewise deletion is most appropriate when data are MCAR. In this situation, no bias will be introduced because the subsample of complete cases is essentially a random sample of the original sample. However, if data are MAR but not MCAR, then a bias may be introduced, especially if a large number of cases are deleted. For example, if men were more likely than women to be responsible for missing data on the variable *shopping preference*, then the results would be biased toward women's shopping preferences.

Pairwise deletion, also called available case analysis, assumes that data are MCAR. In the past, this technique was used frequently with linear models that are functions of means, variances, and covariances. Missing values would be estimated using all cases that had data for each variable or pair of variables in the analysis. Today most experts caution against pairwise deletion, and recommend alternative approaches.

The replacement of missing values with estimated values includes a variety of techniques. This option generally assumes that data are MAR, since the missing values on one variable are predicted from observed values on another variable. a common option available on many software packages is the replacement of missing values with a mean or other central tendency score. This is a simple approach, but has the disadvantage of reducing the variability in the original data, which can cause bias. Another option is to use a regression or likelihood-based method. Such techniques are found in specialty software packages and the procedures for using them are beyond the scope of this module.

1.3. Data Entry

Data entry converts information gathered by secondary or primary methods to a medium for viewing and manipulation. Keyboarding remains a mainstay for researchers who need to create a data file immediately and store it in a minimal space on a variety of media. However, researchers have profited from more efficient ways of speeding up the research process, especially from bar coding and optical character and mark recognition.

Even with these time reductions between data collection and analysis, continuing innovations in multimedia technology are being developed by the personal computer business. The capability to integrate visual images, streaming video, audio, and data may soon replace video equipment as the preferred method for recording an experiment, interview, or focus group. A copy of the response data could be extracted for data analysis, but the audio and visual images would remain for later evaluation. Although technology will never replace researcher judgment, it can reduce data-handling errors,

decrease time between data collection and analysis, and help provide more usable information.

Activity 1

1. The editor's purpose is to guarantee that data are_____.
2. Distinguish between field editing and central editing.
3. What are the four rules which guide the pre-and post coding and categorization of data set?
4. Explain how the collected data could be entered?

2. Exploring, Displaying, and Examining Data

2.1. Exploratory Data Analysis

The convenience of data entry via spreadsheet, or the data editor of a statistical program makes its tempting to move directly to statistical analysis. That temptation is even stronger when the data can be entered and viewed in real time. Why waste time finding out if the data confirms the hypothesis that motivated the study? Why not obtain descriptive statistical summaries and then test hypotheses?

Exploratory data analysis is both a data perspective and a set of techniques. In this unit, we will present unique and conventional techniques including graphical and tabular devices to visualize the data. Data visualization is an integral element in the data analysis process and as a necessary step prior to hypothesis testing. In unit 1, we said research conducted scientifically is a puzzle-solving activity as well as an attitude of curiosity, suspicion, and imagination essential to discovery. It is natural, then, that exploration and examination of the data would be an integral part of our data analysis perspective.

In **Exploratory Data Analysis** (**EDA**) the researcher has the flexibility to respond to the patterns revealed in the preliminary analysis of the data. Thus patterns in the collected data guide the data analysis or suggest revisions to the preliminary data analysis plan.

This flexibility is an important attribute of this approach. When the researcher is attempting to prove causation, however, confirmatory data analysis is required. **Confirmatory data analysis** is an analytical process guided by classical statistics inference in its use of significance testing and confidence.

One authority has compared exploratory data analysis to the role of police detectives and other investigators and confirmatory analysis to that of judges and the judicial system. The former are involved in the search for clues and evidence; the latter are preoccupied with evaluating the strength of the evidence that is found. Exploratory data analysis is the first step in the search for evidence, without which confirmatory analysis has nothing to evaluate. Consistent with that analogy, EDA shares a commonality with exploratory designs, not formalized ones. Because it doesn't follow a rigid structure, it is free to take many paths in unraveling the mysteries in the data-to sift the unpredictable from the predictable.

A major contribution of the exploratory approach lies in the emphasis on visual representations and graphical techniques over summary statistics. Summary statistics, as you will see momentarily, may obscure, conceal, or even misrepresent the underlying structure of the data. When numerical summaries are used exclusively and accepted without visual inspection, the selection of confirmatory models may be based on flawed assumptions. For theses reasons, data analysis should begin with visual inspection. After that, it is not possible but also desirable to cycle between exploratory and confirmatory approaches.

A. Frequency Tables, Bar Charts, and Pie Charts

Several useful techniques for displaying data are not new to EDA. They are essential to any examination of the data. For example, **frequency table** is a simple device for arraying data. Ad recall, a nominal variable, describes the ads that participants remembered seeing or hearing without being prompted by the researcher or the measurement instrument. The values and percentages can be more readily understood in this graphic format. But when the variable of interest is measured on an interval-ratio

scale and is one with many potential values, theses techniques are not particularly informative.

B. Histograms

The histogram is a conventional solution for the display of interval-ratio data. **Histograms** are used when it is possible to group the variable's value into intervals. Histograms are constructed with bars (or asterisks) that represent data values, where each value occupies an equal amount of area within the enclosed area. Data analysts find histograms useful for (1) displaying all intervals in a distribution, even those without observed values, and (2) examining the shape of the distribution for skewness, kurtosis, and the modal pattern. When looking a histogram, one might ask: Is there a single hump (a mode)? Are subgroups identifiable when multiple modes are present? Are straggling data values detached from the central concentration?

The value for the average annual purchases variables possessing an underlying order are similarly appropriate for histograms. A histogram would not be used for a nominal variable that has no order to its categories.

C. Pareto Diagrams

Pareto diagrams derive their name a 19^{th}-century Italian economist. In quality management, J.M.Juran first applied this concept by noting that only a vital few defects account for most problems evaluated for quality and that the trivial may explain the rest. Historically, this has come to known as the 80/20 rule-that is, an 80 percent improvement in quality or performance can be expected by eliminating 20 percent of the causes of unacceptable quality or performance.

The Pareto diagram is a bar chart whose percentages sum to 100 percent. The data are derived from a multiple-choice. Single-response scale; a multiple-choice, multiple-response scale; or frequency counts of words (or themes) from content analysis. The respondents' answers are sorted in decreasing importance, with bar height in descending order from left to right. The pictorial array that results reveals the highest concentration of improvement potential in the fewest number of remedies.

2.2. Cross Tabulation

Depending on the management question, we can gain valuable insights by examining the data with cross tabulation. **Cross-tabulation** is a technique fro comparing data from two or more categorical variables such as gender and selection by one's company fro an overseas assignment. Cross-tabulation is used with demographic variables and the study's target variables (operationalized measurement questions). The technique uses tables having rows and columns that correspond to the levels or code values of each variable's categories.

The use of percentages

Percentages serve two purposes in data presentation. First they simplify the data by reducing all numbers to a range from 0 to 100. Second, they translate the data into standard form, with a base of 100, for relative comparisons. In a sampling situation, the number of cases that fall into a category is meaningless unless it is related to some base. Percentages are used by virtually everyone dealing with numbers-but often incorrectly. The following guidelines, if used during analysis, will help to prevent errors in reporting:

- *Averaging percentages*: percentages cannot be averaged unless each is weighted by the size of the group from which it is derived. Thus, a simple average will not suffice; it is necessary to use a weighted average.
- *Use of too large percentages*: this often defeats the purpose of percentages-which is to simplify. A larger percentage is difficult to understand and is confusing. If a 1,000 percent increase is experienced, it is better to describe this as a 10-fold increase.
- *Using too small a base*: percentages hide the base from which they have been computed. A figure of 60 percent when contrasted with 30 percent would appear to suggest a sizable difference. Yet if there are only three cases in the one category and six in the other, the differences would not be as significance as they have been made to appear with percentages.
- *Percentage decreases can never exceed 100 percent*: this is obvious, but this type of mistake occurs frequently. The higher figure should always be used as the base

or denominator. For example, if a price was reduced from Birr 1 to Birr .25, the decrease would be 75 percent (75/100).

Activity 2

1. Distinguish between bar charts and histograms.
2. What is pareto diagram?
3. Rehearse the guidelines that could used during analysis to prevent errors in reporting in relation to the usage of percentages.

3. Reporting Research Findings

Introduction

As part of the research proposal, the sponsor and the researcher agree on what types of reporting will occur both during and at the research project. Depending on the budget fro the project, a formal oral presentation may not be part of the reporting. A research sponsor; however, is sure to require a written report.

3.1. The written Research Report

It may seem unscientific and even unfair, but a poor final report or presentation can destroy a study. Research technicians may ignore the significance of badly reported content, but most readers will be influenced by the quality of the reporting. This fact should prompt researchers to make special efforts to communicate clearly and fully.

The research report contains findings, analysis of findings, interpretations, conclusions, and sometimes recommendations. The researcher is the expert on the topic and knows the specifics in a way no one else can.

Because a written research report is an authoritative one-way communication, it imposes a special obligation for maintaining objectivity. Even if your findings seem to point to an action, you should demonstrate restraint and caution when proposing that course.

Reports may be defined by their degree of formality and design. The formal report follows a well-delineated and relatively long format. This is in contrast to the informal or short report.

3.2. Research Report Components

Research reports, long or short, have a set of identifiable components. Usually headings and subheadings divide the sections. Each report is individual; sections may be dropped or added, condensed or expanded to meet the needs of the audience.

A. Prefatory Items

Prefatory materials do not have direct bearing on the research itself. Instead, they assist the reader in using the research report.

Letter of Transmittal

When the relationship between the researcher and the client is formal, a **letter of transmittal** should be included. This appropriate when a report is for a specific client (e.g., a company officer) and when it is generated for an outside organization. The letter should refer to the authorization for the project and any specific instructions or limitations placed on the study. It should also state the purpose and the scope of the study. For many internal projects, it is not necessary to include a letter of transmittal.

Title page

The title page should include four items: the title of the report, the date, and fro whom and by whom it was prepared. The title should be brief but include the following three elements: (1) the variables included in the study, (2) the type of relationship among the variables, and (3) the population to which the results may be applied. Redundancies such as "A Report of" and "A Discussion of" add length to the title but little else. Single-word titles are also of little value. Here are three acceptable ways to word report titles:

Descriptive study: The five-year Demand Outlook for Consumer Package Goods in Ethiopia

Correlation Study: The relationship between Relative National Inflation Rates and Household Purchases of Brand X in International Markets

Casual study: The Effect of Various Motivational Methods on Retail Sales Associates' Attitudes and Performance

Authorization Letter

When the report is sent to a public organization, it is common to include a letter of authorization showing the authority for undertaking the research. This is especially true for reports to federal and state governments and nonprofit organizations. The letter not only shows who sponsored the research but also delineates the original request.

Executive Summary

An executive summary can serve two purposes. It may be a report in a miniature (sometimes called a *top line report*), covering all the aspects in the body of the report but in abbreviated form. Or it may be a concise summary of the major findings and conclusions, including recommendations. Two pages are generally sufficient for executive summaries. Write this section after the rest of the report is finished. It should not include new information but may require graphics to present a particular conclusion. Expect the summary to contain a high density of significance terms since it is repeating the highlights of the report.

Table of Contents

As a rough guide, any report of several sections that totals more than 6 to 10 pages should have a table of contents. If there are many tables, charts or other exhibits, they should also be listed after the table of contents in a separate table of illustrations.

B. Introduction

The introduction prepares the reader for the report by describing the parts of the project: the problem statement, research objectives and background material. In most projects, the introduction can be taken from the research proposal with minor editing.

Problem statement

The problem statement contains the need for the research project. The problem is usually represented by a management question. It is followed by a more detailed set of objectives.

Research Objectives

The research objectives address the purpose of the project. These objectives may be research questions and associated research questions. In correlation or causal studies, the hypothesis statements are included. Hypotheses are declarative statements describing the relationship between two or more variables. They state clearly the variables of concern, the relationships among them, and the target group being studied. Operational definitions of critical variables should be included.

Background

Background material may be of two types. It may be the preliminary results of exploration from an experience survey, focus group, or another source. Alternatively, it could be secondary data from the literature review. A traditional organization scheme is to think of the concentric circles of a target. Starting with the outside ring, the writer works toward the center. The bull's eye contains the material directly related to the problem.

Previous research, theory or situations that led to the management question are also discussed in this section. The literature should be organized, integrated, and presented in a way that connects it logically to the problem. The background includes definitions, qualifications, and assumptions. It gives the reader the information needed to understand the remainder of the research report.

Background material may be placed before the problem statement or after the research objectives. If it is composed primarily of literature review and related research, it should follow the objectives. If it contains information pertinent to the management problem or

the situation that led to the study, it can be placed before the problem statement (where it is found in many applied studies).

C. Methodology

In short reports and management reports, the methodology should not have a separate section; it should be mentioned in the introduction, and details should be placed in an appendix. However, for a technical report, the methodology is an important section, containing at least five parts.

Sampling Design

The research explicitly defines the target population being studied and the sampling methods used. For example, was this a probability or non-probability sample? If probability, was it simple random or complex random? How were the elements selected? How was the size determined? How much confidence do we have, and how much error was allowed?

Explanations of the sampling methods, uniqueness of the chosen parameters, or other points that need explanation should be covered with brevity. Calculations should be placed in an appendix in stead of in the body of the report.

Research design

The coverage of the design must be adapted to the purpose. In an experimental study, the materials, tests, equipment, control conditions, and other devices should be described. In descriptive or ex post facto designs, it may be sufficient to cover the rationale for using one design in stead of competing alternatives. Even with a sophisticated design, the strengths and weaknesses should be identified and the instrumentation and materials discussed. Copies of materials are placed in an appendix.

Data Collection

This part of the report describes the specifics of gathering the data. Its contents depend on the selected design. Survey work generally uses a team with field and central supervision.

How many people were involved? What was their training? How were they managed? When were the data collected? How much time did it take? What were the conditions in the field? How were irregularities handled? In an experiment, we would want to know about participant assignment to groups, the use of standard procedures and protocols, the administration of tests or observational forms, manipulation of the variables, and so forth. Typically, you would include a discussion on the relevance of secondary data and guided these decisions. Again, detailed materials such as field instructions should be included in an appendix.

Data Analysis

This section summarizes the methods used to analyze the data and describes the data handling, preliminary analysis, statistical tests, computer programs, and other technical information. The rationale for the choice of analysis approaches should be clear. A brief commentary on assumptions and appropriateness of use should be presented.

Limitations

This topic is often handled with ambivalence. Some people wish to ignore the matter, feeling that mentioning limitations detracts from the impact of the study. This attitude is unprofessional and possibly unethical. Others seem to adopt a masochistic approach of detailing everything. The section should be a thoughtful presentation of significant methodology or implementation problems. An evenhanded approach is one of the hallmarks of an honest and competent investigator. All research studies have their limitations, and the sincere investigator recognizes that readers need aid in judging the study's validity.

D. Findings

This is generally the longest section of the report. The objective is to explain the data rather than draw interpretations or conclusions. When quantitative data can be presented, this should be done as simply as possible with charts, graphics, and tables.

The data need not include everything you have collected. The criterion for conclusion is, "Is this material important to the reader understands of the problem and the findings?"

However, make sure to show findings unfavorable to your hypotheses as well as those that support them, as this reinforces the bond of trust that has developed between researcher and sponsor.

It is useful to present findings in numbered paragraphs or to present one finding per page with the quantitative data supporting the findings presented in a small table or chart on the same page. While this arrangement adds to the bulk of the report, it is convenient for the reader.

E. Conclusions

Summary and conclusions

The summary is a brief statement of the essential findings. Sectional summaries may be used if there are many specific findings. These may be combined into an overall summary. In simple descriptive research, a summary may complete the report, because conclusions and recommendations may not be required.

Findings state facts; conclusions represent inferences drawn from the findings. A writer is sometimes reluctant to make conclusions and leaves the task to the reader. Avoid this temptation when possible. As the researcher, you are the one best informed on the factors that critically influence the findings and conclusions. Good researchers do not draw conclusions that go beyond the data related to the study.

Recommendations

Increasingly, researchers are expected to offer ideas for corrective actions. In applied research the recommendations will usually be for managerial action, with the researcher suggesting one or several alternatives that are supported by the findings. Also, researchers may recommend further research initiatives. In basic or pure research, recommendations are often suggestions for further study that broaden or test the understandings of a subject area.

F. Appendices

The appendices are the places for complex tables, statistical tests, supporting documents, copies of forms and questionnaires, detailed descriptions of the methodology, instructions to field workers, and other evidence important for later support. The reader who wishes to learn about the technical aspects of the study and to look at statistical breakdowns will want a complex appendix.

G. Bibliography

The use of secondary data requires a bibliography. Long reports, particularly technical ones, require a bibliography. A bibliography documents the sources used by the writer. Although bibliographies may contain work used from preparing the report.

Activity 3

1. What items are included in the research report as prefatory items?
2. List the parts of a research report to be included in an introduction.
3. What does methodology of a research mean?
4. What is the benefit of including appendices as part of the research report?

3.3. Writing the Report

Students often give inadequate attention to reporting their findings and conclusions. This is unfortunate. A well-presented study will often impress the reader more than a study with greater scientific quality but with a weaker presentation. Judging a report as competently written is often the key first step to a manager's decision to use the findings in decision making and also to consider implementation of the researcher's recommendations. Report-writing skills are especially valuable to the junior executive or researcher who aspires to rise in an organization. A well-written study frequently enhances career prospects.

A. Prewriting Concerns

Before writing, one should ask again, "What is the purpose of this report?" Responding to this question is one way to crystallize the problem.

The second prewriting concern is, "Who will read the report?" Thought should be given to the needs, temperament, and biases of the audience. You should not distort facts to meet these needs and biases but should consider them while developing the presentation. Knowing who will read the report may suggest its appropriate length. Generally, the higher the report goes in an organization, the shorter it should be.

Another consideration is technical background-the gap in subject knowledge between the reader and the writer. The greater the gap, the more difficult it is to convey the full findings meaningfully and concisely.

The third prewriting question is, "What are the circumstances and limitations under which I am writing?" Is the nature of the subject highly technical? Do you need statistics? Charts? What is the importance of the topic? A crucial subject justifies more effort than a minor one. What should be the scope of the report? How much time is available? Deadlines often impose limitations on the report.

Finally, "How will the report be used?" Try to visualize the reader using the report. How can the information be made more convenient? How much effort must be given to getting the attention and interest of the reader? Will be the report be ready by more than one person? If so, how many copies should be made? What will be the distribution of the report?

B. Presentation Considerations

The final consideration in the report-writing process is production. Reports can be typed; printed on an ink-jet, laser, color, or other printer, or sent out for typesetting. Most student and small research reports are typed or produced on a computer printer. The presentation of the report conveys to the readers the professional approach used

throughout the project. Care should be taken to use compatible fonts throughout the entire report. The printer should produce consistent, easy-to-read letters on quality paper. When reports are photocopied for more than one reader, make sure the copies are clean and have no black streak or gray areas.

Overcrowding the text creates an appearance problem. Readers need the visual relief provided by ample white space. I define "ample" as 1 inch of white space at the bottom, and right-hand margins. On the left side, the margin should be at least $1_{1/4}$ inches to provide room for binding or punched holes. Even greater margins will often improve report appearance and help to highlight key points or sections. Overcrowding also occurs when the report contains page after page of large blocks of unbroken text. This produces an unpleasant psychological effect on readers because of its formidable appearance. Overcrowding text, however, may be avoided in the following ways:

- Use shorter paragraphs. As a rough guide, any paragraph longer than half a page is suspect. Remember that each paragraph should represent a distinct thought.
- Indent parts of text that represent listings, long quotations, or examples.
- Use headings and subheadings to divide the report and its major sections into homogeneous topical parts.
- Use vertical listings of points (such as this list).

Inadequate labeling creates another physical problem. Each graph or table should contain enough information to be self-explanatory. Text headings and subheadings also help with labeling. They function as signs for the audience, describing the organization of the report and indicating the progress of discussion. They also help readers to skim the material and to return easily to particular sections of the report.

3.4. Oral Presentations

Researchers often present their findings orally. Such a presentation sometimes called a briefing, has some unique characteristics that distinguish it from most other kinds of public speaking: Only a small group of people is involved; statistics normally constitute an important portion of the topic; the audience members are usually managers with an interest in the topic, but they want to hear only the critical elements; speaking time will

often be as short as 20 minutes but may run longer than an hour; and the presentation is normally followed by questions and discussion.

A. Preparation

A successful briefing typically requires condensing a lengthy and complex body of information. Since speaking rates should not exceed 100 to 150 words per a minute, a 20-minute presentation limits you to about 2,000 to 3,000 words. If you are to communicate effectively under such conditions, you must plan carefully. Begin by asking two questions. First, how long should you plan to talk? Usually there is an indication of the acceptable presentation length. It may be the custom in an organization to take a given allotted time for a briefing. If the time is severely limited, then the need for topical priorities is obvious. This leads to the second question: What are the purposes of the briefing? Is it to raise concern about problems that have been uncovered? Is it to add to the knowledge of the audience members? Is it to give them conclusions and recommendations for their decision making? Questions such as these illustrate the general objectives of the report. After answering these questions, you should develop a detailed outline of what you are going to say. Such an outline should contain the following major parts:

1. *Opening*: a brief statement, probably not more than 10 percent of the allotted time, sets the stage for the body of the report. The opening should be direct, get attention, and introduce the nature of the discussion that follows. It should explain the nature of the project, how it came about, and what it attempted to do.
2. *Findings and conclusions*: the conclusions may be stated immediately after the opening remarks, with each conclusion followed by the findings that support it.
3. *Recommendations*: where appropriate, these are stated in the third stage; each recommendation may be followed by references to the conclusions leading to it. Presented in this manner, they provide a natural climax to the report. At the end of the presentation, it may be appropriate to call for questions from the audience.

Early in the planning stage you need to make two further decisions. The first concerns the type of audiovisuals that will be used and the role they will play in the presentation.

Audiovisual decisions are important enough that they are often made *before* the briefing outline and text are developed.

Presenting your research findings using PowerPoint™ or other presentation software requires preparation similar to presenting with nonelectronic visual aids. The researcher must still determine his or her style of presentation, the order of findings, and which findings will be presented graphically, in tabular format, or verbally. As most visual aids are prepared using computer software, the key hyperlink files are already available. It might seem as though the presenter could bypass the costly printing of visual aids, which can be a time consuming task. However, the electronic presenter must have a contingency plan for a malfunctioning computer. Color transparencies are the low tech backup but clearly don't allow the full range of possibilities that electronic hyperlinks afford. Having a second laptop and projection system is the usual high tech insurance plan. The same general rule applied to all presentations is critical for electronic ones-practice, practice, practices-but a caveat is added: Practice with your equipment so that movement between files, hyperlinks, and your PowerPoint™ presentation seems effortless.

The second discussion you must make as you plan for your presentation is what type it will be. Will it be memorized, read from your manuscript, or given extemporaneously? The impromptu briefing is rule out here because impromptu speaking does not involve preparation. Your reputation and the research effort should not be jeopardized by "winging it".

Memorization is a risky and time-consuming course to follow. Any memory slip during the presentation can be a catastrophe, and the delivery sounds stilted and distant. Memorization virtually precludes establishing rapport with the audience members and adapting to their reactions while you speak. It produces a self-or speaker-centered approach and is not recommended.

Reading a manuscript is also not advisable, even though many professors seem to reward students who do so (perhaps because they themselves get way with it at professional

meetings). The delivery sounds dull and lifeless because most people are not trained to read aloud, and therefore they do it badly. They become focused on the manuscript to the exclusion of the audience. This head-down preoccupation with the text is clearly inappropriate for management presentations.

The **extemporaneous presentation** is audience-centered and made from minimal notes on an outline. This mode permits the speaker to be natural, conversational, and flexible. Clearly, it is the best choice for an organizational setting. Preparation consists of writing a draft along with a complete sentence outline and converting the main points to notes. In this way, you can try lines of argument, experiment with various ways of expressing thoughts, and develop phraseology. Along the way, the main points are fixed sequentially in your mind, and supporting connections are made.

After the outline and audiovisual aids comes the final stage of preparation-the rehearsal. Rehearsal, a prerequisite to effective briefing, is too often slighted, especially by inexperienced speakers. Giving a briefing is an artistic performance, and nothing improves it more than for the speaker to demonstrate mastery of the art. First rehearsal efforts should concentrate on those parts of the presentation that are awkward or poorly developed. After the problem areas have been worked out, there should be at least a few full-scale practices under simulated presentation conditions. All parts should be timed and edited until the time target is met.

B. Delivery

While the content of a report is the chief concern, the speaker's delivery is also important. A polished presentation adds to the receptiveness of the audience, but there is some danger that the presentation may overpower the message. Fortunately, the typical research audience knows why it is assembled, has a high level of interest, and does not need to be entertained. Even so, the speaker faces a real challenge in communicating effectively. The delivery should be restrained. Demeanor, posture, dress, and total appearance should be appropriate for the occasion. Speed of speech, clarity of enunciation, pauses, and gestures all play their part. Voice pitch, tone quality, and

inflections are proper subjects for concern. There is little time for anecdotes and other rapport-developing techniques, yet the speaker must get and hold audience attention.

Speaker Problems

Inexperienced speakers have many difficulties in making presentations. They often are nervous at the start of a presentation and may even find breathing difficulties. This is natural and should not be of undue concern. It may help to take a deep breath or two, holding each of a brief time before exhaling as fully as possible. This can be done inconspicuously on the way to the podium.

Several characteristics of inexperienced speakers may be summarized as questions. Even if you are an accomplished speaker, it is helpful to review them as you watch a video of your presentation.

1. Vocal characteristics:
    a) Do you speak so softly that someone cannot hear you well? It is helpful to have someone in the back of the room who can signal if your voice is not carrying far enough.
    b) Do you speak too rapidly? Remind yourself to slow down. Make deliberate pauses before sentences. Speak words with precision without exaggerating. However, some people talk too slowly, and this can make the audience restive.
    c) Do you vary volume, tone quality, and rate of speaking? Any of these can be used successfully to add interest to the message and engage audience attention. Speakers should not let their words trial off as they complete a sentence.
2. Physical characteristics:
    a) Do you rock back and forth, roll or twist from side to side, or lean too much on the lectern?
    b) Do you hitch or tug on clothing, or fiddle with pocket change, keys, pencils, or other devices?
    c) Do you stare into space? Lack of eye contact is particularly bothersome to listeners and is common with inexperienced speakers. Many seem to choose a spot above the heads of the audience and continue to stare at this spot except

when looking at notes. *Eye contact is important*. Audience members need to feel that you are looking at them. It may be helpful to pick out three people in the audience (left, right, and center) and practice looking at them successively as you talk.

d) Do you misuse visuals by fumbling or putting them on in incorrect order or upside down? Do you turn your back to the audience to read from visuals?

Review Questions

1. Explain:
   a) Coding rules;
   b) Missing data;
   c) Field editing;
   d) Central editing.
2. Define Pareto diagram.
3. Suppose you were preparing two-way tables for percentages for the following pairs of variables. How would you run the percentages?
   a) Age and consumption of breakfast cereal;
   b) Marital status and sports participation;
   c) Crime rate and unemployment rate.
4. Distinguish between speaker-centered presentation and extemporaneous presentation.
5. Write the components of a research report.

ARBA MINCH UNIVERSITY

COLLEGE OF BUSINESS & ECONOMICS

DEPARTMENT OF ACCOUNTING & FINANCE

CONTINUING AND DISTANCE EDUCATION CENTER

*RESEARCH METHODS IN ACCOUNTING & FINANCE (*AcFn 2131*)*

ASSIGNMENT

Name_____ Semester:_____

ID No_____ Batch:_____

Department_____ Max. Mark: 30%

Center _____

N.B. The assignment should be submitted to the continuing and distance education center / coordination units of your center on or before_____.

Part One: True or False

Instruction: Write TRUE if the statement is correct and FALSE if the statement is incorrect. (1 point each)

_____1. In convenience sampling method, participants refer researchers to others who have characteristics, experiences, or attitudes similar to or different from their own.

_____2. All information is not of equal value and hence primary sources have more value than secondary sources, and secondary sources have more value than tertiary sources.

_____3. A concept is an image or abstract idea specifically invented for a given research and/or theory-building purpose.

_____4. A pilot test is conducted to detect weaknesses in design and instrumentation and to provide proxy data for selection of a probability sample.

_____5. In non-probability sampling, items for the sample are selected deliberately by the researcher; his choice concerning the items remains supreme.

Part Two: Multiple Choices

Instruction: Choose the best answer from the alternatives provided and write the letter of your choice on the space provided. (1 Point each)

_____6. Unstructured interview implies that

A. No specific questions or order of topics to be discussed, with each interview customized to each participant; generally starts with a participant narrative.
B. Generally starts with a few specific questions and follows the individual's tangents of thought with interviewer probes.
C. Often uses a detailed interview guide similar to a questionnaire to guide the question order and the specific way the questions are asked, but the questions generally remain open-ended.
D. None of the above

_____7. Which one of the following descriptors doesn't fall under the classification of research design based on the research environment?

A. Simulation
B. Laboratory research
C. Statistical study
D. Field setting

_____8. Among the following which factor could stimulate an interest in studying research methods?

A. The managers' increasing need for more and better information and for greater insights from that information.
B. The stakeholders demanding greater influence.
C. The competition, both global and domestic, is growing and often coming from unexpected sources.
D. All of the above

_____9. I n causal analysis,_____relationship exists when two variables mutually influence or reinforce each other.

A. Symmetrical
B. Reciprocal
C. Asymmetrical
D. None of the above

_____10. is a second independent variable that is included because it is believed to have a significant contributory or contingent effect on the originally stated IV_DV relationship.

A. Extraneous variable
B. Moderating variable
C. Intervening variable
D. Dependent variable

_____11. All are the major sources of error in communication research except,

A. Measurement questions and survey instruments
B. Participants
C. Interviewers
D. None of the above

_____12._____is a sampling technique applied in order to obtain a representative sample if a population from which a sample is to be drawn does not constitute a homogeneous group.

A. Cluster sampling
B. Multi-stage sampling
C. Stratified sampling
D. Sequential Sampling

_____13. Which of the following statements is true about percentages?

A. A larger percentage is easy to understand and is clear.
B. Percentages cannot be averaged unless each is weighted by the size of the group from which it is derived.
C. Percentages don't hide the base from which they have been computed.
D. Percentage decreases can exceed 100 percent.

_____14. One of the following parts of a research report is not included in the methodology section.

A. Data Collection
B. Sampling Design
C. Data Analysis
D. Problem statement

_____15. Among the following ways one is vital to avoid overcrowding text in research report production.

A. Using shorter paragraphs.
B. Indenting parts of text that represent listings, long quotations, or examples.
C. Using headings and subheadings to divide the report and its major sections into homogeneous topical parts.
D. All of the above

Part Three: Project Work

Instruction: Based on the lesson that you have learned in chapter three of the module about research proposals accomplish the following tasks. (15 Points)

1. Identify your own research topic.

2. Write the introduction section of the topic you identified.

3. Write the statement of problem for your topic.

4. Review literature for your identified topic.

5. Define clearly your general and specific objectives.

6. Write clearly the research design/methods for the identified topic.

7. Prepare estimated budget for your research.