HANDBOOK ON RESEARCH PROCEDURES

Prepared by

the
Research Committee
MCPA
in
Cooperation with

The Collegiate Employment Research Institute
Michigan State University
Career Development and Placement Services
113 Student Services Building
East Lansing, MI 48824-1113
(517) 355-9510

April, 1990

MCPA Research Committee Members

Philip D. Gardner
Robert C. Pebworth
David Baumgartner
Marily Bury
Michael P. Carey
Susan H. Cox
Carol N. Doelling
Jean E. Eisel
Richard Getch
JoAnn A. Goodyear
Jim Hurley
Sue Keihn
Batt Maher
John W. Mosser
Carol Olorunsola
Scott M. Preissler
Loretta Schneider
Lewis Schumaker

Any comments on this handbook or requests for copies can be directed to:

Dr. Phil Gardner
Collegiate Employment Research Institute
113 Student Services Building
Michigan State University
East Lansing, Michigan 48824
(517) 355-2211

Foreword

Daily we make decisions that affect our students, our employers, our staff, and ourselves. These decisions need to be based on the best information available to us. Our professional responsibility requests us to seek out this information; usually from someone else, but also our own personal knowledge and experiences provide a wealth of information. We face the challenge of minimizing the error, bias, misrepresentation and incompleteness associated with the information we use. To improve the usefulness of available information, we need an understanding of research procedures to the extent that we can evaluate the quality of information received from others and can follow the proper methods for acquiring our own information.

Research may appear threatening, because research is often considered synonymous with statistics. It is commonly assumed that research, like statistics, requires endless calculations. Since the expression "math was never by best subject," pervades our profession, there is a built in resistance to doing something that we dread. Why be frustrated? Why be stressed out? In reality we deal with numbers every day. The key to surviving in the face of statistical information is to know how to interpret statistics and when to use certain statistical methods. The calculations can be left to a computer programmer or statistician. Gaining an understanding of research procedures allows you to defend your own numerical findings and avoid being hoodwinked by the numbers supplied by others.

Recognizing the need to familiarize MCPA members with research procedures, the Research Committee of MCPA has developed resources designed to aid members in their information-gathering. The Committee administers two awards: the Steele and the Singleton, that recognize research efforts; provides assistance through mentors to those who are seeking research assistance; and offers workshops at annual meetings to those who want to brush up on research methods.

This handbook grew out of our first workshop held at Pheasant Run. The contents, in tersely outlined form, highlight the basic steps in the research process. Material has been collected from various sources and is not original material from the Committee. References are listed at the end of the handbook. Users of this handbook are encouraged to examine books on appropriate research methods and statistics for more details. This handbook is under constant revision; so comments, directed to the chair of the Research Committee, are always appreciated.

It behooves each of us to contribute to the profession through involvement in research. In fact, we (as a profession) should generate more of our own information rather than continue to rely on the work of others. We hope that through MCPA you have opportunities to expand your research skills.

Once Upon A Time

Mark Twain penned a story about Eve in which Eve commented on the acquisition of knowledge in a new world. Excerpts from Eve's diary, revealed the following:

> But studying, learning, inquiring into the cause and nature and purpose of everything we came across, were passions with us, and this research filled our days with brilliant and absorbing interest ... Each was ambitious to beat the other in scientific discovery, and this incentive added a spur to our friendly rivalry, and effectively protected us against falling into idle and unprofitable ways and frivolous pleasure seeking. (p.71)
>
> Our first memorable scientific discovery was the law that water and like fluids run downhill, not up. It was Adam that found this out. Days and days he conducted his experiments secretly, saying nothing to me about it; for he wanted to make perfectly sure before he spoke... My astonishment was his triumph, his reward. He took me from rill to rill - dozens of them - saying always, "There - you see it runs downhill - in every case it runs downhill, never up. My theory was right; it is proven, it is established, nothing can controvert it."
>
> In the present day, no child wonders to see the water run down and not up, but it was an amazing thing then, and as hard to believe as any fact I have ever encountered... (p.71-22)
>
> I scored the next great triumph for science myself: to wit, how the milk gets into the cow. Both of us had marveled over that mystery a long time. We have followed the cows around for years - that is, in the daytime - but had never caught them drinking a fluid of that color. And so, at last we said they undoubtedly procured it at night. Then we took turns and watched them by night. The result was the same - the puzzle remained unsolved... One night as I lay musing, and looking at the stars, a grand idea flashed through my head, and I saw my way!... deep in the woods I chose a small grassy spot and wattled in it, making a secure pen; then I enclosed a cow in it. I milked her dry, then

left her there, a prisoner. There was nothing there to drink - she must get milk by her secret alchemy, or stay dry.

All day I was in a fidget, and could not talk connectedly I was so preoccupied; but Adam was busy trying to invent the multiplication table, and did not notice. Toward sunset he had got as far as 6 times 9 are 27, and while he was drunk with the joy of his achievement and dead to my presence and all things else, I stole away to my cow. My hand shook so with excitement and with dread failure that some moments I could not get a grip on a teat; then I succeeded, and the milk came! Two gallons. Two gallons, and nothing to make it out of. I knew at once the explanation: The milk was not taken in by the mouth, it was condensed from the atmosphere through the cow's hair. I ran and told Adam, and his happiness was a great as mine, and his pride in me inexpressible.

Presently he said, "Do you know, you have not made merely one weighty and far-reaching contribution to science, but two!"

And that was true. By a series of experiments we had long ago arrived at the conclusion that atmospheric air consisted of water in invisible suspension; also, that the components of water were hydrogen and oxygen, the proportion of two parts of the former to one of the latter, and expressible by the symbol H_2O. My discovery revealed the fact that there was still another ingredient - milk. We enlarged to symbol to H_2O,M. (p.73)

From "Eve's Autobiography" see Mark Twain, Letters From the Earth, B. DeVoto (ed.), Greenwich, Conn: Faucett Publications (Crest), 1938 (1962). Also cited by C. Saslow, Basic Research Methods. New York: Random House. 1982.

TABLE OF CONTENTS

I

OVERVIEW

Twain's story captures the basic steps in the research process. First, a question arises that requires an answer. When methodical observation fails to produce an answer, an experiment is undertaken, using a prescribed set of methods, that will collect the necessary information. After analyzing the collected data and verifying the results, results can be reported to associates and friends. The final step involves generalizing to a wider body of knowledge.

Underlying this whimsical story of Eve's research efforts rests the strength of the scientific process: knowledge once gained is always subject to further testing. Experiments may need to be administered to a different population or undertaken for a longer period of time. No single research project can ever provide the final answer (Saslow); thus the scientific process can be considered self-correcting.

In opening this handbook, this section presents a general overview of some basic concepts and definitions commonly encountered when deciding on which paths to follow during the research process. This section also addresses some initial questions that a researcher needs to answer before beginning.

Saslow has described research as a systematic operation that requires a researcher to answer these questions before collecting data:

1. What are you looking for?
2. Whom are you going to observe or research?
3. When and where is the research to be done?
4. How are the observations to be made?
5. In what form are the observations to be recorded?

The systematic research can be conducted in several different formats. A case study offers an opportunity to develop a narrative description of a situation that may only apply to a small group of people or that is unusual in its occurrence. Observations are usually limited, as well as the information necessary to support conclusions. Case studies are common in medicine and psychoanalysis and are used as a teaching device in business schools. Case studies can be valuable to us as suggestions for future experimental research. The problem with case studies is that they are suggestive rather than conclusive.

It is not advisable to make strong conclusions based on a case study.

Ethnography, the observation of people (in our case) in their natural setting, is no longer the sole domain of anthropologists among social scientists. Educators, psychologists, and others have reassessed research paradigms to include more qualitative methods. Ethnography provides an excellent way to determine what happens in a situation. Many research questions also want to examine the "whys" of an event which ethnography does not do very well. Information that is gained from ethnographic observations can be used to direct experimental research into the "whys."

The approach most of us are more familiar with is the survey of large groups of individuals. This type of observational approach can target small to very large groups of people and approach them through structured interviews (face-to-face) or questionnaires (mail or phone). Surveys are useful in summarizing the characteristics of different groups of people or to estimate their feelings and attitudes about issues (Saslow, 13).

Each of these approaches can play an important role in our research efforts, depending on how we define the problem and on what we wish to accomplish. Because survey research is the most common approach in our profession, the material in this handbook will focus on this approach.

Before introducing specific topics on survey research, several distinctions need to be drawn for some words that are often commonly interchanged. First, consider two different types of surveys: opinion (marketing) and attitudinal. While the research steps are the same, the purpose and outcomes of these two surveys are different. Marketing researchers expend a lot of energy defining and targeting their sample to insure representativeness. The target group usually is administered a short, defined survey. In attitudinal research, the sample is important but more energy is given to the causal model of the behavior in question and the questions designed to support or refute that model. Attitudinal questionnaires tend to be longer and some would consider, repetitive in order to address issues of validity and reliability. Because people encounter marketing surveys more often, they may be resistent to attitudinal studies because of length and type of questions asked. Since placement professionals do not conduct a lot of attitudinal research and because marketing is currently a hot topic, we need to be careful that we do not confuse these two different approaches.

Research undertaken in the placement office or personnel department usually attempts to find a solution to a pressing practical problem. This type of research has been labeled applied research. However, conclusions drawn from an applied study often tend to make more general statements. When research

begins to move away from practical problem solving to a more general understanding, the research has taken on a basic flavor. A well thought-out applied project can contain an element of basic research. Solving a practical problem and adding to the general body of knowledge are not mutually exclusive goals. It does require the researcher to have a clear problem statement and a good understanding of the existing body of literature that pertains to the research question.

In today's educational headlines, few can miss the references to evaluation and assessment of higher education. Evaluation procedures differ from experimental research, though they are frequently lumped together. The purpose of evaluation is to determine if an existing program has produced a product or desired outcome. While evaluators use many research techniques, they have no control over the process (program) being studied; cannot set up the study with the intent of gathering dependable information; have no control over events prior to the initiation of the program or while it was in progress; and are called upon to make value judgments (Saslow, 25-26). Evaluators need research skills and others not often required of a researcher, e.g. political skills. If you are called on to do evaluation, remember you are not doing research per se; be prepared to modify your research procedures and to deal with a wider set of issues.

With these introductory comments completed, the more detailed information can be presented. The order of the handbook follows closely the general steps in the research process. Rephrasing Saslow's basic questions into the following section topics:

1. Problem Statement (What)
2. Research Design (How)
3. Variables (What)
4. Sampling (Whom)
5. Data Collection (When and Where)
6. Questionnaire Construction and Administration (How, When and Where)
7. Measurement (How)
8. Analysis and Interpretation (In What Form)
9. The Research Report

II

PROBLEM STATEMENT

Many people think that science is basically a fact-gathering activity.

> Get as many numbers as you can in your data base and you are bound to have something to report.

Scientific inquiry can not function without some preconceived ideas.

> It is these ideas that allow data to bear fruit.

The first step, like all first steps, is the most important step in research:

> An adequate statement of the problem.

Concede that it is not always possible to formulate a clear, simple problem statement -- recognize the complexity of the issues that we face.

> Nevertheless one's ability to define the problem will dictate one's ability to solve the problem.

This section will cover the criteria of a good problem statement, including the development of hypotheses, framing the problem through the literature review, and determining the feasibility of the problem.

A. Criteria of Good Problem Statements:

1. The problem should express a relation between two or more variables.

   Is A related to B?

   Are A and B related to C?

2. The problem should be stated clearly and unambiguously in question form.

   Questions have the virtue of posing problems directly.

3. The problem statement should be such as to imply possibilities for empirical testing.

   If a problem can not be tested, it is not a scientific

question.

Many important questions are not scientific questions, because they are not amenable to testing.

B. Problems are articulated and refined as hypothesis

An hypothesis: (a) conjectural statement of the relation between two or more variables.

(b) carries clear implications for testing the stated relations.

Restated: Variables are measurable and the relationship is specified (increase - null - decrease)

e.g. Individuals having the same or similar occupational role will hold similar attitudes toward cognitive objects significantly related to occupational role. (Cognitive objects -- concrete/abstract things perceived and "known" by individuals)

Relation: occupational roles - attitudes
(educator) (toward education)

Test: A comparison of at least two different occupational groups' attitudes toward some cognitive construct. e.g. teachers and businessmen attitudes toward education.

C. Importance of Hypotheses

1. Working instruments of theory.
2. Can be tested and shown to be probably true or false.
3. Exist apart from scientist's values and opinions.

D. Value of Problem Statement and Hypotheses

1. Directs Investigation - the problem statement and hypothesis serve as a road map to guide the investigator.
2. Enables researcher to deduce specific empirical evidence implied by the problem.

E. Differences Between a Statement and a Hypothesis

1. Hypotheses can be tested

Facts and variables are not tested as such.

Problems can not be scientifically solved (usually broad; plus questions are not tested)

2. Use of hypothesis is like playing a game of chance - the rules are established in advance and bets made. One can not change the rules after an outcome nor change one's bet after making them. (e.g. can't throw the dice and then bet, or, if you gather data first and then select a datum (dependent variable) and come to a conclusion on the basis of the datum, one has violated the rules of the game . . . in a fair game everything is counted).

F. Common Errors in Forming Problem Statement

1. Scientific problems are not moral and ethical questions.

2. Listing of methodological points as sub-problems.

3. Problems are too vague/too general.

4. Problem statement too specific.

G. Formulating Your Research Project

1. The Literature Review

While the problem may appear to be original to you, likely someone has already looked at the question or a variation of it.

The literature review permits a better frame to be built around the problem.

Review theories that may be relevant to the problem.

The literature review opens these opportunities:

(1) May be able to test theory itself.

(2) May suggest an approach to understanding your question.

(3) Identifies an earlier study(ies) that can be repeated to see if its (their) findings hold up.

(4) Permits the testing of whether a relationship found in one area extends to another area.

(5) Provides explanations to account for unexpected findings or failure to confirm predictions.

(6) Shows how techniques developed for solving the problem may be applied to the investigation of a different problem.

2. Why the emphasis on relating your study to other knowledge?

a. Clarify your problem.

b. Translate your concepts into operations.

c. Ensure that your effort will enhance existing body of knowledge.

d. Generalizability of results.

Single studies that do not seem to add anything to knowledge will not be published or widely used; the opportunities for pulling together such findings into a generalizable principle may be missed.

In many areas, a bibliographical survey will undoubtedly be more time consuming than rewarding.

Might find that no research of interest exists, but this is rather uncommon.

3. Need to look thoroughly through journals

1. Psychological Abstracts

Student Personnel Abstract

Higher Education Abstract

Current Contents/Social and Behavior Sciences

2. Recent issues of relevant journals (Journal of Vocational Behavior -- annual review)

3. ERIC

4. Doctoral Dissertation Abstracts

5. Associations (Academy of Management and AERA Division Newsletters)

H. Feasibility of the Project

Now that you have defined your problem and placed the problem in the context of the existing body of knowledge you can determine if your project is feasible.

Consider these factors:

1. Scope of the problem -- Is it appropriate?
2. Time (e.g., recruit subjects, construct questionnaire, collate data, conduct analyses, write report) -- Do you have enough?
3. Money -- Do you have enough?
4. Cooperation from others -- Is your support team available?
5. Availability of research subjects -- Are there enough within easy access?
6. Equipment -- Do you have access to necessary equipment, e.g. computers?
7. Avoidance of trouble -- Is it politically, socially, culturally feasible?

I. Decision Time

Choose One:

Continue - Review Other Options - Abandon

Let's Go!!

II

Research Design

Design: The plan and structure of investigation so conceived as to obtain answers to questions.

This step of the research process can involve very simple or very complex designs. As designs become more complex, they are harder to explain in abbreviated form; so you may want to refer to one of the references on research methods.

The purpose of the design step is to specify the independent variables (including their levels); to determine the grouping of your subjects; and to establish when measurements of your dependent variable will occur. Since the amount of information you obtain is influenced by the design you choose, careful attention needs to be given to each step.

A. Plan

1. Overall scheme or program of research.

   a. Outline of what is done.

      Hypothesis to operationalize implications to the final analysis.

   b. <u>Structure</u> (not often clearly defined)

      Framework, organization, or configuration of elements of the structure related in specific ways, e.g. math equation.

      or

      A paradigm or model of the relationships among the variables of a study.

2. <u>Purpose</u> (1) to provide answers to research questions
   (2) to control variance

3. <u>Plan</u> <u>Suggests</u>
   (1) What observations to make (how many).
   (2) How to make observations.
   (3) How to analyze the quantitative representations.
   (4) Designation of variables.

example Treatments

A_1	A_2
MA_1	MA_2

Acceptance Scores

Note: This one-variable randomized design with two partitions permits only a statistical test of the difference between two statistics: only one possible test.

with 2 x 2, design, for example, can test (A, A) (B, B) (A/B).

The research design is not a static plan. The initial design may not be adequate to the demands being made upon it. If information is sought on both A and B together in one experiment, it would be necessary to test for interactions.

4. Controlling Variance

a. Maximize the variance of the variable or variables of the substantive research hypothesis.

b. Control the variance of extraneous or "unwanted" variables that may have an effect on the experimental outcomes.

c. Minimize error or random variance.

Whenever we talk of variance we need to know what variance we are talking about.

We always mean variance of the dependent variable by manipulation and control of the independent variables' influence on the variance of the dependent variable.

A sub principal: Design, plan, and conduct research so that the experimental conditions are as different as possible

i.e., Independent variables -- need to vary as much as possible (5 point scale VS. 7 or 9 point scale)

5. Control of Extraneous Variables.

a. Eliminate Variables

To eliminate the effect of a possible influential

independent variable on a dependent variable, choose subjects so that they are as homogeneous as possible on that independent variable.

Remember: Lose ability to generalize.

b. Randomization

c. Build the extraneous variable into the design as an independent variable:

> An extraneous variable can be controlled by building it into the research design as an attribute variable thus achieving control and yielding additional research information about the effect of the variable on the dependent variable and about its possible interaction with other interdependent variables.

d. Match subjects

6. Minimization of Error Variance

Variability of measures due to random fluctuations whose basic characteristic is self-compensating (varying now this way, now that way) generally <u>cancel out</u>.

Measurement errors e.g., lapses of memory, guessing

a. Control conditions

b. Increase reliability of measure (accuracy).

<u>Without content - good theory, good hypothesis, good problem statement -- research design is empty</u>.

B. Experimental and non-experimental

Experiment: A scientific investigation in which an investigator manipulates and controls one or more independent variables and observes the dependent variable.

Key Word: Manipulates

Non-experimental: Cannot manipulate variables or assign subjects or treatments at random.

Experimental favored if possible - because we are more confident that the relationships being studied are the relations we think they are.

C. Faulty Designs

1. One-group

X	Y	Experimental
X	Y	Non-experimental

2. One-group, before-after

Yb	X	Ya	Experimental
Yb	X	Ya	Non-experimental

While all looks fine

a. Measurement effect
- measuring changes subjects - "reactive measures"

b. History between Yb and Ya. Things can happen, esp. the longer the period of time between measurements, that can affect the experiment. Such events become specific to the experiment.

3. Simulated before-after

	X	Ya
Yb		

Use pretest measures, such as the measures of another group, which are chosen to be similar as possible to the experimental group (acts as a control group).

Unfortunately don't know if two groups are equivalent before X.

4. Two Groups, No Control (very common)

(a)

X	Y	Experimental
X	Y	

(b)

X	Y	Non-experimental
X	Y	

Groups/subjects taken as they are.

Weaknesses

(1) Two groups assumed to be equal in the independent variables -- other than X

Groups not random.

D. Criteria for Research Design

1. Does the design answer the research questions?
   or
   Does the design adequately test the hypotheses?

   Design research to answer research questions.

2. Does this design adequately control independent variables?

   Randomize wherever possible: select subjects at random; assign subjects to groups at random; assign experimental treatments to groups at random.

   Control the independent variable so that extraneous and unwanted sources of systematic variance have minimal opportunity to operate.

3. Can we generalize the results of a study to other subjects, other groups, and other conditions?

   (How much can we generalize the results of the study?) Applied research has become much more concerned about generalizability.

4. Did X, the experimental manipulation, really make a difference?

   Internal Validity

5. When an experiment has been completed and a relation found, to what population can it be generalized?

   External Validity

6. Will relation hold if setting changed?

   Ecological representativeness

7. Are the variables of this research representative?

   Variable representativeness

E. General Designs

"Design is data discipline."

1. Conceptual Foundations

   Cartesian sets - (see measurement)

2. Designs

a. Experimental group - control group, randomized subjects

R

X Y Experimental
X Y Control

Advantages

a) Best theoretical control system

b) Flexible - extension to any number of groups with any number of variables.

c) Test several hypotheses at one time.

d) Statistically and structurally elegant.

b. Experimental Group-Control Group: Matched Subjects

Mr

X Y (Experimental)
X Y (Control)

Randomization must be incorporated -

c. Before and After Control Group (Pretest/Post-test)

R or Mr

Yb X Ya Experimental
Yb X Ya Control

This design has many advantages
Randomized
Control Group (controls for history and maturation)
Pretest Feature

Disadvantages
Pretest -
Can have a sensitizing effect on subjects; e.g., alerted to certain events in environment

May not be responding to X influence as to increased sensitivity

Run into a problem of generalizability.

d. Simulated Before-After Randomized

R Ya
Yb

Value of this design doubtful, even though acceptable: Improvement -- randomization

When to use: (a) When worried about the reactive effect of pretesting.

(b) No other choice.

e. Three-Group, Before-After

R

Yb	X	Ya	(Experimental)
Yb	∂X	Ya	(Control 1)
X	Ya		(Control 2)

Helps avoid problems of interactive pretests.

f. Four-Group, Before-After

R

Yb	X	Ya	(Experiment)
Yb	∂X	Ya	(Control 1)
	X	Ya	(Control 2)
	∂X	Ya	(Control 3)

Combination of two-best designs

Strong design: (a) Demand for comparison satisfied.

(b) Randomization enhances statistical equiv.

(c) History & maturation controlled.

(d) Interaction effect controlled

(e) Contemporaneous effects controlled.

Weaknesses: (a) Practicality -- hard to run two experiments and find subjects.

(b) Statistical -- lack of balance -- not "4" complete sets of measures

Save for very important studies.

Time Designs (Variants)

Yb	X	Ya
Yb	X	Ya
	X	Ya
	X	Yb

g. Compromise Experimental Group - Control Group

Yb	X	Ya	(Experimental)
Yb	X	Ya	(Control)

Very common -- Notice no randomization.

Weakness results from lack of equivalence between groups.

Attempt to obtain equivalence by other means -- (Don't give up if you can't equate groups by random selection.)

(a) Use the same population or samples alike as possible --

(b) Assign experiment at random --

(c) Similarity of groups should be checked -- e.g., sex, age, social class, etc.

(d) Equivalence of groups should be checked using means and standard deviations of the pretests:

t tests and F tests
Also check distributions

Selection is one of the troublesome problems of research.

Self-selection: i.e., volunteers an experiment when subjects are recruited on bases extraneous to the research program.

h. Incorporating Time

Y_1 Y_2 X Y_3 Y_4

(a) Separate reactive measurement effects.

(b) Statistical analysis of time measures is a special problem.
- Can yield spurious results.

F. **Basic difference between experimental and non-experimental research.**

In research, you are using some method to manipulate or measure X. You then look to see if Y has varied from the variation in X.

In an experimental design, generally predict from a controlled X to Y.

Using the principal of randomization and active manipulation of X can assume Y is varying as a result of X.

In non-experimental research Y is observed and X (or several X's) are observed at a time either before, after, or at the same time as Y.

So far the same as the experimental approach.

Essential difference is the direct control of X; in non-experimental research, direct control is not possible.

Neither manipulation or random assignment can be used. Therefore, cannot assert with confidence that X and Y related.

Self-Selection: Occurs when the members of the groups being studied are in the groups because they differentially possess characteristics extraneous to the research problem.

"Assign themselves."

Self selection can be subtle -

(a) samples -- selected in a non-random way.

(b) Comparison groups -- selected because they are in one group or another "possess" dependent variable in great/lesser degree.

These problems do not mean that non-experimental research is not important. Just need to be aware when we are using it -- two criteria to judge whether study is worthy: (1) methodological soundness and (2) substantive interest.

Major weaknesses of non-experimental research

(a) Inability to manipulate independent variables

(b) Lack of power to randomize

(c) Risk of improper interpretation

Value of non-experimental research

(a) Problem does not lend itself to experimental inquiry.

IV

VARIABLES

A word common to the research vocabulary is variable. Variable should have been introduced formally in the problem statement where the hypotheses usually define the variables to be used in the study. In fact, the entire design of your research project is done around variables. Therefore it is important that we are able to distinguish between different types of variables.

A variable is something that can be measured along a dimension; in other words, the variable has more than one value. Sex, for example, is a variable while female is not. (Note: A scale measuring one's degree of femininity can be created.)

Defining the variables in your study is very important. A clear definition allows (1) you to know what information to look for, and (2) other people to interpret your results. This step is referred to as "operationally defining the variables" in the study. Since a variable can be defined in a number of ways, it is necessary to derive a complete set of definitions as soon as possible. Consider how often you may have interpreted the results of a study differently than the author's because you employed a different operational definition.

Imprecise operational definitions are a very common problem in some research.

Variables can function in different ways. Some variables are chosen by the researcher prior to the experiment and the appropriate levels are established. For example, sex and academic major may be used to select participants. These type of variables are classified as independent variables and can be remembered as input variables. The variables which measure the responses of participants or can only be obtained after an experiment are referred to as dependent variables. In elaborate attitudinal studies, especially causal models, initial dependent variables can later become independent variables.

To be operalized, an independent variable must specify at least two levels. For independent variables with more than two levels, the researcher needs to be very explicit about the levels being used. The operational definitions for dependent variables are required to be more specific. For each dependent variable the definition should include what measures are to be taken, the scales to be used, and the procedure by which the measurements will be taken.

In many research situations, multiple independent variables are

included. With a multiplicative increase in experimental conditions, the entire design (see sample) is affected. Increasing the number of dependent variables is much easier and does not complicate things as much in the design. Simply, to increase the number of dependent variables, the survey instrument, for example, would have to be longer.

In assigning the levels to an independent variable, the researcher sets the stage for much of what follows in future steps. If the independent variable is defined in a non-subject manner (levels are created by direct manipulation of subjects - Saslow), subjects can be randomly assigned to levels. If the independent variable is defined in a subject manner (levels are defined according to characteristics of the subject - Saslow), subjects much be sought who possess those characteristics. The researcher loses some ability to randomize with subject defined independent variables. Note: some variables can be defined either way. Usually variables such as age, sex and grade point average, can only be subject independent variables.

Two other types of variables are also present in most survey research. Some variables, classified as relevant variables, may have an influence on the dependent variables. Relevant variables need to be either eliminated or controlled in the study. If not, these variables can cause changes in the dependent variable that are attributed to the effects of the independent variables(s). When relevant variables are not controlled, they are considered confounding -- essentially they confuse the effects of the independent variables.

A good research project carefully defines the variables to be used. If relevant variables are present they need to be controlled to prevent confounding the study.

V

SAMPLING

Sampling: Taking any portion of a population or universe as representative of that population or universe.

not is representative

rather considering it to be representative

Random Sample: A method of drawing a portion of a population so that each member of the population has an equal chance of being selected.

Reworded: Method of drawing a sample of a population so that all possible samples of fixed size n have the same probability of being selected.

Can never be sure that a random sample is truly representative. Representative means: exemplify the characteristics of the population. Therefore, if sex is an independent variable, the representative sample will have approximately the same proportion of men and women, e.g., as the population.

Randomization: Assignment to experimental treatments of members of a universe in such a way that for any given assignment every member of the universe has an equal probability of being chosen.

A. Sample Size

Often taught rule: get as large as sample as possible.

Real question: How much error is likely to be in statistics calculated from samples of different sizes.

1. Large numbers advocated not because they are good in and of themselves.

   Rather to give the principal of randomization a chance to work:

2. Small samples easier to obtain deviant sample. Therefore, the bottom line accepts 30 minimum as the acceptable number of observations per cell of your independent variable matrix. Note: This is after taking into account your response.

Addendum: In some statistical analyses, especially factor analysis and regression, sample size becomes extremely important.

(a) Measurement and sampling error

(b) Reliable identification/replication

Need to reduce intercorrelations

Rule-of-Thumb: Ten subjects for each variable (item) in a set.

B. **Kinds of Samples:**

Probability:
- (a) stratified
- (b) cluster sampling
- (c) systematic

Non-probability (often necessary and unavoidable)
- (a) Quota sampling
- (b) Purposive
- (c) Convenience (those at hand)

C. **Defining the Population**

1. Step One: Need a clear idea of the population from which the sample is to be selected. The first step in defining a population is to decide whether it is a population of individuals, households, institutions, transactions, or whatever.

2. Step Two: What units to include?

   a. Geography

   b. Age of Individuals

   c. Sex, race, marital status, education

   d. Citizenship, voter registration, intent to vote

   e. Household variables --

      First define a household

      (1) Census definition (everyone in a housing unit; unit occupied by tenants not eating with others; direct access from outside; complete kitchen facilities).

(2) Characteristics - e.g., mixed, single-head.

(3) Income.

Avoid over-defining the population. The simpler the universe definition, the easier and less costly to find the sample.

Your population definition must be able to be operationalized, e.g., all women of childbearing age or women 12 to 50; all residents affected by airport noise or residents within one mile of the airport

D. Small-Scale Sampling with Limited Resources

1. Credibility Scale - Used to judge small samples - includes the factors that samplers generally consider in looking at a sample.

   (1) Generalizability (geography, special populations, limitations).

   A careful discussion of the sample limitations is useful; carefully stated sample bias gains credibility when results are presented.

   (2) Sample size.

   (3) Sample execution.

   (4) Use of resources.

   The following table can be used to evaluate your sample.

   The weighting scales are personal judgments and depends on what is being sampled. The credibility score can range from 0 to 35 (your score/total possible) with higher scores being more credible.

   Can generalize by comparing the results from other studies.

CREDIBILITY SCALE FOR EVALUATING SAMPLE

CHARACTERISTICS	SCORE
A. Generalizability	
1. Geographic spread	
Single location	0
Several locations combined or Several locations compared	2
Limited geography	4
Widespread geography	6
Total universe	10
2. Discussion of limitations	
No discussion	0
Brief discussion	3
Detailed discussion	5
3. Use of special populations	
Obvious biases in sample that could affect results	-5
Used for convenience, no obvious bias	0
Necessary to test theory	5
General population	5
B. Sample size	
Too small even in total for meaningful analysis	0
Adequate for some, but not all major analyses	3
Adequate for purpose of study	5
C. Sample Execution	
Poor response rate, haphazard sample	0
Some evidence of careless field work	3
Reasonable response rate, controlled field operations	5
D. Use of resources	
Poor use of resources	0
Fair use of resources	3
Optimum use of resources	5
Maximum total points possible	35

Source: Sudman, Seymour. 19 . "Applied Sampling". In P. Rossi, J. Wright, and A. Anderson (eds) Handbook of Survey Research, New York: Academic Press.

E. Random and Systematic Sampling

To draw a random sample can assign numbers to the population and employ a random number table to generate sample.

By taking every eighth case, can systematically draw a sample from a larger number. Need two things:

(a) Sampling internal (i = N/n)
(b) Random starting point.

In most cases the systematic approach is appropriate as a sampling technique.

F. The Uses and Limitations of Lists

- The most difficult task in sampling is finding the appropriate list --

  There are no lists for the entire population or households (the Census Bureau never releases its list), men, women, young people, Blacks, etc.

- Mailing lists are primarily derived from memberships, e.g., a sample of boys under 16 from subscription list of Boy's Life would not be representative of all sixteen year old boys.

- At local level population and household lists are available -- e.g., Polk directories.

- Professional organizations, business establishments, etc. may have lists.

Common problems with lists.

(a) Contain blank or ineligible units.
Incorrect ways for overcoming blanks and ineligibles

(1) Take next name on list.

(2) Count down a fixed number of eligible names from the top of page.

Correct approach: ignore and not include in sample; thus you have to adjust the sampling interval.

(b) Duplications

Must have an equal probability of selection; duplicates introduce a bias.

(c) Omissions

The biggest problem with any list, e.g., a telephone list does not include those with no phones and unlisted numbers. Can take these steps:

(1) discarding the list

(2) using the list and ignoring commission

(3) using the list with supplementation

G. Appropriate Samples

In doing research in our field we operate on a limited budget which limits what we can do. As a result, we opt to use convenience samples -- captive students and alumni. This approach often is criticized as not representative. We are fortunate in that the issues we study are directly related to the population at hand. We are not asking students to make business decisions or household buying decisions. We are generally removed from the classroom, thus avoiding the instructor-student relationship where students may have a tendency to exaggerate response.

Need to be very careful then in choosing your sample. One easy way to overcome sample bias is to have several colleagues on other campuses collaborate in your research project drawing appropriate samples from their campus and administering the instrument to their sample.

see
Sudman, S. 1976. Applied Sampling. New York: Academic Press.

VI

DATA COLLECTION

A. Interviews

Interviews are the most widely used means of obtaining information. This technique is very practical -- just reflect on all the questions you answer each day (and who asks the questions).

Data - collection methods are grouped by the degree of their directness. Interviews are generally direct while ink blots, for example, are considered an indirect method.

Strength of Direct Methods

1) Gather a great deal of information

2) Time

3) Interpretation

Weaknesses of Direct Methods

1) Unwilling/reluctant/unable to give answers

2) Invalid data

1. Tools of science - used as:

   a. Exploratory device to help identify variables and relations, to suggest hypotheses, and to guide other phases of the research.

   b. Main instrument of research.

   c. Supplement with other methods.

Must be sure that reliability and validity are accounted for -- especially in training interviewers.

Remember an interview is a face-to face encounter where one person asks another questions of two types: structured or unstructured or standardized or unstandardized. In a standardized interview, the questions, their sequence, and wording are fixed. Unstandardized are more flexible, permitting the interviewer to follow certain lines of questioning.

2. Interview Schedule.

The planning and writing of an interview schedule is difficult because of the multiple meaning and ambiguity of words and lack of clear focus on problems. Many of us lack the background and experience to construct a schedule.

a. Kinds of schedule information

(1) Fixed-alternative items -- offer a choice.

(2) Open-end items - supply framework for response.

(a) funnel -- special sequence of open-ended questions that begins general and then focuses on specific topic.

(3) Scales - verbal items to each of which an individual responds to expressing degrees of agreement or disagreement or some other response category. Social researchers have increasingly used scale items in their work.

a. Criteria of Question-writing

(1) Is the question related to the research problem and objectives?

(2) Is the type of question appropriate?

(3) Is the item clear and unambiguous?

(4) Is the question a leading question?

(5) Does the question demand personal or delicate material that the respondent may resist?

(6) Does the question demand knowledge and information that the respondent does not have?

(7) Is the question loaded with <u>social desirability</u>?

b. Value of Interviews

This method can yield data unavailable from any other source; it is adaptable and can be used with

all types of respondents.

How does this method compare to other methods?

Lets compare with the questionnaire.

(1) Questionnaire is closed (generally), uniform, has good reliability. It is also anonymous and economical. Yet, the response rate is often low; can be confusing/ambiguous; respondents may not provide answers to open-ended questions (can not express themselves in writing).

(2) Interviews allow people to express their feelings, future intentions, etc. especially when scale items are used. Yet, best handled by experienced staff; takes time, energy and money. Most importantly -- need skilled interviewers.

see

C. Cannell and R. Kahn. 1968. "Interviewing". In G. Lindsey and E. Aronson (eds) The Handbook of Social Psychology (2nd ed). Vol II, Chapter 15. Reading, MA: Addison-Wesey.

D. Warwick and C. Lininger. 1975. The Sample Survey: Theory and Practice. New York: McGraw-Hill. Institute for Social Research, 1987. Interviewer's Manual. Ann Arbor: University of Michigan.

B. Objective Tests and Scales

This is the most used method in the behavioral sciences, indicated by the considerable time researchers spend constructing scales.

1. Objectivity

Objectivity is agreement among expert judges on what is observed. Thus, objective measures of observation finds everyone following the same rules in assigning numerals to objects.

Therefore, objective procedure is one in which agreement among observers is at a maximum (in other words, variance is at a minimum).

All methods of observation are inferential and all methods have some objectivity. The methods mentioned happen to be more objective than other methods.

2. Tests and Scales

A test is a systematic procedure in which the individuals tested are presented with a set of constructed stimuli to which they respond (simply a measurement instrument).

A scale is a set of symbols or numerals so constructed that they symbols (numerals) can be assigned by rule to the individual to whom the scale is applied -- the assignment being indicated by the individuals' possession of whatever the scale is supposed to measure.

A scale can be used in two ways:.

(1) to indicate a measuring instrument

(2) to indicate the systematized numerals of the measuring device

3. Types of Objective Measures

a. Intelligence and Aptitude Tests

b. Achievement Tests

c. Personality Measures (biggest problem is establishing validity -- that is why we use repeated measures)

d. Attitude scales

(1) Major Scales: somatic rating scales (Likert) equal - appearing internal scales cumulative (Gutman) scales.

e. Value Scales

(Culturally weighted preferences for things)

Because we have not scientifically studied values; few measures of values exist. Interestingly, values are a large part of our daily communication.

f. Other Scales

(1) F scale (measures authoritarianism)

(2) Interests (Strong-Campbell)

(3) Needs

(4) Dominance

(5) Social Responsibility

4. Social and Educational Indicators

Indicators seem to have two meanings:

(1) Statistics that reflect social conditions

(2) All forms of evidence on technological-social change, crime, investments, etc.

Social indicators are concepts and associated statistics that reflect social conditions and human status and that under certain conditions can be used as variables in behavioral research.

5. Types of Objective Scales

Independent: means that a person's response to an item is unrelated to his response to another item. e.g., True-False, Yes-No, Likert items.

Non-independent: forces respondent to choose one item or alternative that precludes the choice of other items or alternatives, e.g. forced-choice.

Advantages of independent scales: economic,; applicability of most statistical analyses; maximum information obtained as each item is answered;

Disadvantage: response - set bias (tendency to give the same answer)

Advantages of non-independent scales: reduces response bias.

Disadvantage: cost more; frustrating for respondent (patience); complexity

(a) Agreement-Disagreement Items

(1) Those permitting one of two choices

(2) Those permitting one of three or more choices

(3) Those permitting one or more of three or more choices

Main thing to keep in mind: The scoring system has to yield interpretable data congruent with the scoring system.

(b) Rank-order Scales

Order a set of items to some predetermined criterion.

Advantages

(1) Scales of individuals can easily be inter-correlated and analyzed

Composite rank orders of groups of individuals can be correlated.

(2) Scale values of a set of items can be calculated using one of the rank-order methods of scaling.

(3) Partially escape response bias.

(c) Forced-Choice Items

Paired-comparisons - must choose between set of alternatives on the basis of some criterion

May irk some respondents who are not comfortable making choices from what's given.

A typical forced-choice item is a TETRAD which consists of two pairs of items, one pair high in preference value, the other low with one member of each pair being a discriminator (valid) and the other irrelevant (not valid).

Forced-Choice Scheme

High preference - discriminator
High preference - irrelevant

Low preference - discriminator

Low preference - irrelevant

Pick item most descriptive of a person; then select least descriptive item.

Advantage - response bias and social desirability are controlled; validity presumably increased.

Disadvantage - too complex; difficult to make choices.

(d) Ipsative and Normative Measures

Normative measures are the usual kind of measures obtained with test and scales -- they are independent and subject to analysis of the mean.

Ipsative measures are systematically affected by other measures and are referred for interpretation of the same mean (each individual's set of measures has same mean and standard deviation)

Ranks are ipsative measures, therefore, usual statistics are not appropriate.

see

Adkins, D. 1974. Test Construction: Development and Interpretation or Achievement Tests, 2nd ed. Columbus: Charles E. Merrill.

Liker, R. 19832. "A Technique for the Measurement of Attitudes." Archives of Psychology, No. 140.

Robinson, J. and Shaver, P. 1969. Measures of Social Psychological Attitudes. New York: McGraw-Hill, 1967.

C. Available Materials and Content Analysis

1. Available Material

Purposes

(a) Can explore the nature of the data and the subjects -- from this can obtain an insight to the entire situation.

(b) Can suggest hypotheses.
(c) Can test hypotheses.

(d) Can check sample data.

(e) Can draw samples (voting lists, school registration, etc.)

Types and Sources

(a) Census and registration data.

(b) Archives

(c) Available databases (achievement test scores, longitudinal studies)

(d) Human relations area files

(e) <see> county clerks, local school districts, Department of Education, UNESCO, newspapers

2. Content Analysis

A method of studying and analyzing communications in a systematic, objective, and quantitative manner.

Instead of observing people directly, we study the words they use in communicating answers to questions or about ideas. The words become actual variables that can be quantified and systematically analyzed.

Letters, diaries, newspapers, minutes of meetings, open-ended questions are a few examples of what can be content analyzed.

(a) Categorization of the universe

Define the universe to be analyzed. Then you partition this universe into usable variables. The verbal cues are then categorized into subheadings. This first step takes a great deal of thought.

(b) Units of Analysis

(1) Word - smallest unit which is the easiest to work with. e.g., value and non-value words,; difficult, medium, easy words. Words are counted and assigned a category.

(2) Theme - Often a sentence/a proposition about something; thus it is more difficult to use. Themes are combined into sets of themes. If the themes are complex, content is difficult and possibly unreliable.

(3) Character - an individual in a literary production, e.g., analyzing stories.

(4) Space-time - physical measurement of content, e.g., number of inches, number of pages.

(5) Item - an important content unit. The item represents the entire production, e.g., essay, television program, news story. Can be used to measure creativity.

(c) Quantification

All materials are potentially quantifiable. It is not easy, but can be done.

(1) Most common method - nominal measurement -- count the number of objects in a category.

(2) Ranking - ordinal measurement - working with less than 30 objects can rank according to a set criterion.

(3) Rating - rated on degrees of some measure, e.g., creativity.

Certain conditions must be meet before quantification is possible:

(1) Count carefully when materials to be analyzed are representative

(2) Count carefully when the category items appear in the materials in sufficient numbers to justify counting.

If materials are not representative or the items appear infrequently, cannot generalize from statistics.

Computers have greatly boosted the role of content analysis in research by removing the onerous tasks.

see

Berelson, B. 1954. "Content Analysis" in G. Lindzey (ed) Handbook of Social Psychology. Reading, MA: Addison-Wesley.

Holsti, O. 196S. "Content Analysis" In G. Lindzey and E. Aronson (eds), The Handbook of Social Psychology (2nd ed). Reading, MA: Addison-Wesley

D. Other Methods

There are other data collection methods that a researcher could use:

1. Projective methods
2. Vignettes
3. Observing behavior
4. Sociometry

Check these methods references for more information on these topics:

Kerlinger, F. 1986. Foundations of Behavioral Research. New York: Holt, Rinehart and Winston, Inc.

Selltiz, C., L. S. Wrightsman, and S. W. Cook. 1976. Research Methods in Social Relations. New York: Holt, Rinehart and Winston, Inc.

VII

QUESTIONNAIRE CONSTRUCTION AND ADMINISTRATION

Survey research may require a questionnaire to be administered; but questionnaires are one of a group of tools that can be employed in survey reasearch. Consider:

1. Statistics compiled by government agencies or in archives, e.g., High School Longitudinal Study.

2. Direct observation or measurement.

3. Field experiments.

4. Content Analysis.

Therefore, do not assume that a brand new questionnaire is necessary for each new project.

Let's assume that the questionnaire appears to be the best instrument for the job; and:

1. That careful sample selection has taken place (sample will understand the questions).

2. That respondents will be willing to reply.

A. Questionnaire Construction

1. Standardized Questionnaires.

   To be used with large samples and data will be analyzed -- will need a standardized instrument so that each respondent receives the same stimulus.

   Additional conditions:

   1. Prescribed order for asking questions.

   2. Exact wording.

   3. Prescribed definitions/explanations.

2. Qualities of a good questionnaire.

   If you put four qualified researchers in a room and asked each to prepare an instrument for the same research problem, you would likely obtain four different instruments. The point being that there is no "How To" approach to designing instruments. A good survey, nevertheless, will meet these conditions:

1. Meet the objectives of the research.

2. Obtain the most complete and accurate information possible.

3. Do this within the limits of available time and resources.

    1. Content:

        a) Decide what information is required.

        b) Draft some questions to elicit information.

        c) Put them into meaningful order and format.

        d) Pretest.

        e) Go back to a.

    You don't start by writing questions; if your conceptual model is well developed then the variables for which you seek information will stand out.

    Finding what factors are relevant to the problem is the key.

    Go back to your literature review to see what factors have been employed and have not.

    Do not be afraid to discuss your project with other experts; people familiar with the situation -- they can be provide a great deal of help.

    Finally, hold conversations with several members of the population you tend to survey. This provides a glimpse of reality.

Variable checklist: What should be included?

(1) Knowledge or awareness of issue.

(2) Interest in the problem.

(3) Attitudes toward issue.

(4) Why does respondent feel that way?

(5) How strong (intensity) of feelings?

(6) Saliency.

(7) Expectations for the future.

(8) Readiness to act (bridges gap between attitudes and behavior).

(9) Perceptions.

(10) Demographic factors.

Since you can not ask everything, your instrument will not contain all these elements. Plus, the list is only a partial one that is limited only by your creativity.

B. Drafting Questions

Think about all the people that ask you questions everyday-most are indirect/not to the point and even biased. Think about your answer -- they are not very specific, evasive, and possibly misleading.

This type of question-answer repartee is not useful in scientific studies. Interestingly, as important as question writing is, little has been written on the topic.

1. Open versus closed questions.

(1) Open: advantage -- allow for responses to be framed from the respondents perspective without aid of cues, and reveal what is salient to respondent; disadvantages-- elicit irrelevant material, the point of the question can be missed, differences in responses difficult to standardize.

(2) Closed: advantage -- less variability in responses, meaning more explicit, easier to code,less potential for interviewing error; disadvantage -- suggest answers not thought of before, force respondents into unnatural frame of reference.

In most cases, you will use closed questions unless (a) building rapport with respondent, (b) pretesting, and (c) not sure of all possible response categories.

2. Response Categories.

Many questions naturally demand a dichotous response (do you own a VCR?). Often we want to sort into two groups. However, two response categories provide a poor response distribution-- respondents tend to lump

together making it difficult to discern differences.

Has become popular to use five point Likert scales because (1) they are flexible and (2) they can be combined into subgroups.

Note: scales with more definition (e.g., 7, 9) do not necessarily improve the situation:

(a) Difficult to assign anchors (some people can not make distinctions without anchors).

(b) Too often on longer scales people tend to pick the middle.

(c) Can be difficult to deal with large categories in data analysis.

Common response categories (that can be 5-point scales).

1. Excellent - good - fair - poor (useful in ratings).

2. Approve - disapprove (how strongly one feels).

3. Agree - disagree.

4. Too many (much) - about right - not enough (little) (measuring people's satisfaction).

5. Better - worse - about the same (comparisons).

6. Very - fairly - not at all (importance).

7. Regularly - often - seldom - never (frequency).

8. Always - most of the time - rarely (frequency).

9. More likely - less likely (measure of probability).

All these scales work well with list of items so that you can ask a number of questions)provide additional stimuli).

One area of concern is whether to include the neutral or middle position. There are divergent positions on this issue. Generally, if the researcher has expectations on the response distribution, the neutral alternative should be omitted if the majority of respondents may select it because it would be a "safe" answer. The same rational applies to the use of "don't know" -- this is a "lazy answer" that respondents tend to use if made available. The bad side to this decision is that some people will be forced to give answers to questions that they have not thought about and

have not formulated an answer.

The "other" option also can present problems because it qualifies answers with the suggestion "it depends". "Other" is useful in a list of items that may not be inclusive.

3. Principles of Question Writing

    1. Keep it simple -- be aware of the vocabulary of your audience. Confusing/complex surveys result in a lot of don't knows.

    2. Avoid lengthy questions.

    3. Specify alternative and make respondents aware that any position is acceptable.

    Avoid the "loaded question" unless you absolutely feel it is the only way to get a response.

4. Common Errors in writing questions.

    1. Double-barreled questions.

    2. False premises.

    3. Vague, ambiguous words.

    4. Overlapping alternatives.

    5. Double negatives.

    6. Intentions to act.

    Demographic data are requested as the needs of the research dictate.

C. Question Order and Format

A major task in questionnaire construction is to group questions into some reasonable order and then, place in a readable format.

Implicit in this task is the need for standardized introduction that includes general instructions.

The introduction is crucial.

When can alleviate concerns about the tasks to be done.

(a) Who the survey is for.

(b) What is it about.

(c) Who is doing the survey and your association.

1. Ordering the questions.

   (a) An absolutely inflexible rule: make the opening question an easy, non-threatening one.

   (b) Similarly, only ask the difficult or sensitive question when the respondent is well into the survey. Usually demographic questions come at the end.

   (c) Questions should follow some logical order -- one leads into the next.

   Prepare respondent for a shift in the orientation of the questions when they occur.

   "Now I have some different questions".

2. Length of questionnaire.

   There are no golden rules on the length of a survey. A long survey has advantages in that more data can be obtained (expanded opportunities). Long surveys are costly though in a number of aspects: preparation, coding, processing. Also respondents can become distracted and not complete the survey.

   In general, surveys tend to be too long. Avoid asking questions that may be "interesting". You have to justify every item in the instrument.

3. Format.

   Conventions vary widely -- depends on your preference. Be consistent! Look at several surveys for the style you like.

   e.g. Horizontal format works best with a list of items; if using an open-ended item be sure to leave enough lines to accommodate handwriting styles.

Make directions clear and concise, especially when leading respondent (asking to skip questions if not applicable).

Booklet format recommended (printed double-sided).

D. Pretesting

Pretest! Pretest! Pretest!

Try out your questionnaire preferably on strangers. If respondents have difficulties note them down. Usually you only need 6 to 12 to get a feel for how people are responding. (A good representative sample will make the real" thing go smoother.) Go back to the drawing board if necessary.

E. Administration

There are two ways to administer your survey: personal interview or self-administered. Personal interviews are commonly used with small samples (less than 100) and in conjunction with special topics (e.g., dual careers) where the interviewer can probe areas that are not suitable for closed questions. There are a number of issues in personal interviewing, particularly interviewer training, that have to be addressed before administration can begin. The following discussion focuses on self-administered (mail) surveys.

What has to be overcome in self-administered surveys are: low response rates, incomplete samples, and misinterpretation of questions.

A. Dillman's: Total Design Methods

This set of standard procedures can assist in overcoming the problems of mailed surveys.

Part I: Identifying and designing each aspect of the survey process that may affect response in a way that maximizes response rates.

Part II: Organizing the survey effort in a way that assures that the design intentions are carried out in complete detail.

Basic Assumption: A person is most likely to respond to a questionnaire when the perceived costs of doing so are minimized, the perceived rewards

are maximized, and the respondent trusts that the expected rewards will be delivered.

General Principles for Design:

(1) Design questionnaire as booklet (ideally 6 1/2 x 8 1/4 inches).

(2) Questionnaire typed regularly then reduced.

(3) Avoid resemblance to an advertising brochure.

(4) No questions on first page; have an interesting title, graphics and any necessary instructions.

(5) No questions on last page; use for comments and thank you's.

(6) Order questions so that most interesting questions come first; place objectionable questions at the end with demographic questions last.

(7) Give special attention to first question.

(8) Format each page with great care; lowercase letters for questions/uppercase letters for answers, attempt to answer in straight vertical line, try to avoid splitting questions from one page to another, use transition to guide respondent.

Length can influence response rate: a 10 page survey versus a 14 page survey can increase the response by 10%.

B. General Procedures for Implementation

(1) One-page cover letter is prepared that explains.

1. The social usefulness of study.

2. Why each respondent is important.

3. Who should complete the questionnaire.

4. The confidentiality of results and the number system which is for follow-up purposes.

(2) Exact mailing date is placed on letter.

(3) Individual names and addresses are typed onto the printer letters; researchers individually sign each letter.

(4) Questionnaires are stamped with an identification number (explained in cover letter).

(5) Mailing packet consists of cover letter, questionnaire, and business reply envelope which are placed in monarch envelope on which the recipient's name and address have been typed (never use mailing labels).

(6) Mail using first-class postage if at all possible.

(7) One (1) week after mailing, send follow-up reminders.

(8) Three (3) weeks after first mailing, a second cover letter and questionnaire is sent everyone who has not responded.

(9) Eight (8) weeks after first mailing, a complete set is sent to non-respondents via certified mail.

C. Limitations of Mail Surveys

(1) Assessing a representative sample of a population esp, the general population.

(2) Among those likely to refuse will be those with lower educations.

(3) Difficulty in handling certain kinds of questions, esp open-ended items, repetitious attitude items,and screening questions.

(4) Length of time needed for implementation.

D. Costs

Not an easy question to answer.

(a) Are all costs to be included (every aspect)
VS. (b) Counting costs only above normal daily business (additional costs)

In general, smaller surveys are relatively more expensive (economies of scale).

REFERENCES

Bradburn, N. M. et. al. 1979. Improving Interview Methods and Questionnaire Design. San Francisco: Jossey-Bass.

Dillman, D. A. 1978. Mail and Telephone Surveys: The Total Design Method. New York: Wiley-Interscience.

Payne, S. L. 1951. The Art of Asking Questions. Princeton: Princeton University Press.

VII

MEASUREMENT

A considerable amount of a researcher's time can be and often is spent (sometimes preoccupied) on measurement. Nevertheless, measurement is one of the most important topics in research design.

A. Why? These four issues are usually cited:

(1) Philosophical concerns over the nature of measurement and its role in scientific process.

(2) Reliability and validity of measurement.

(3) Development of formal models of measurement.

(4) Specific techniques for constructing measurements.

Stevens has defined measurement as:

> The assignment of numerals to objects or events according to rules.

The only rule that does not apply to measurement is randomization. The rule is important because the rule determines what kinds of mathematical operations we can legitimately perform on the measurement.

> The rule is the guide, method, or command that tells us what to do.

Underlying measurement is a theory grounded in set or Cartesian principles. For a background, several concepts will be reviewed.

> Binary relations is any statement about two objects that is meaningful in the sense that it can be classified as true or false, e.g., preference relation.
>
> A binary relation has three key attributes --
> Reflexivity (aPa for all a in set A)
> Symmetric (aPb implies bPa for all a and b in A)
> Transitive (aPb and bPc implies aPc for all a, b, c in A)
>
> A relation that has the above attributes is termed an equivalence relation (=, equal to, same as)
>
> *The importance of this equivalence relation is that it enables a given set to be portioned into

non-overlapping equivalence classes, e.g., formation of mutually exclusive categories.

*The transitivity axiom is especially important because it allows elements to be ranked and compared. Without transitivity, even if aPb and bPc, the element a and c can not be directly compared. Transitivity is the key to ordering.

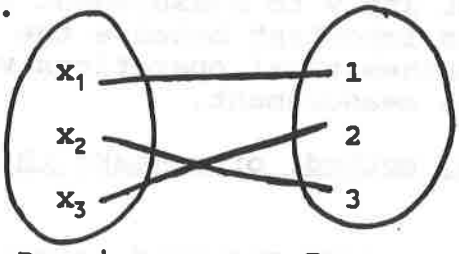
A number is assigned to each member of the set for quantitative meaning. This assignment process is called mapping; mapping allows us to associate one set with another

e.g., sex 1 = male 0 = female
nice 1 = not nice to 5 = very nice
can bring these two sets together.

A binary operation occurs when two elements of A are combined to produce a third element in the set.

A general equation for any measurement procedure
f = (x, y); x = any object, and y = a numeral

The function, f, is equal to the set of ordered pairs (x, y) such that x is an object and each corresponding y is a numeral.



Domain e.g., children | Range (Rank)

One last principle: morphism

Let A be a set on which the binary operation O is defined, and let B be another set on which the binary operation Q is defined.

Morphism is the mapping from set A to set B (called u) so that u (a, Oa_2) - ua_1Qua_2

Note: to work the mapping must preserve the operations.

This insures Isomorphism - which means similarity of form.

*Important because it permits algebraic systems to operate.

B. Scale

First step is to define the objects in the universe being studied (e.g., your sample population). Next, the properties of the objects of the universe must be defined. Their objects are classified according to the possession of these properties (have, not have) -- e.g., male, female. Finally, a count is made of the members of each classification.

1. Types of Scales

a. Nominal (one-to-one mapping). This is the lowest level of measurement. The numbers assigned to objects are numerical without having number meaning; however, they cannot be ordered or added. Assignment to groups, team, occupation, sex permits counting the objects in a cell or subset.

b. Ordinal (set of all monotonically increasing, continuous mappings). The objects of a set can be rank-ordered on an operationally defined characteristic or property.

(a) a > b > c > . . . > n on some property.

(b) Combined criteria, e.g., gpa and SAT for academic achievement.

(c) Greater than - less than, > or <.

Numbers assigned are rank values. These numbers indicate order not absolute quantifies nor are the intervals necessarily equal (though you may imply so).

c. Interval (a set of all positive linear transformations of f). Possesses the characteristics of both nominal and ordinal scales. Do assume that the difference between numbers are equi-distant (equally spaced). Intervals can be added and subtracted and meaningfully compared.

d. Ratio (set of mappings differing from f by a positive multiple). Possesses qualities of all previous scales and has an absolute (natural) zero that has empirical meaning.

2. Meaningful Statistics

Scale types vary on the admissible transformations permitted. Thus, the interpretation of statistics depends on the scale type.

Meaningfulness of the results depends upon the particular measurement models by which numbers are assigned.

*The calculation of the arithmetic mean is meaningless for ordinal scales since the mean statistic is destroyed by the monotone transformations. However, the median is meaningful.

*The sum of two values x1+x2 is meaningful if both from the same interval scale, but is not when measured on a ratio scale.

*Standard deviation is meaningful statistic for interval scales. However the statistic coefficient of variation is not -- is meaningful for ratio scales.

e.g., Temperature can be measured on two interval scales with arbitrary zero points. The two are related by linear transformation.

A meaningless statement would be:

25 degrees Celsius yesterday and 50 degrees Celsius today -- it is twice as hot today. The 50/25 = 2 ratio is meaningless for interval scales. Look what happens when converting to Fahrenheit degrees Celsius equivalents.

25 degrees Celsius = 77 degrees Fahrenheit 50 degrees Celsius = 122 degrees Fahrenheit 122/77 =/ 50/25

Note: Must be careful when using scales and statistics.

C. Scaling Techniques

1. Single-Item Measurement: based on a single item that elicits a single response. Problem with this scale is that we may not be measuring relevant attitudes.

2. Thurstone Scales: Method of paired comparisons. Presents judge with every possible pair, asking which is favored. From the judging a scale is created.

3. Likert Scales (summated scales). Most commonly used scales in social science research.

Basic Assumptions

(a) Continuous underlying attitude dimension and that each item is monotonically related to that continuum.

(b) Sum of the item scores is monotonically related to the attitude.

(c) A single common factor (measuring only one underlying attitude)

Steps:

1. Selection of the items. There is a wide variety of attitude statements that could be used which means you need to have a theoretical understanding of the attitude to be measured (see how to write good questions also).

2. Pretest statements. You can discard items where no variation occurs. The examination of the relationship of each item to the total score (correlations) -- want items that correlate highly to total score.

3. Scale scores are computed. Then reliability measure (see below) is made.

4. Guttman Scaling. Guttman's scales have a accumulative property which is based on the notion of unidimensionality. The underlying assumption of this approach is that attitude is a unidimensional phenomenon and that statements favorable or unfavorable about the attitude object can be ordered. This approach is complicated, particularly the statistical test, though it is widely used.

see
McIver, John P. and E. G. Garmines. 1981. Unidimensional Scaling. No. 24 in Quantitative Applications in the Social Sciences. Beverly Hills: Sage Publications.

5. Latent Structure Analysis

This approach accounts for all the interitem relations. The model is based on the assumption that the observed interrelations among the items are due to respondents

falling into two or more distinct subgroups. Due to its complexity, not widely used.

6. Other Approaches

Semantic Differential
Conjoint Measurement
Magnitude Scaling

D. Reliability

Synonyms: dependability, stability, consistency, predictability, accuracy.

Several ways to define:

1. If we repeatedly measure the same set of objects with similar instrument, will the results be the same or similar? Stability/dependability/predictability.

2. Are the measures obtained from a measuring instrument the "true" measures of the property measured? Accuracy.

3. How much error of measurement is there in a measuring instrument?

Systematic variance (constant or biased). Random variance (self-compensating).

Reliability is the accuracy or precision of a measuring instrument.

The underlying theory of reliability ties into variance. The total obtained variance has two components: true variance component, error variance component.

What we want to look at is the true variance.

The reliability then is:

(a) The proportion of the true variance to the total obtained variance of the data yielded by a measuring instrument.

(b) The proportion of error variance to the total variance yielded by a measuring instrument subtracted from 1.00, the index 1.00 indicating perfect reliability:

$r_{tt} = 1 - \frac{Ve}{Vt}$ reliability coefficient

Interpretation of the reliability coefficient.

If r, the coefficient of correlation, is squared, it becomes a coefficient of determination. The coefficient of determination gives the proportion or percentage of the variance shared by two variables.

e.g., r = .90 then variables share (.90)2 = 81% of the total variance of the two common variables.

*The reliability coefficient is also a coefficient of determination which tells how much variance of the total variance of a measured variable is "true" variance.

If r_t represents the coefficient of correlation between the obtained score and the "true" scores, X

$r_{tt} = r_t^2$ reliability coefficient

Another way to explain this concept is through the concept of internal consistency. If sub samples are taken, correlate the scores. If the test items are homogenous the reliability should be high. If they are not, the test is not reliable.

We strive to improve reliability (Maximincon principle) by maximizing the variance of the individual differences while minimizing error variance.

(1) Write the items of psychological and educational measuring instruments unambiguously.

(2) If an instrument is not reliable enough, add more items of equal kind and quality.

(3) Clear and standard instructions tend to reduce errors of measuring.

To be interpretable, a test must be reliable. Without reliability you can not determine the relationship between variables.

Confused: don't worry - SPSSx will calculate reliability coefficients for you. Rule of thumb: coefficients greater than .75 are acceptable between 6.0 and 7.5 coefficients are useable but should be aware that measure is not as strong as you would like.

D. Validity

A complex and controversial topic in behavioral science.

Validity deals with the reality or meaning of variables.

Are we measuring what we think we are measuring? Emphasis on <u>what</u>.

A test may be reliable, but not valid --

reliability refers to <u>factual</u> validity refers to <u>understanding</u>

Three types of validity:

1. Content Validity

- The representativeness or sampling adequacy of the content
  - is the substance or content of this measure representative of the content or the universe of content of the property being measured?

- Basically judgmental; have several competent judges review content.

2. Criterion-Related Validity

is studied by comparing test or scale scores with one or more external variables known or believed to measure the attribute under study.

How well does the test (or tests) predict an outcome? Look at predictive ability.

- The higher the correlation between a measure or measures of academic aptitude and the criterion the better the validity.

3. Construct Validity

Usually want to know what psychological or other property or properties can explain the variance of tests.

"Meaning of the Test"

- Interest is more in the properties being measured than in the tests -- or the theory behind the test.

Thus, construct validity is preoccupied with theory.

I will leave the mathematical statistical concepts for your future reading.

see
Kerlinger, F. 1986 Foundations of Behavioral Research New York: Holt, Rinehart and Winston.

Poor measurement can invalidate any scientific study. Most of the criticism by all observers centers on validity. Achieving reliability is mainly a technical matter. Validity is more than adjusting numbers and items because it incorporates the theoretical or philosophical underpinnings of what is being carried out. More attention is being given these two concepts as they pertain to all measurement instruments. The bottom line is that poor/inadequate measurement issues will not be tolerated.

IX

ANALYSIS AND INTERPRETATION

The purpose of analysis is to summarize the completed observations in such a manner that they yield answers to the research questions.

The purpose of interpretation is to search for the broader meaning that links your answers to the general body of knowledge.

Everything you have done prior to these steps were done with the intention of fulfilling these research conditions. This is why we started in the first place.

Remember: The relationship between analysis and interpretation and the problems of analysis and interpretation will vary according to the research design being used.

Statistics is the theory and method of analyzing data obtained from samples of observations in order to study and compare sources of variance of phenomena, to help make decisions to accept or reject huypothesized relations between the phenomena, and to aid in making reliable inferences from epirical observations.

A. Interpretation

Interpretation often is melded right into the analysis and rightly so; but the process of interpretation needs to be clarified.

(1) Interpretation is done in order to establish continuity -- the linking of your results with other studies. The establishment of continuity provides a path from confirmation to new hypotheses.

(2) Interpretation leads to the establishment of explanatory concepts.

B. Analysis and Interpretation

There are numerous ways to analyze data: use of non-quantifiable data, classification, tabulation and statistical analysis. The overwhelming emphasis of the research literature is on statistical analysis. In this section, each approach will be briefly considered.

1. Non-Quantified Data

There is a on-going philosophical argument between positivists and phenomenologists over the use of non-quantifiable data. Given the doubt about the adequacy of some of the measurement scales we employ -- non-quantifiable data can be very helpful. The unsolicited comment on a survey, an oral commentary on a particular issue, and other types of observations not planned for in the research design are classified as non-quantifiable. While non-quantifiable comments can embellish the report they are not necessarily scientific. Some ways they are important:

(a) Lead to new insights - they may help in clarifying a relationship among measured variables that have been shown statistically important. The perceptions of observers can prove important clues. A clue or observation, however, does not necessarily demonstrate that the possible explanation is correct unless a suitable measure can be found or developed for statistical analysis.

The search for explanatory concepts may come through non-quantifiable data, especially for cases that do not display the same patterns or trends as most cases in the study, e.g., one ethnic group varies. These cases usually are small in number.

(b) Illustrate the meaning of categories. Often a category can have a wide range of meaning; comments can be used to make distinctions between members of a category.

2. Statistical Analysis

Descriptive Statistics: Characterize the methods employed in summarizing the obtained data.

Inferential Statistics: Characterize the methods employed in making and evaluating generalizations from the data.

(a) Descriptive

(1) To characterize what is typical in the group (mean, average, median, mode).

(2) To indicate how widely individuals in the group vary (range, quartile deviation, standard deviation).

(3) To show other aspects of how the individuals are distributed with respect to the variable being measured. (Frequency)

(4) To show the relation of the different variables in the data to one another. (Cross-tabulations, correlations).

(5) To describe the differences among two or more groups of individuals. (analysis of variance (simple), chi-square)

C. Generalization

When comparing samples, generally want to test some hypothesis about the nature of the difference.

Tests of significance (statistical) have been devised which tell whether two or more samples actually differ in a manner that is significant.

In testing we deal with inferences and probabilities; thus there exist some chance of making an error, e.g., accepting the hypothesis as true when in fact it isn't or vice versa.

Type I error (reject null hypothesis when in fact it is true): risk of making this error is determined by the level of significance used in the test. Type I error can be minimized by making the level of significance more extreme.

However, by tightening up on Type I error increase the risk of Type II error.

Type II error (accept the null hypothesis when it is false): This error means we have failed to recognize a difference when it actually existed. To reduce Type II error -- use large samples or statistical tests that can use all available information.

Given the existence of these errors, a researcher can not rely entirely on the statistical evidence for the interpretation of the result. Additional evidence, usually indirect, will need to be used, e.g., other research results and existing knowledge.

While statistical tests may indicate a difference exists, they do not reveal <u>necessarily</u> the <u>reason</u> for

the relationship. If we want to establish causal inferences, we must meet additional assumptions and conditions above those for statistical tests of a relationship.

D. Inferring Causal Relations

Three types of evidence are required.

(1) That X and Y vary together in the way predicted by the specific hypothesis.

(2) That Y did not precede X in time.

(3) That other factors did not determine Y.

In experimental design, these conditions are met. In non-experimental design safeguards need to be built into the study (commonly add additional variables to control for their effect -- thus the investigator must know what other factors may influence Y).

Spurious relationships between X and Y exists if their variation stems, not from a connection between them, but from the fact that each of them is related to a third variable or combination of variables. Modeling techniques help avoid interpreting spurious relationships as causal.

E. Statistics

It is at this point that it is necessary to have a firm understanding of statistics -- or find a friend that does.

Purposes

(1) To reduce large quantities of data to manageable and understandable form.

(2) To aid in study of populations and samples.

(3) To aid in decision making.

(4) To aid in making reliable inferences.

a. Normal probability curve and the standard deviation

- basis for all statistical procedures -
- serves as a bench mark -

b. Standard Error

Another basic principal: The standard deviation is the sampling distribution of any given measure -- e.g., the mean.

F. Analysis of Variance

The most widely used method to analyze data is analysis of variance or one of its variations.

Concept: Differences of more than two groups can be tested for significance (t-test only applied to two groups).

There are several analysis procedures:

One-way
Two-way (multiple)
or
Factorial
Correlated Groups
Nonparametric (not normally distributed)

G. Analytical Procedures

I mention these analytical procedures because they are the backbone of many analyses. Time does not permit a discussion of each; so you are referred to a good text. However, some attempt should be made to understand multiple regression analysis.

(1) Multiple Regression

(2) Analysis of Covariance

(3) Discriminant Analysis

(4) Path Analysis

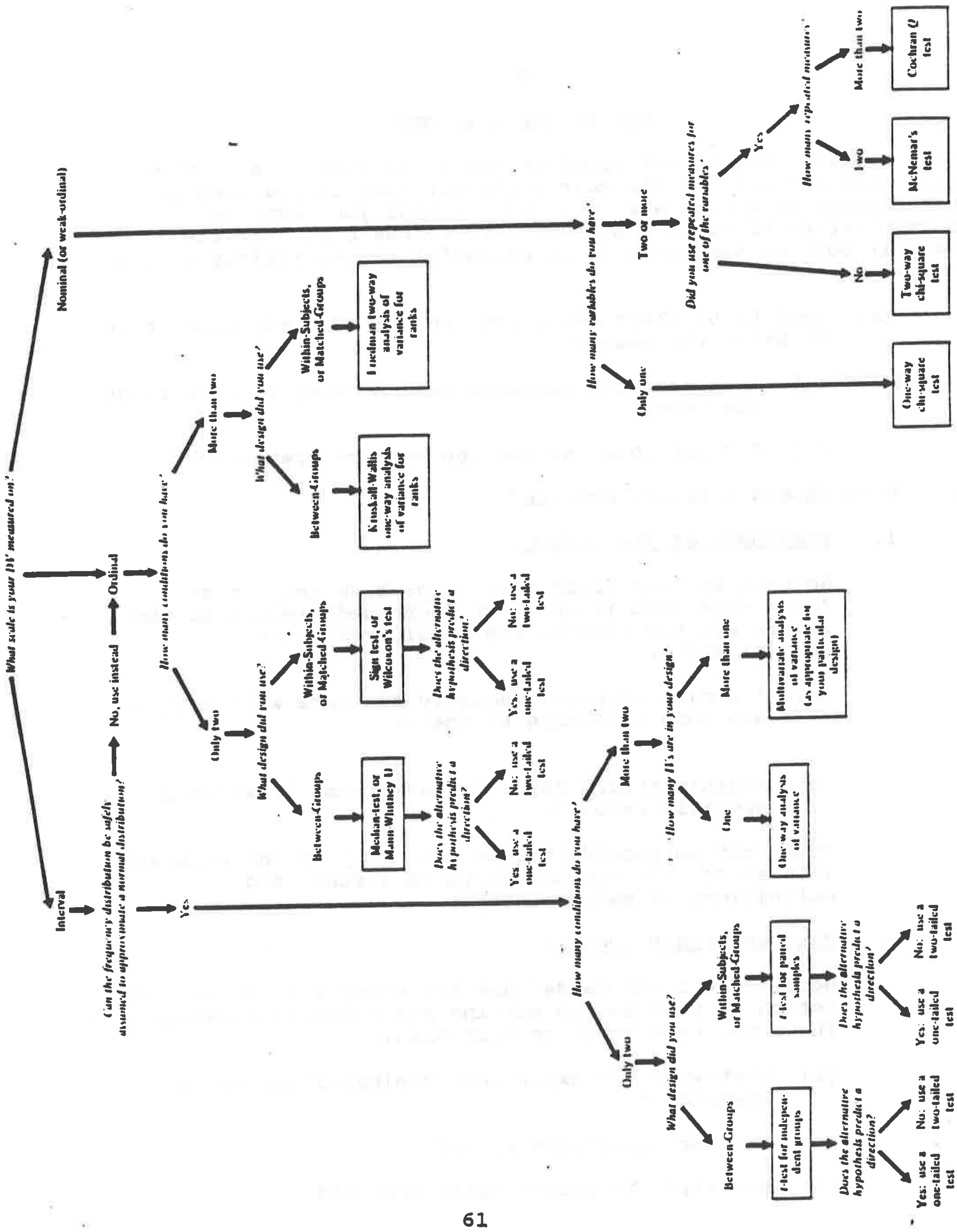
Causal Modeling — Now very similar

(5) Factor Analysis

(6) Analysis of Covariance Structures
LISREL

*The chart on Page 61 serves as a map for making decisions concerning statistical tests. The chart does not cover regression analysis; but for the means test, the chart is very appropriate.

Decision tree for choosing a statistical test.



Taken from C. Saslow. 1982. Basic Research Methods. New York: Random House.

X

THE RESEARCH REPORT

The research task is not complete until the report is written. This step may well be the most difficult because you have to communicate to others what you did and what you found out. Communication of results so that they become a part of the general body of knowledge is an essential responsibility of the investigator.

(a) Need to be clear about your audience -- you might have to write for several

(1) What does the audience want or need to know about the study?

(2) How can this information best be presented?

A. What Should a Report Contain?

1. Statement of the Problem

Go back to your first step -- problem definition -- (make sure that is what you addressed) relate to the reader why the problem was considered worth investigating.

(a) A social science scholarly audience will want to see some relevance to theory.

(b) Administrators may want background to solve a specific problem.

This section should include a summary of the relevant literature, the hypotheses to be tested, and definitions of major concepts.

2. The Research Procedure

Need to tell the reader how the study was carried out. Return to your design outline and sampling strategy (have you been loyal to your plan.)

(a) What were the experiment manipulations or the questions?

(b) How were questions scaled?

(c) How were the interviewers trained?

(d) Who were the subjects? and how were they selected?

Establish the generalization of the results

(e) What techniques were used to analyze the data? What level of confidence will be employed?

3. The Results

Rule: Give all the evidence relevant to the research question asked, whether or not the results are in accord with the investigator's views.

You are not free to choose what will be included and what will be left out.

You do decide what is relevant. How is this done?

(1) Look at the research problem and hypotheses for guidance -- if things didn't work or data could not be collected -- say so!

Remember, the research problem is seldom completely firmed up when the research problem is formulated - you may refine your statement, add new hypotheses, and investigate unforeseen relationships the original problem still remains the basic point of reference.

For every comparison that is made between groups or a relationship, the statistical significance should be reported.

In an exploratory study, the content of the report is less directed by the design statement and data analysis; the investigator must rely more on his or her own judgement. Nevertheless -- always begin with the problem statement.

4. Discussion of Implications

To convey the broader meaning of your study to your audience the implications of the results need to be elaborated.

(a) A statement of the inferences drawn from the findings in this particular situation that may apply in similar situations.

(b) Note conditions of their studies that limit the extent of legitimate generalization.

(c) Mention relevant questions that are still unanswered or new questions raised by the study.

5. The Summary

Customary to review briefly the problem, procedures, major finding, and major conclusions. In journal articles the abstract often serves as the summary. For an administrative report, an executive summary placed up front may be used.

*In some cases, a shorter version of the report may be warranted (study may not be as important as others) then: (1) minimize discussion of previous research, except the most important; (2) relate study to theoretical details in brief detail; (3) limit review of instrument and select only a few questions (more detail from author upon request); reference method of data collection; be selective on what findings you report; use tables, charts, graphs sparingly..

B. Style of Paper

Basic qualities are accuracy and clarity -- a pleasing style is a bonus.

Step 1: Decide on what information you want to tell to readers and how the various points are related. (An outline is an effective device to organize thoughts)

Step 2: Don't worry about writing style during first draft -- just write.

Step 3: Review content and give attention to style. Avoid jargon.

Step 4: Prepare tables, charts, etc. carefully. Titles should briefly state the subject matter; row and column headings should be clear and concise; everything should be clearly labeled.

Step 5: Review manuscript for clarity and meaning. Also check grammar. If possible, have a qualified colleague (friendly critic) review the report to point out unclear or illogical passages.

REFERENCES

These books on research methods have been used to pull together the material in this handbook. Saslow is an excellent introductory text while Kerlinger would be more appropriate for advanced students.

Kerlinger, F. 1986. Foundations of Behavioral Research. New York: Holt, Rinehart and Winston, Inc.

Saslow, C. 1982. Basic Research Methods. New York: Random House.

Selltiz, C., L. S. Wrightsman, and S. W. Cook. 1976. Research Methods in Social Relations. New York: Holt, Rinehart and Winston, Inc.

A number of good statistic books are available; but if you are apprehensive, you might want to try:

Rowntree, D. 1981. Statistics Without Tears. A Primer for Non-Mathematicians. New York: S. Charles Scribner's Sons.