# **RESEARCH SKILLS 4**



# Table of Contents

DATA COLLECTION METHODS	1
SOURCES OF DATA	1
PRIMARY SOURCES OF DATA	1
SECONDARY SOURCES OF DATA	2
DATA COLLECTION METHODS	2
THE QUESTIONING TECHNIQUE	4
FACE-TO-FACE INTERVIEWS	4
TELEPHONE INTERVIEWS	4
ISSUES IN DATA COLLECTION	7
ETHICS FOR THE RESEARCHER	7
ETHICS FOR RESPONDENTS	7
MEASUREMENT OF VARIABLES: OPERATIONAL DEFINITION	8
HOW VARIABLES ARE MEASURED	8
OPERATIONALIZATION OF VARIABLES	8
OPERATIONALIZATION: DIMENSIONS AND ELEMENTS	9
OPERATIONALIZING THE MULTIDIMENSIONAL CONCEPT OF ACHIEVEMENT MOTIVATION	9
DIMENSIONS AND ELEMENTS OF ACHIEVEMENT MOTIVATION	9
MEASUREMENT: SCALING, RELIABILITY, VALIDITY	11
Scales	11
RATING SCALES	12
INTERNATIONAL DIMENSIONS OF SCALING	14
SAMPLING	17
NORMALITY OF DISTRIBUTIONS	17
SAMPLING PROCESS	17
PROBABILITY SAMPLING	18
NONPROBABILITY SAMPLING	19
ISSUES OF PRECISION AND CONFIDENCE IN DETERMINING SAMPLE SIZE	21
TRADE-OFF BETWEEN CONFIDENCE AND PRECISION	22
IMPORTANCE OF SAMPLING DESIGN AND SAMPLE SIZE	22
SAMPLING AS RELATED TO QUALITATIVE STUDIES	23
THE RESEARCH PROCESS: ELEMENTS OF RESEARCH DESIGN	24
TESTING	24
EXPLORATORY STUDY	25
DESCRIPTIVE STUDY	25
CASE STUDY ANALYSIS	25
CASUAL VERSUS CORRELATIONAL RELATIONSHIP	25
EXTENT OF RESEARCHER INTERFERENCE WITH THE STUDY	26
CONTRIVED AND NONCONTRIVED STUDY SETTING	26
UNIT OF ANALYSIS: INDIVIDUALS, DYADS, GROUPS, ORGANIZATIONS, CULTURES	26
CROSS-SECTIONAL STUDIES	26
LONGITUDINAL STUDIES	26
MANAGERIAL IMPLICATIONS	27

QUANTITATIVE DATA ANALYSIS	28
GETTING THE DATA READY FOR ANALYSIS	28
GETTING A FEEL FOR THE DATA	29
MEASURES OF CENTRAL TENDENCY AND DISPERSION	29
MEASURES OF DISPERSION	29
RELATIONSHIP BETWEEN VARIABLES	29
<b>R</b> ELATIONSHIP BETWEEN <b>2</b> NOMINAL VARIABLES, $x^2$ test	30
TESTING GOODNESS OF DATA	30
QUANTITATIVE DATA ANALYSIS: HYPOTHESIS TESTING	32
TYPE I AND II ERRORS, AND STATISTICAL POWER	32
CHOOSING THE APPROPRIATE STATISTICAL TECHNIQUE	32
TESTING A HYPOTHESIS ABOUT A SINGLE MEAN	32
TESTING HYPOTHESES ABOUT TWO RELATED MEANS	33
TESTING HYPOTHESES ABOUT TWO UNRELATED MEANS	33
TESTING HYPOTHESES ABOUT SEVERAL MEANS	33
REGRESSION ANALYSIS	34
STANDARDIZED REGRESSION COEFFICIENTS	34
REGRESSION WITH DUMMY VARIABLES	34
MULTICOLLINEARITY	34
OTHER MULTIVARIATE TESTS AND ANALYSES	35
DATA WAREHOUSING, DATA MINING, AND OPERATIONS RESEARCH	36
SOME SOFTWARE PACKAGES USEFUL FOR DATA ANALYSIS	37
THE RESEARCH REPORT	38
Тне перопт	38
THE WRITTEN REPORT	38
ORAL PRESENTATION	41
FINK 1	43
Surveys	43
FIND 2	46
THE SURVEY FORM	46
FINK 3	49
PRACTICAL CONCERNS	49
FINK 4	51
Sampling	51
FINK 5	55
SURVEY DESIGN	55
	50
	58
PRESENTING THE SURVEY RESULTS	58

# Data collection methods

#### Sources of data

- **Primary data** refer to information obtained first-hand by the researcher on the variables of interest for the specific purpose of the study.
  - Examples are individuals, focus groups, panels, and the internet from time to time.
- Secondary data refer to information gathered from sources that already exist.
  - Examples are company records, archives, government publications, industry analyses, websites, the internet, etc.

#### Primary sources of data

- **Individuals** provide information when interviewed, administered questionnaires, or observed.
- **Focus groups** consist typically of 8 10 members with a moderator leading the discussions for about 2 hours.
  - Meet for a one-time group session
  - Members are generally chosen on the basis of their expertise in the topic of focus.
  - Unstructured and spontaneous responses are expected to reflect genuine opinions, ideas, and feelings of the members.
  - Relatively inexpensive and can provide fairly dependable data within a short time frame.
  - The moderator introduces the topic, observes, and takes note whilst never becoming an integral part of the discussion. Also ensures that all members participate in the discussion and no member dominates the group.
  - Data obtained through these homogeneous group members provides only qualitative information.
  - Members are not selected scientifically, thus their opinions cannot be considered to be truly representative.
  - Focus groups are used for:
    - Exploratory studies
    - Making generalizations based on the information generated by them
    - Conducting sample surveys
  - Online focus groups are also common thanks to videoconferencing.
- Panels
  - Meet more than once, unlike focus groups.
  - Individuals are randomly chosen.
  - **Panel studies** are panels composed of expert members in a given field of study.
  - Static panels: the same members serve on the panel over extended periods of time.
    - Advantages:

- Offers a good and sensitive measurement of the changes that take place between two points in time.
- Disadvantages:
  - Panel members could become so sensitized to the changes that their opinions might no longer be representative of the rest of the population.
- **Dynamic panels**: the panel members change from time to time as various phases of the study are in progress.
  - Advantages and disadvantages are the reverse of static panels.
- **Delphi Technique** is a forecasting method that uses a cautiously selected panel of experts in a systematic, interactive manner.
  - Experts answer questionnaires in 2 or more rounds.
    - In the 1<sup>st</sup> round they are asked to answer a series of questions on the likelihood of a future scenario or any other issue about which there is unsure or incomplete knowledge.
      - The contributions are collected, summarized, and fed back to form a 2<sup>nd</sup> round questionnaire.
      - Process goes on until it is stopped by the researcher.
- **Unobtrusive methods**, trace measures, originate from a primary source that doesn't involve people.

### Secondary sources of data

- **Secondary data** refer to information gathered by someone other than the researcher conducting the current study.
  - This data can be internal or external to the organization and accessed through the internet or perusal of recorded or published information.
  - Data can be used for forecasting sales by constructing models based on past sales figures, and through extrapolation.
  - Sources can be books, periodicals, publications, economic indicators, census data, statistical abstracts, databases, media, annual reports, etc.
  - Mostly qualitative data.
  - Advantage:
    - Savings in time and costs of acquiring information.
  - Disadvantage:
    - Secondary data as sole source can become obsolete.

### Data collection methods

Data can be collected in a variety of ways, settings, and from different sources. These methods include:

- Interviewing:
  - Advantage: flexibility in terms of adapting, adopting, and changing the questions.

- Especially useful during the exploratory stages of research.
- Important to train the interviewers and good training decreases bias.
- Interviews may be structured or unstructured, face-to-face, by telephone, or online
  - Unstructured interviews: the interviewed doesn't enter the interview setting with a planned sequence of questions to be asked.
    - Objective is to bring some preliminary issues to the surface in order to determine what needs further investigation.
  - **Structured interviews**: conducted when it's known at the outset what information is needed.
    - Questions are asked either personally, by phone, or PC.
    - Questions are likely to focus on factors that surfaced during the unstructured interviews and relevant to the problem.
    - Interviewer must comprehend the purpose and goal of each question to be able to recognize probable response.
    - Once finished the information is tabulated and the data analyzed.
  - Bias refers to errors or inaccuracies in the data collected. It could be introduced by the interviewer, the interviewee, or the situation.
    - Bias could be situational as well:
      - *Nonparticipation* can bias data inasmuch as the responses of the participants may be different from those of nonparticipants.
      - Differing trust levels and rapport established.
      - The physical setting of the interview.
    - Interviewer can reduce bias by being consistent with the questioning mode, by not distorting or falsifying the information received, and by not influencing the responses.
    - Biased data will be obtained when respondents are extremely busy or not in a good humor.
    - Personality of interviewer, introductory sentence, inflection of voice, etc. also affect bias.
    - CAI Computer Assisted Interviews. There are 2 types:
      - CATI Computer Assisted Telephone Interview
        - Responses from people can be obtained all over the world
      - CAPI Computer Assisted Personal Interview
        - Involves big investments in hardware and software
        - VCS Voice Capture System allows the computer to capture the respondents' answers
    - Advantages of computer-aided survey services:

- Can start analyzing the data even as the field survey is in progress
- Data can be automatically cleaned up and errors fixed while they are being collected
- Bias due to ordering questions in a particular way can be eliminated
- $\circ$   $\;$  Skip patterns can be programmed into the process  $\;$
- Questions can be customized to incorporate the respondent's terminology
- Advantages of software packages automatic indexing of data can be done with these programs.
  - Indexing such that specific responses are coded in a particular way
  - Retrieval of data with a fast search speed
- The researcher can establish rapport by being pleasant, sincere, sensitive, and non-evaluative.

## The questioning technique

- Funneling: The transition from broad to narrow themes, from open-ended to focused questions, is called the funneling technique
- Unbiased questions
- Clarifying issues
- Helping the respondents to think through issues
- Taking notes

#### Face-to-face interviews

- Advantages: researcher can adapt the question as necessary, clarify doubt, ensure the responses are properly understood, and pick up nonverbal cues from the respondent.
- Disadvantages: geographical limitations, vast resources needed, costs of training, and making respondents uneasy due to anonymity concerns.

### **Telephone interviews**

- Advantages: for interviewer number of different people can be reached in a short time, for respondent – eliminates any discomfort in having to face the interviewer, might feel less uncomfortable disclosing personal information.
- Disadvantages: respondent could unilaterally terminate the interview without warning or explanation, researcher will not be able to see the respondent to read nonverbal ques.

### Administering questionnaires:

- A **questionnaire** is a pre-formulated written set of questions to which respondents record their answers.
  - Can be administered personally, mailed to respondents, or electronically distributed.
- Personally administered, sent through mail, electronically administered, observation of individuals and events, other motivational techniques such as projective tests.
- Advantage: obtaining data more efficiently in terms of research time, energy, and costs.

- **CATI**: Computer-assisted telephone interviewing help both the interviewing process and with preparing and administering questionnaires electronically.
- Guidelines for questionnaire design
  - Principles of wording
    - The appropriateness of the content
    - How questions are worded/language use
    - Type and form of questions asked
    - Sequencing of the questions
    - Personal data sought from the respondents
  - Content and purpose of questions
  - Language and wording of the questionnaire should approximate the level of understanding of the respondents
- Type and form of questions:
  - **Open-ended questions** allow respondents to answer them in any way they choose.
  - **Closed questions** ask the respondents to make choices among a set of alternatives given by the researcher.
    - All items using a nominal, ordinal, Likert, or ratio scale
    - Help respondents make quick decisions to choose among several alternatives
    - Help researcher code the information more easily for analysis
  - It is advisable to include some negatively worded questions.
  - **Double-barreled questions** lend themselves to different possible responses to its subparts. These questions should be avoided.
  - Ambiguous questions have built-in bias inasmuch as different respondents might interpret such items on the questionnaire differently.
  - Recall-dependent questions require respondents to recall experiences from the past, which might have bias due to hazy memory.
  - Leading questions, which signal and pressure a respondent to answer a certain way, shouldn't be used to avoid bias.
  - **Loaded questions** are another type of bias in questions that are phrased in an emotionally charged manner.
  - Questions should not be worded such that they elicit socially desirable responses.
  - The sequence of questions in the questionnaire should be such that the respondent is led from questions of a general nature to more specific ones.
  - Classification data, aka demographics. It is best not to ask the name or income of respondent, but if it must happen, this should be asked toward the end of the questionnaire.
- Principles of measurement
  - The interval and ratio scales should be used in preference to nominal or ordinal scales
  - Once data are obtained, the goodness of data should be assessed through tests of validity and reliability

- Validity establishes how well a technique, instrument, or process measures a particular concept
- Reliability indicated how stably and consistently the instrument taps the variable
- Data has to be obtained in a manner that makes for easy organization and coding
- Questionnaire's should have:
  - A good introduction
  - Organized questions, with instructions and guidance, and good alignment
  - Limit information on income and other sensitive personal data
  - Open-ended questions at the end
  - Conclude the questionnaire with courtesy
- Pretest the questionnaire
- Other methods of data collection:
  - Observational studies
    - Nonparticipant-observer
    - Participant-observer
    - **Structured observational study**: observer has a predetermined set of categories of activities or phenomena to be studied
    - Unstructured observational study: observer will record practically everything that is observed. This is claimed to be the hallmark of qualitative research
    - Advantages of observational studies:
      - Data is more reliable
      - Easier to note the effects of environmental influences on specific outcomes
      - Easier to observe certain groups of individuals
  - Disadvantages of observational studies:
    - Necessary for the observer to be physically present
    - This method is slow, tedious, and expensive.
    - Observer fatigue can easily set in
    - Cognitive thought process of individuals can't be captured
    - Observers have to be trained in what and how to observe, and avoid bias
      - Bias can come in the form of recording errors, memory lapses, errors in interpreting activities, behaviors, events, and nonverbal cues.
  - Projective methods
    - Word association techniques are often used to get at true attitudes and feelings
    - **Thematic apperception tests** (TAT) call for the respondent to weave a story around a picture that is shown.
      - Several need patterns and personality characteristics of employees can be traced.
    - **Inkblot tests**, another form of motivational research, use colored inkblots that are interpreted by the respondents.

Data can be collected in the natural environment of the workplace, in an artificial lab setting where there are controlled variables, it can be gathered in the homes of respondents, on the street, malls, or online.

When translating an instrument ensure the translation is done accurately and then translated back to the original language to ensure vocabulary equivalence.

#### Issues in data collection

- Response equivalence
  - Ensured by adopting uniform data collection procedures
- Timing of data collection
  - Data collection should be completed within acceptable time frames
- Status of the individual collecting the data

### Ethics for the researcher

- 1. Treating the information given by the respondent as strictly confidential and guarding their privacy is one of the primary responsibilities of the researcher.
- 2. The researcher should not misrepresent the nature of the study to subjects, especially in lab experiments.
- 3. Personal or seemingly intrusive information should not be solicited, and if it is absolutely necessary for the project, it should be tapped with high sensitivity to the respondent, offering specific reasons.
- 4. Whatever the nature of the data collection method, the self-esteem and self-respect of the subjects should never be violated.
- 5. No one should be forced to respond to the survey and if someone does not want to avail themselves of the opportunity to participate, the individual's desire should be respected. Informed consent of the subjects should be the goal of the researcher.
- 6. Nonparticipant-observers should be as uninstructive as possible.
- 7. In lab studies, the subjects should be debriefed with full disclosure of the reason for the experiment after they have participated in the study.
- 8. Subjects should never be exposed to situations where they could be subject to physical or mental harm.
- 9. There should be absolutely no misrepresentation or distortion in reporting the data collected during the study.

### Ethics for respondents

- 1. The subject, once having exercised the choice to participate in a study, should cooperate fully in the tasks ahead, such as responding to a survey or taking part in an experiment.
- 2. The respondent also has an obligation to be truthful and honest in the responses.

# Measurement of variables: operational definition

#### How variables are measured

Measurement is the assignment of numbers or other symbols to *characteristics* (*attributes*) of objects according to a pre-specified set of rules.

**Objects include:** 

- Persons
- Business units
- Companies
- Countries
- Items
- Etc.

Examples of characteristics of objects:

- Arousal-seeking tendency
- Achievement motivation
- Organizational effectiveness
- Shopping enjoyment
- Length
- Weight
- Ethnic diversity
- Service quality
- Conditioning effects
- Taste
- Etc.

There are at least two types of variables: one lends itself to objective and precise measurement; the other is more nebulous and does not lend itself to accurate measurement because of its abstract and subjective nature.

### Operationalization of variables

One technique to tap the previously mentioned variables is to reduce abstract notions or concepts to observable behavior and/or characteristics. Reduction of abstract concepts to render them measurable in a tangible way is called **operationalizing** the concepts.

Operationalizing is done by looking at the behavioral dimensions, facets, or properties denoted by the concept. These are then translated into observable and measurable elements to develop an index of measurement of the concept.

- 1. Come up with a definition of the construct you want to measure
- 2. Think about the content of the measure; an instrument that actually measures the concept that you want to measure has to be developed
- 3. A response format is needed (e.g.7-point rating)

4. The validity and reliability of the measurement has to be assessed

### Operationalization: dimensions and elements

A valid measure of need for cognition contains 34 items even though need for cognition is a unidimensional construct.

Operationalizing the multidimensional concept of achievement motivation After defining the construct, the next step in the process of measuring abstract constructs is to go through the literature to find out whether there are any existing measures of the concept.

Scientific journals and scale handbooks are important sources of existing measures.

Empirical articles published in academic journals provide a detailed description of how specific constructs were measured.

#### Dimensions and elements of achievement motivation

Figure 6.2 Dimensions (D) and elements (E) of the concept (C) "achievement motivation"



Operationalization is likely to:

- Exclude some of the important dimensions and elements arising from failure to recognize or conceptualize them
- Include certain irrelevant features, mistakenly thought to be relevant

Operationalization consists of the reduction of the concept from its level of abstraction, by breaking it into dimensions and elements.

An operationalization does not describe the correlated of the concept.

Operationalizing a concept does not consist of delineating the reasons, antecedents, consequences, or correlates of the concept. It describes its observable characteristics in order to be able to measure the concept.

Only a well-developed instrument, which has been operationalized with care, will be accepted and frequently used by other researchers.

In conducting transnational research, it is important to remember that certain variables have different meanings and connotations in different cultures.

# Measurement: scaling, reliability, validity

There are two main categories of attitudinal scales, the rating scale and the ranking scale.

- **Rating scales** have several response categories and are used to elicit responses with regard to the object, even, or person studied.
- **Ranking scales** make comparisons between or among objects, events, or persons and elicit the preferred choices and ranking among them.

## Scales

A **scale** is a tool or mechanism by which individuals are distinguished as to how they differ from one another on the variables of interest to the study.

- The scale or tool may be a gross one (it only broadly categorizes individuals on the variables with varying degrees of sophistication)
- 4 basic types of scales (with the degree of sophistication increasingly progressively from nominal to ratio):
  - **Nominal scale**: one that allows the researcher to assign subjects to certain categories or groups.
    - The information that can be generated from nominal scaling is the calculation of the % or frequency.
    - Gives basic, categorical, gross information.
  - **Ordinal scale**: not only categorizes the variables in such a way as to denote differences among the various categories, it also rank-orders the categories in some meaningful way.
    - With any variable for which the categories are the be ordered according to some preference, the ordinal scale would be used.
    - The preference would be ranked, numbered, etc.
    - Goes beyond differentiating the categories to providing information on how respondents distinguish them by rankordering them.
    - Does not give any indication of the magnitude of the differences among the ranks.
  - **Interval scale**: allows us to perform certain arithmetical operations on the data collected from the respondents.
    - Lets us measure the distance between any two points on the scale, which helps to compute the means and the standard deviations of the responses on the variables.
    - The interval scale not only groups individuals according to certain categories and taps the order of these groups, it also measures the magnitude of the differences in the preferences among the individuals.
    - Taps the differences, the order, and the equality of the magnitude of the differences in the variable.
    - It is a more powerful scale than the nominal and ordinal scale and has for its measure of central tendency the arithmetic mean.

- Its measures of dispersion are the range, the standard deviation, and the variance.
- **Ratio scale**: not only measures the magnitude of the differences between points on the scale, it also taps the proportions in the differences.
  - Overcomes the disadvantage of the arbitrary origin point of the interval scale, in that it has an absolute zero point.
  - The most powerful of the 4 scales because it has a unique zero origin and subsumes all the properties of the other three scales.
  - Some variables, such as gender, can be measured only on the nominal scale, while others, such as temperature, can be measured on various scales.
  - Whenever possible the more powerful scale should be used.

## Rating scales

The following rating scales are often used in organizational research:

- Dichotomous scale
  - Used to elicit a yes or no answer
- Category scale
  - Uses multiple items to elicit a single response
- Semantic differential scale
  - Used to assess respondents' attitudes toward a particular brand, advertisement, object, or individual.
  - Responses can be plotted to obtain a good idea of their perceptions
  - Treated as an interval scale
- Numerical scale
  - Similar to the semantic differential scale, with the difference that numbers on a 5 or 7-point scale are provided, with bipolar adjectives at both ends
  - o An interval scale

#### Itemized rating scale

- $\circ~$  A 5 or 7-point scale with anchors is provided for each item and the respondent states the appropriate number on the side of each item
- o Interval scale
- Provides the flexibility to use as many points in the scale as considered necessary
- Also possible to use different anchors
- When a neutral point is provided, it is a balanced rating scale, and when it is not, it is an **unbalanced rating scale**.
- Frequently used in business research, since it adapts itself to the number of points the researcher wishes to use.
- Likert scale
  - Designed to examine how strongly subjects agree or disagree with statements on a 5-point scale
  - The summated approach is widely used; thus the Likert scale is referred to as a summated scale
  - o Treated as interval scales

#### • Fixed or constant sum rating scale

- Respondents are asked to distribute a given number of points across various items.
- Ordinal scale
- Staple scale
  - Simultaneously measures both the direction and intensity of the attitude toward the items under study.
  - $\circ~$  The characteristic of interest to the study is placed at the center with a numerical scale ranging from +x to -x
  - Interval scale

### • Graphic rating scale

- Graphical representation helps the respondents to indicate on this scale their answers to a particular question by placing a mark at the appropriate point on the line
- Ordinal scale
- Easy to respond to
- Brief descriptions on the scale points are meant to serve as a guide in locating the rating rather than representing discrete categories.
- $\circ$   $\,$  Faces scale is also a graphic rating scale  $\,$

#### Consensus scale

- A panel of judges selects certain items, which in its view measure the relevant concept.
- $\circ\;$  Items are chosen particularly based on their pertinence of relevance to the concept.
- Thurstone Equal Appearing Interval Scale: a concept is measured by a complex process followed by a panel of judges
- Rarely used for measuring organizational concepts because of the time necessary to develop it

#### Other scales

- Advanced scaling methods such as multidimensional scaling, where objects, people, or both, are visually scaled, and a conjoint analysis is performed.
  - Provides visual image of the relationships in space among the dimensions of a construct.
- Likert or some form of numerical scale is the one most frequently used to measure attitudes and behaviors in organizational research.

#### Ranking scales

- Used to tap preferences between two or among more objects or items (ordinal in nature)
- Such rankings may not give definitive clues to some of the answers sought
- Alternative methods used are paired comparisons, forced choice, and the comparative scale

### Paired comparison

- Used when respondents are asked to choose between 2 objects at a time
- Helps to assess preferences
- Forced choice

- Enables respondents to rank objects relative to one another, among the alternatives provided.
- Easier for the respondents

#### Comparative scale

- Provides a benchmark or a point of reference to assess attitudes toward the current object, event, or situation under study.
- Nominal data lend themselves to dichotomous or category scales
- o Ordinal data to any one of the ranking scales
- Interval or interval-like data to the other rating scales
- The semantic differential and the numerical scale are not interval scales
- Rating scales are used to measure most behavioral concepts
- Ranking scales are used to make comparisons or rank the variables that have been tapped on a nominal scale

# International dimensions of scaling

Different cultures reach differently to issues of scaling.

The use of better instruments will ensure more accuracy in results, which in turn will enhance the scientific quality of the research.

Item analysis: carried out to see if the items in the instrument belong there or not.

- The means between the high-score group and the low-score group are tested to detect significant differences through the t-values.
- **Reliability** is a test of how consistently a measuring instrument measures whatever concept it is measuring.
  - Reliability measure indicates the extent to which it is without bias and hence ensures consistent measurement across time and across the various items in the instrument.
  - The ability of a measure to remain the same over time is indicative of its stability and low vulnerability to changes in the situation
  - Two tests of stability are:
    - **Test-retest reliability**: The reliability coefficient is obtained by repetition of the same measure on a second occasion.
    - **Parallel-form reliability**: When responses on 2 comparable sets of measures tapping the same construct are highly correlated
  - The **internal consistency** of measures is indicative of the homogeneity of the items in the measure that tap the construct.
  - The **interitem consistency reliability** is a test of the consistency of respondents' answers to all the items in a measure
  - **Splif-half reliability** reflects the correlations between two halves of an instrument. The estimates will vary depending on how the items n the measure are split into two halves.
- **Validity** is a test of how well an instrument that is developed measures the particular concept it is intended to measure
  - Several types of validity tests are used

- **Content validity**: ensures that the measure includes an adequate and representative set of items that tap the concept
- the more the scale items represent the domain or universe of the concept being measured, the greater the content validity.
- Face validity indicates that the items that are intended to measure a concept, do, on the face of it, look like they measure the concept.
  - Face validity is considered by some a basic and minimum index of content validity
- **Criterion-related validity** is established when the measure differentiates individuals on a criterion it is expected to predict.
  - Can be done by establishing concurrent validity or predictive validity
  - **Concurrent validity** is established when the scale discriminates individuals who are known to be different
  - **Predictive validity**: indicated the ability of the measuring instrument to differentiate among individuals with reference to a future criterion
- **Construct validity** testifies to how well the results obtained from the use of the measure fit the theories around which the test is designed.
- Convergent validity is established when the scores obtained with 2 different instruments measuring the same concept are highly correlated.
- Discriminant validity is established when, based on theory, 2 variables are predicted to be uncorrelated, and the scores obtained by measuring them are indeed empirically found to be so.
- Some of the ways the above forms of validity can be established are through:
  - Correlational analysis
  - Factor analysis, a multivariate technique that confirms the dimensions of the concept that have been operationally defined
  - The multitrait, multimethod matrix of correlations derived from measuring concepts by different forms and different methods

It is not unusual for 3 or more equally good measures to be developed for the same concept. When more than one scale exists for any variable, it is preferable to use the measure that has better reliability and validity and is also more frequently used.

In a **reflective scale**, the items are expected to correlate. Each item in a reflective scale is assumed to share a common basis in the underlying construct of interests.

A **formative scale** is used when a construct is viewed as an explanatory combination of its indicators.

The job description index includes 5 dimensions and 72 items. These 72 items are not necessarily related to each other, because the 5 dimensions they represent do not necessarily hang together.

A scale that contains items that are not necessarily related is called a formative scale. Formative scales are used when a construct is viewed as an explanatory combination of its indicators.

The job description index is formative in nature.

A good (valid) formative scale is one that represents the entire domain of the construct.

While it makes sense to test the interitem consistency of reflective scales, it does not make sense to test the interitem consistency of formative scales.

# Sampling

If data are not collected from the people, events, or objects that can provide the correct answers, the survey will be in vain.

The **population** refers to the entire group of people, events, or things of interest that the researcher wishes to investigate.

An **element** is a single member of the population. The census is a count of all elements in the human population.

A **sample** is a subset of the population.

The **sampling unit** is the element or set of elements that is available for selection in some stage of the sampling process.

A **subject** is a single member of the sample.

The characteristics of the population such as  $\mu$  (the population mean),  $\sigma$  (*the population standard deviation*), and  $\sigma^2$  (the population variance) are referred to as its parameters. The central tendencies, the dispersions, and other statistics in the sample of interest are treated as approximations of the central tendencies, dispersions, and other parameters of the population.

Reasons for using samples rather than the entire population is because samples are likely to produce more reliable results mostly because fatigue is reduced and fewer errors occur.

### Normality of distributions

The sample has to be chosen so that the distribution of the characteristics of interest follows the same pattern of normal distribution in the sample as it does in the population.

- Central limit theorem: the sampling distribution of the sample mean is normally distributed. As the sample size n increases, the means of the random samples taken from practically any population approach a normal distribution with mean  $\mu$  and standard deviation  $\sigma$ .
- Important issues in sampling are the sample size and the sampling design.
- Generalizability is one of the hallmarks of scientific research.

#### Sampling process

Sampling is the process of selecting a sufficient number of the right elements from the population. Major steps in sampling include:

- Define the population
  - Target population must be defined in terms of elements, geographical boundaries, and time.
- Determine the sample frame

- When the sampling frame doesn't exactly match the population coverage error occurs.
- The researcher should deal with this error by either redefining the target population in terms of the sampling frame, screening the respondents, or adjusting the collected data by a weighting scheme.

#### • Determine the sampling design

- **Probability sampling**: the elements in the population have some known, non-zero chance.
- **Nonprobability sampling**: the elements do not have a known or predetermined chance of being selected as subjects.
- Determine the appropriate sample size
  - Factors affecting decision on sample size
    - The research objective
    - The extent of precision desired
    - The acceptable risk in predicting that level of precision
    - The amount of variability in the population itself
    - The cost and time constraints
    - The size of the population itself
- Execute the sampling process
  - Non-response error exists to the extent that those who did respond to your survey are different from those who did not on characteristics of interest in your study. Important sources:
    - Not-at-homes
    - Refusals
      - Rate of refusals depends on length of survey, collection method, patronage of the research.

### **Probability sampling**

**Unrestricted probability sampling,** aka **simple random sampling**, is when every element in the population has a known and equal chance of being selected as a subject.

**Complex probability sampling**, aka **restricted probability**, offer a viable, and sometimes more efficient, alternative to unrestricted design. Efficiency is improved since more information can be obtained for a given sample size.

The **systematic sampling** design involves drawing every nth element in the population starting with randomly chosen element between 1 and n.

**Stratified random sampling** involves a process of stratification, segregation, followed by random selection of subjects from each stratum. The population is first divided into mutually exclusive groups that are relevant, appropriate, and meaningful in the context of the study.

- **Proportionate stratified random sampling**: members represented in the sample from each stratum will be proportionate to the total number of elements in the respective strata.
- **Disproportionate stratified random sampling**: the number of subjects from each stratum would be altered, while keeping the sample size unchanged.

- Made either when strata are too small or large, or when there is more variability suspected within a particular stratum
- $\circ$   $\,$  Also done when it's easier, simpler, and less expensive to collect data  $\,$

Cluster samples are samples gathered in groups or chunks of elements that are natural aggregates of elements in the population.

- In **cluster sampling**, the target population is first divided into clusters, then a random sample of clusters is drawn from each cluster.
  - **Area sampling**: clusters consist of geographical areas such as countries, city blocks, or particular boundaries within a locality
    - Less expensive
  - $\circ \quad \text{Unit costs are lower} \\$
  - $\circ$   $\;$  Exposes itself to greater bias
  - Least generalizable of all probability sampling designs, because naturally occurring clusters do not contain heterogeneous elements
  - $\circ$   $\;$  Not very common in organizational research
  - Offers convenience
- Multistage cluster sampling: cluster sampling done in several stages.

**Double sampling**: A sampling design where initially a sample is use in a study to collect some preliminary information of interest, and later a subsample is used to examine the matter in more detail.

Nonprobability sampling: the elements in the population do not have any probability attached to their being chosen as sample subjects.

• Sometimes only way to obtain data.

**Convenience sampling** refers to the collection of information from members of the population who are conveniently available to provide it.

- Most often used during the exploratory phase
- Best way of getting basic information quickly and efficiently

**Purposive sampling**: obtaining information from specific target groups. Two major types of purposive sampling:

- **Judgment sampling** involves the choice of subjects who are most advantageously placed to provide the required information
  - Judgment sampling design is used when limited number or category of people have the information sought
  - $\circ$  May curtail the generalizability of the findings, due to sample of experts conveniently available to us
- **Quota sampling** ensures that certain groups are adequately represented in the study through the assignment of a quota
  - Can be considered a form of proportionate stratified sampling
  - The sample may not be totally representative of the population, however the convenience it offers in terms of effort, cost, and times makes it attractive for some research efforts
  - Becomes a necessity when a subset of the population is underrepresented in the organization

 Basically stratified samples from which subjects are selected nonrandomly

#### Simple random sampling

Where cost is a primary consideration, and the number of elements in the population is very large and/or geographically dispersed, this sampling method may not be the most desirable, because it could become expensive.

#### Stratified random sampling

Most efficient sampling design, is a good choice when differentiated information is needed.

#### Systematic sampling

If the sampling frame is large, and a listing of the elements is conveniently available in one place, this approach will offer the advantages of ease and quickness in developing the sample.

#### Note

Systematic sampling is inadvisable where systematic bias can be anticipated to be present. Systematic sampling will have to be avoided in cases where known biases are possible.

#### **Cluster sampling**

Most useful when a heterogeneous group is to be studied at one time.

#### Area sampling

Best suited when the goal of the research is confined to a particular locality or area.

#### Double sampling

Provides added information at minimal additional expenditure.

#### Convenience sampling

Nonprobability design, which isn't generalizable, used to obtain some quick information and get a feel for the phenomenon or variables of interest.

#### Note

Convenience sampling should be resorted to in the interest of expediency, with the knowledge that the results aren't generalizable.

#### Judgment sampling

Used where the collection of specialized informed inputs on the topic area researched is vital, and the use of another sampling design wouldn't offer opportunities to obtain the specialized information.

#### Quota sampling

Allows for the inclusion of all groups in the system researched.

#### Note

Decisions on which design to use depend on many factors, including:

- Extent of prior knowledge in the area of research
- The main objective of the study
- Cost considerations

#### Issues of precision and confidence in determining sample size

A reliable and valid sample should enable us to generalize the findings from the sample to the population under investigation.

- Sample statistics should be reliable estimates and reflect the population parameters as closely as possible within a narrow margin of error.
- No sample statistic is going to be exactly the same as the population parameter.
- **Precision** refers to how close our estimate is to the true population characteristic.
  - Precision is a function of the range of variability in the sampling distribution of the sample mean.
  - Standard error is calculated by:  $S_{\overline{x}} = \frac{S}{\sqrt{n}}$  where S is the standard deviation, n is the sample size, and  $S_{\overline{x}}$  indicates the standard error or precision offered by the sample
  - $\circ$  Standard error varies inversely with the √ of the sample size.
  - $\circ$   $\,$  The smaller the variation in the population, the smaller the standard error.
- **Confidence** denotes how certain we are that our estimates will really hold true for the population.
  - Reflects the level of certainty with which we can state that our estimates of the population parameters will hold true.
  - $\circ~95\%$  confidence is the conventionally accepted level for most business research.

#### Sample data, precision, and confidence estimation

We do an interval estimation to ensure a relatively accurate estimation of the population parameter. The standard error  $S_{\overline{x}}$  and the % or level of confidence we require will determine the width of the interval, represented by:  $\mu = X \pm KS$  where K is the t statistic for the level of confidence desired.

Critical values for t in any statistics book:

- 90% = K value of 1.645
- 95% = K value of 1.96
- 99% = K value of 2.576

The sample size, n, is a function of:

- the variability in the population
- precision or accuracy needed
- confidence level desired
- type of sampling plan used

#### Trade-off between confidence and precision

If the sample size, n, cannot be increased, then the only way to maintain the same level of precision is to forsake the confidence with which we can predict our estimates.

Necessary for researchers to consider at least 4 aspects while making decisions on the sample size needed to do the research:

- How much precision is really needed in estimating the population characteristics of interest
- How much confidence is really needed?
- To what extent is there variability in the population on the characteristics investigated?
- What is the cost-benefit analysis of increasing the sample size?

#### Sample data and hypothesis testing

Sample data can also be used to test hypotheses about population values.

$$H_0: \mu A - \mu B =$$

The alternate hypothesis of differences will be stated non-directionally  $H_0: \mu A - \mu B \neq 0$ 

The t statistic can be calculated for testing the hypothesis:

$$t = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - (\mu 1 - \mu 2)}{S_{\overline{X}_1} - S_{\overline{X}_2}}$$

$$S_{\overline{X}_1} - S_{\overline{X}_2} = \sqrt{\frac{n1s\frac{2}{1} + n2s\frac{2}{2}}{(n1+n2-2)}} \begin{pmatrix} 1 & & 1\\ - & + & -\\ n_1 & & n_2 \end{pmatrix}$$

Sample data can be used not only for estimating the population parameters, but also for testing hypotheses about population values, population correlations, etc.

Correction formula:

$$S\overline{X} = \frac{S}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$$

Where N is the total number of elements in the population, n is the sample size to be estimates,  $S\overline{X}$  is the standard error of the estimate of the mean, and S is the standard deviation of the sample mean.

#### Importance of sampling design and sample size

If the important sampling design is not used, a large sample size will not allow the findings to be generalized to the population. Unless the sample size is adequate for the desired level of precision and confidence, no sampling design will be useful to the researcher in meeting the objectives of the study. Sampling decisions should consider both sampling design and size. Another point to consider is whether statistical significance is more relevant than practical significance.

Roscou proposes the following rules for determining sample size:

- 30 > Sample sizes < 500 = appropriate for most research
- where samples are to be broken into subsamples; minimum sample size of 30 for each category is necessary
- in multivariate research, the sample size should be several times as large as the number of variables in the study
- for simple experimental research with tight experimental controls successful research is possible with samples as small as 10 in size

Efficiency in sampling is attained when the sample size could be reduced, or for a given sample size, n, the level of precision could be increased.

- A stratified random sampling plan is often the most efficient, and a disproportionate stratified random sampling design has been shown to be more efficient that a proportionate sampling design.
- Cluster sampling is less efficient than simple random sampling because there is more homogeneity among the subjects in the clusters.
- Multistage cluster sampling is more efficient than single-stage cluster sampling when there is more heterogeneity found in the earlier stages.
- There is often a trade-off between time and cost efficiencies and precision efficiencies.

### Sampling as related to qualitative studies

- Qualitative sampling begins with precisely defining the target population.
- Purposive sampling is one technique that is often employed in qualitative investigation: subjects are selected on the basis of expertise.
  - **Grounded theory** expresses the idea that theory will emerge from data through an interactive process that involves repeated sampling, collection of data, and analysis of data until theoretical saturation is reached.
    - Theoretical saturation is reached when no new information about the subject emerges in repeated cases.

# The research process: elements of research design

Issues relating to decisions regarding the purpose of the study, it's location, the type it should conform to, the extent to which it is manipulated and controlled by the researcher, its temporal aspects, and the level at which the data will be analyzed, are integral to research design.

Decisions have to be made as to the type of sample to be used, how the data will be collected, how variables will be measured, and how they will be analyzed to test the hypotheses.



Each component of the research design offers several critical choice points.

The more sophisticated and rigorous the research design is, the greater the time, costs, and other resources expended on it will be.

#### Testing

Studies may be either exploratory in nature or descriptive, or may be conducted to test hypotheses.

**Case study** is an examination of studies done in other similar organizational situations. It is also a method of solving problems or understanding phenomena of interest. The nature of the study depends on the stage to which knowledge about the research topic has advanced.

### Exploratory study

An **exploratory study** is undertaken when not much is known about the situation at hand, or no information is available on how similar problems or research issues have been solved in the past.

Extensive preliminary work needs to be done to gain familiarity with the phenomena before developing a model. Exploratory studies are undertaken to better comprehend the nature of the problem.

Doing a study for the first time in a particular organization does not make the research exploratory in nature; only when knowledge is scant and deeper understanding is sought does the study become exploratory.

#### **Descriptive study**

A **descriptive study** is undertaken in order to ascertain and be able to describe the characteristics of the variables of interest in a situation.

Descriptive studies are undertaken in organizations to learn about and describe the characteristics of a group of employees. They are also undertaken to understand the characteristics of organizations that follow certain common practices.

Descriptive studies that present data in a meaningful form help to:

- Understand the characteristics of a group in a given situation
- Think systemically about aspects in a given situation
- Offer ideas for further probe and research
- Help make certain simple decisions

Qualitative data obtained by interviewing individuals may help the understanding of phenomena at the exploratory stages of a study, quantitative data in terms of frequencies, or mean and standard deviations, become necessary for descriptive studies.

#### Case study analysis

Case studies involve in-depth, contextual analyses of matters relating to similar situations in other organizations.

In exploratory studies:

- Researcher explores situational factors to understand the characteristics of the phenomena
- Pilot studies on a small scale are not uncommon

Descriptive studies:

- Undertaken when the characteristics in a situation are known to exist
- Hypothesis testing offers an enhanced understanding of the relationship that exists among variables
  - May establish cause and effect relationships
  - Can be done with qualitative and quantitative data

### Casual versus correlational relationship

**Casual study**: A study in which the researcher wants to delineate the cause of one or more problems

**Correlational study**: A study in which the researcher is interested in delineating the important variables associated with the problem

#### Extent of researcher interference with the study

The extent of interference by the researcher with the normal flow of work has a direct bearing on whether the study is casual or correlational.

A correlational study is conducted in the natural environment with minimal interference.

If trying to establish cause and effect relationships, the researcher tries to manipulate certain variables to study the effects on the dependent variable.

#### Contrived and noncontrived study setting

**Noncontrived settings**: Organizational research done in the natural environment where work proceeds normally **Contrived settings**: Artificial settings

Correlational studies are invariably conducted in noncontrived settings Most rigorous causal studies are done in contrived settings

**Field studies**: Correlational studies done in organizations **Field experiments**: Studies conducted to establish cause-and-effect relationships using the same natural environment as default **Lab experiments**: Experiments done to establish cause and effect relationship

**Lab experiments**: Experiments done to establish cause-and-effect relationship beyond the possibility of the least doubt which require the creation of an artificial, contrived environment in which all the extraneous factors are strictly controlled.

#### Unit of analysis: individuals, dyads, groups, organizations, cultures

**Unit of analysis**: The level of aggregation of the data collected during the subsequent data analysis stage

Dyads: Several two-person groups

If the problem statement is related to group effectiveness, then the unit of analysis will be at the group level.

#### **Cross-sectional studies**

**Cross-sectional studies (one-shot)**: A study undertaken in which data are gathered just once in order to answer a research question.

#### Longitudinal studies

**Longitudinal studies**: Studies where data on the dependent variable are gathered at two or more points in time to answer the research question.

- Take more time and effort & cost more than cross-sectional studies
- Help to identify cause-and-effect relationships

Most field studies conducted are cross-sectional because of time, effort, and costs involved.

#### Managerial implications

Knowledge about research design issues helps the manager to understand what the research is attempting to do.

One of the important decisions a manger has to make before starting a study pertains to how rigorous the study ought to be.

# Quantitative data analysis

After data have been collected from a representative sample of the population, the next step is to analyze them to test the research hypotheses. Preliminary tests must first be completed:

#### Getting the data ready for analysis

- **Data coding** involves assigning a number to the participants' responses so they can be entered into a database.
- Code the response in the same way.
  - At least 10% of the coded questionnaires should be checked for accuracy.
- Data entry
  - $\circ$   $\;$  Each row of the editor represents a case or observation
  - Each column represents a variable
  - Always use the first column for identification purposes; assign a number to every questionnaire, write this number on the first page of the questionnaire, and enter this number in the first column of your data file.
- Editing data
  - Deals with detecting and correcting illogical, inconsistent, or illegal data and omissions in the information returned by the participants of the study.
  - Illogical response is an outlier response.
    - An outlier is an observation that is substantially different from the other observations.
    - Outliers aren't always an error.
    - You can check the dispersion of nominal and/or ordinal variables by obtaining minimum and maximum values and frequency tables.
  - Inconsistent responses are responses that are not in harmony with other information.
  - Illegal codes are values that are not specified in the coding instructions.
  - One way to handle a blank response is to ignore it when the analyses are done. A disadvantage of this is it will reduce the sample size.
  - Another way to handle a blank response it to look at respondent's pattern of responses to other questions and deduce a logical answer.
- **Data transformation** is the process of changing the original numerical representation of a quantitative value to another value.
  - Data are typically changed to avoid problems in the next stage of the data analysis process.
  - Reverse scoring is another type of data transformation.
  - Necessary when several questions have been used to measure a single concept.
  - If questions are scattered, care has to be taken to include all the items without any omission or wrong inclusion.

Getting a feel for the data is the necessary first step in all data analysis.

- Frequencies refer to the number of times various subcategories of a certain phenomenon occur, from which the % and the cumulative % of their occurrence can be easily calculated.
  - Frequencies can also be visually displayed as bar charts, histograms, or pie charts.

#### Measures of central tendency and dispersion

There are 3 measures of central tendency:

- The **mean**, average, is a measure of central tendency that offers a general picture of the data without inundating one with each of the observations in a data set.
- The **median** is the central item in a group of observations when they are arrayed in either an ascending or descending order.
- The **mode** is the most frequently occurring group.

#### Measures of dispersion

- The measure of dispersion is also unique to nominal and interval data. The 3 measurements of dispersion connected with the mean are:
  - **Range** refers to the extreme values in a set of observations.
  - **Variance** is calculated by subtracting the mean from each of the observations in the data set, taking the square of the difference, and dividing the total of these by the number of observations.
  - **Standard deviation** offers an index of the spread of a distribution or the variability in the data.
    - Very commonly used
    - The square root of the variance
    - The mean and standard deviation are the most common descriptive statistics for interval and ratio scaled data. The standard deviation, in conjunction with the mean, is a very useful tool because:
      - Practically all observations fall within 3 standard deviations of the average or the mean.
      - More than 90% of the observations are within 2 standard deviations of the mean.
      - More than half of the observations are within one standard deviation of the mean.
- When the median is the measure of central tendency, percentiles, deciles, and quartiles become meaningful. The quartile divides into 4 equal parts, decile into 10, percentile into 100.
- **Interquartile range** is the measure of dispersion for the median and consists of the middle 50% of the observations.
  - Very useful when comparisons are to be made among several groups.

#### Relationship between variables

Nonparametric tests are available to assess the relationship between variables measured on a nominal or an ordinal scale. Spearman's and Kendall's rank

correlation are used to examine the relationship between 2 ordinal variables. A correlation matrix is used to examine the relationship between interval and/or ratio variables.

## Relationship between 2 nominal variables, $x^2$ test

- **Chi-square**  $x^2$  **test**: a nonparametric test which indicated whether or not the observed pattern is due to chance.

  - $x^2 = \sum \frac{(Oi Ei)^2}{Ei}$   $x^2$  = chi-square statistic
    - Oi is the observed frequency of the *i*th cell
    - Ei is the expected frequency
  - The chi-square statistic is associated with the degrees of freedom, df, which denote whether or not a significant relationship exists between two nominal variables.
  - Helps us see whether or not 2 nominal variables are related.
- Correlations
  - A Pearson correlation matrix will indicate the direction, strength, and significance of the bivariate relationship among all the variables that were measured at an interval or ratio level.
  - A correlation coefficient that indicates the strength and direction of the relationship can be computed by applying a formula that takes into consideration the 2 sets of figures.
    - Theoretically, there should be a perfect positive correlation between 2 variables, which is represented by 1.0 and perfect negative, represented by -1.0
  - The Pearson correlation is appropriate for interval and ratio-scaled variables, and the Spearman Rank or Kendall's Tau coefficients are appropriate when variables are measured on an ordinal scale.

# Testing goodness of data

- Reliability of a measure is established by testing for both consistency and stability. Consistency indicates how well the items measuring a concept hang together as a set.
  - Test-retest reliability can be established by computing the correlation between the same tests administered at 2 different time periods.
  - Reliabilities < .60 are considered poor, those in the .70 range are acceptable, and those > .80 are good.
  - All the negatively worded items in the questionnaire should first be reversed before the items are submitted for reliability tests.
  - Taking out an item, though improving the reliability of our measure, 0 affects the validity of our measure in a negative way.
- Validity
  - Factorial validity can be established by submitting the data for factor analysis.
    - The results will confirm whether or not the theorized dimensions emerge.
    - Measures are developed by first delineating the dimensions so as to operationalize the concept.

- **Criterion-related validity** can be established by testing for the power of the measure to differentiate individuals who are known to be different.
- **Convergent validity** can be established when there is a high degree of correlation between two different sources responding to the same measure.
- **Discriminant validity** can be established when 2 distinctly different concepts are not correlated to each other.
  - **Convergent** and **discriminant** validity can be established through the multi-trait multi-method matrix.
  - When well-validated measures are used there is no need to establish their validity again for each study.

Descriptive statistics such as maximum, minimum, mean, standard deviations, and variance can now be obtained from the multi-item, interval scaled independent and dependent variables.

After we have obtained descriptive statistics for the independent and dependent variables in our study, we can test our hypotheses.

# Quantitative data analysis: hypothesis testing

Steps to be followed in hypothesis development and testing:

- State the null and the alternate hypotheses
- Determine the level of significance desired
- Choose the appropriate statistical test depending on the type of scales that have been used (nominal, ordinal, interval, or ratio)
- See if the output results from computer analysis indicate that the significance level is met. When the result is larger than the critical value, the hypothesis is rejected

#### Type I and II errors, and statistical power

The hypothetico-deductive method requires hypotheses to be falsifiable.

- Null hypotheses  $H_0$  are thus set up to be rejected in order to support the alternate hypothesis, termed  $H_a$
- The null hypothesis is presumed true until statistical evidence indicates otherwise
- The purpose of hypothesis testing is to determine accurately if the null hypothesis can be rejected in favor of the alternate hypothesis
- There are 2 kinds of errors:
  - A **type I error**, aka alpha  $\alpha$ , is the probability of rejecting the null hypothesis, when it's actually true
    - The probability of type I error, aka significance level, is determined by the researcher
    - Typical significance levels in business research are 5% and 1%
  - A **type II error**, aka beta  $\beta$ , is the probability of failing to reject the null hypothesis given that the alternate hypothesis is actually true
    - The probability of type II error is inversely related to the probability of type I error: the smaller the risk of one of these types of error, the higher the risk of other type of error
  - **Statistical power** is the probability of correctly rejecting the null hypothesis. Statistical power depends on:
    - Alpha
    - Effect size
    - The size of the sample

#### Choosing the appropriate statistical technique

The choice of the appropriate statistical technique largely depends on the number of variables you are examining and the scale of measurement of your variables.

*Univariate statistical techniques* are used when you want to examine two-variable relationships.

#### Testing a hypothesis about a single mean

The **one sample t** –**test** is used to test the hypothesis that the mean of the population from which a sample is drawn is equal to a comparison standard.

The following formula is used to compute the t-value:

$$t_{n-1} = \frac{X - \mu}{s / \sqrt{n}}$$

Having calculated the t-statistic, we can now compare the t-value with a standard table of t-values with n-1 degrees of freedom to determine whether the t-statistic reaches the threshold of statistical significance. When the t-statistic is larger than the appropriate table value, the null hypothesis is rejected.

## Testing hypotheses about two related means

We can also do a (paired samples) t-test to examine the differences in the same group before and after a treatment.

The **Wilcoxon signed-rank test** is a nonparametric test for examining significant differences between two related samples or repeated measurements on a single sample.

**McNemar's test** is a nonparametric method used on nominal data. It assesses the significance of the difference between two dependent samples when the variable of interest is dichotomous. It is used primarily in before-after studies to test for an experimental effect.

- A rather straightforward technique to test marginal homogeneity
- *Marginal homogeneity* refers to equality (or the lack of a significant difference) between one or more of the marginal row totals and the corresponding marginal column totals
- Marginal homogeneity would mean there was no effect of the treatment

# Testing hypotheses about two unrelated means

An **independent samples t-test** is carried out to see if there are any significant differences in the means for two groups in the variable of interest

• A **nominal** variable that is split into two subgroups is tested to see if there is a significant mean difference between the two split groups on a dependable variable, which is measured on an **interval** or **ratio** scale

### Testing hypotheses about several means

An **analysis of variance (ANOVA)** helps to examine the significant mean differences among more than two groups on an interval or ratio-scaled dependent variable.

- The results of ANOVA show whether or not the means of the various groups are significantly different from one another, as indicated by the F statistic
- The F statistic shows whether two sample variances differ from each other or are from the same population
- The F distribution is a probability distribution of sample variances and the family of distributions changes with changes in the sample size
- It's unwise to use multiple t-tests, taking two groups at a time, because the greater the number of t-tests done, the lower the confidence we can place on results

## **Regression analysis**

Simple **regression analysis** is used in a situation where on independent variable is hypothesized to affect one dependent variable.

- We can model this linear relationship by a *least squares* function
- A simple linear regression equation represents a straight line
- We can also express this relationship in an equation:
  - $\circ \quad Y_i = \beta_0 + \beta_1 X_\mu + \mu$
  - The parameters  $B_0$  and  $B_1$  are called regression coefficients. There are the intercept  $B_0$  and the slope  $B_1$  of the straight line relating propensity to buy (Y) to perceived quality (X<sub>1</sub>)
  - The coefficient of determination, R<sup>2</sup>, provides information about the goodness of fit of that regression model: it is statistical measure of how well the regression line approximates the real data points
  - R<sup>2</sup> is the percentage of variance in the dependent variable that is explained by the variation in the independent variable
  - If R<sup>2</sup> is near to 1, most of the variation in the dependent variable can be explained by the regression model
- The basic idea of **multiple regression analysis** is similar to that of simple regression analysis
  - In this case, we use more than one independent variable to explain variance in the dependent variable
  - Multiple regression analysis provides a means of objectively assessing the degree and the character of the relationship between the independent variables and the dependent variable:
    - The regression coefficients indicate the relative importance of each of the independent variables in the prediction of the dependent variable
- The square of multiple r, R-square, or R<sup>2</sup> as it is commonly known, is the amount of variance explained in the dependent variable by the predictors

### Standardized regression coefficients

**Standardized regression coefficients** or **beta coefficients** are the estimates resulting from a multiple regression analysis performed on variables that have been standardized.

### Regression with dummy variables

A **dummy variable** is a variable that has two or more distinct levels, which are coded 0 or 1. Dummy variables allow us to use nominal or ordinal variables as independent variables to explain, understand, or predict the dependent variable.

# Multicollinearity

**Multicollinearity** is an often encountered statistical phenomenon in which two or more independent variables in a multiple regression model are highly correlated.

- In its most severe case multicollinearity makes the estimation of the regression coefficients impossible
- In all other cases it makes the estimates of the regression coefficients unreliable

- The simplest and most obvious way to detect multicollinearity is to check the correlation matrix for the independent variables
- The presence of high correlations is a first sign of sizeable multicollinearity
- More common measures for identifying multicollinearity are therefore the tolerance value and the variance inflation factor (VIF)
- These measures indicate the degree to which one independent variable is explained by the other independent variables
- Multicollinearity is not a serious problem if the purpose of the study is to predict or forecast future values of the dependent variable
- If the objective of the study is to reliably estimate the individual regression coefficients, multicollinearity is a problem
- Following methods to reduce it:
  - Reduce the set of independent variable to a set that are not collinear
  - $\circ\;$  Use more sophisticated ways to analyze the data, such as ridge regression
  - Create a new variable that is a composite of the highly correlated variables

# Other multivariate tests and analyses

- **Discriminant analysis** helps to identify the independent variables that discriminate a nominally scaled dependent variable of interest
  - The independent variables measured on an interval or ratio scale discriminate the groups of interest to the study
- Logistic regression is also used when the dependent variable is nonmetric
  - When the dependent variable has only two groups, logistic regression is often preferred, because it doesn't face the strict assumptions that discriminant analysis faces and because it's very similar to regression analysis
  - Very different from a statistical point of view, alike from a practical viewpoint
  - Both methods produce prediction equations and in both cases the regression coefficients measure the predictive capability of the independent variables
  - Logistic regression allows the researcher to predict a discrete outcome from a set of variables that may be continuous, discrete, or dichotomous
- **Conjoin analysis** is a statistical technique that is used in many fields including marketing, product management, and operations research.
  - Requires participants to make a series of trade-offs
  - In marketing, conjoint analysis is used to understand how consumers develop preferences for products or services
  - By asking for enough choices, it's possible to establish how important each of the levels is relative to the others; known as the utility of the level
  - Conjoint analysis is traditionally carried out with some form of multiple regression analysis
- **Two-way ANOVA** can be used to examine the effect of two nonmetric independent variables on a single metric dependent variable

- In this context, an independent variable is often referred to as a factor and this is why a design that aims to examine the effect of two nonmetric independent variables on a single metric dependent variable is often called a *factorial design*
  - Very popular in the social sciences
- Two-way ANOVA enables us to examine main effects but also interaction effects that exist between the independent variables
- **MANOVA** is similar to ANOVA, with the difference that ANOVA tests the mean difference of more than 2 groups on *one* dependent variable, whereas MANOVA tests mean differences among groups across *several* dependent variables simultaneously, by using sums of squares and cross-product matrices
  - Multiple ANOVA tests, using one dependent variable at a time, would also bias the results, since the dependent variables are likely to be interrelated
  - MANOVA circumvents this bias by simultaneously testing all the dependent variables, cancelling out the effects of any inter-correlations among them
  - The independent variable is measured on a nominal scale and the dependent variables on an interval or ratio scale
- **Canonical correlation** examines the relationship between two or more dependent variables and several independent variables

## Data warehousing, data mining, and operations research

A **data warehouse** serves as the central repository of all data collected from disparate sources including those pertaining to the company's finance, manufacturing, sales, etc.

- Usually built from data collected through the different departments of the enterprise and can be accessed through various online analytical processing (OLAP) tools
- Can be described as the process of extracting, transferring, and integrating data spread across multiple external databases and ever operating systems, with a view to facilitating analysis and decision making

Using algorithms to analyze data in a meaningful way, **data mining** more effectively leverages the data warehouse by identifying hidden relations and patterns in the data stored in it.

**Operation research** (OR) or *management science* (MS) is another sophisticated tool used to simplify and thus clarify certain types of complex problem that lend themselves to quantification.

- Uses higher mathematics and statistics to identify, analyze, and ultimately solve intricate problems of great complexity faced by the manager
- Provides an additional tool to the manager by using quantification to supplement personal judgment
- Areas of problem solving that easily lend themselves to OR include those relating to inventory, queuing, sequencing, routing, and search and replacement

• OR helps to minimize costs and increase efficiency by resorting to decision trees, linear programming, network analysis, and mathematical models

It should be ensured that the data in the information system are error-free and are frequently updated

## Some software packages useful for data analysis

Based on your specific needs, your research problem, and/or conceptual model you might consider the following software packages:

- LISREL
  - Designed to estimate and test structural equation models
  - Carry out exploratory factor analysis and confirmatory factor analysis
- MATLAB
  - A computer program that was originally designed to simplify the implementation of numerical linear algebra routines
  - Used to implement numerical algorithms for a wide range of applications
- SAS/STAT
  - An integrated system of software products, capable of performing a broad range of statistical analyses such as descriptive statistics, multivariate techniques, and time series analyses
  - Used in many disciplines, including medical sciences, biological sciences, social sciences, and education
- SPSS
  - A data management and analyses program designed to do statistical data analysis, including descriptive statistics as well as sophisticated inferential and multivariate statistical procedures like analysis of variance (ANOVA), factor analysis, cluster analysis, and categorical data analysis
- Stata
  - A general purpose statistical software package that supports various statistical and econometric methods, graphics, and enhanced features for data manipulation, programming, and matrix manipulation

# The research report

#### The report

It's important that the results of the study and the recommendations for solving the problem are effectively communicated to the sponsor, so that suggestions made are accepted and implemented.

- Writing the report concisely, convincingly, and with clarity is perhaps as important than conducting a perfect research study
- The contents and organization of both modes of communication depend on the purpose of the research study, and the audience to which it's targeted

#### The written report

The written report enables the manager to weigh the facts and arguments presented herein, and implement the acceptable recommendations

#### The purpose of the written account

- It's important to identify the purpose of the report
  - If the purpose is simply to offer details on some specific areas of interest requested by a manager, the report can be very narrowly focused and provide the desired information to the manager in a brief format
  - If the report is intended to sell an idea to management, then it has to be more detailed and convincing as to how the proposed idea is an improvement and why it should be adopted
  - A different form of report will be prescribed in some cases, where a manager asks for several alternative solutions or recommendations to rectify a problem in a given situation
    - Here the researcher provides the requested information and the manager chooses from among the alternatives
    - In this case, a more detailed report surveying past studies, the methodology used for the present study, different perspectives generated from interviews and current data analyses, and alternative solutions based on the conclusions drawn therefrom will have to be provided
    - The advantage and disadvantage of each of the proposed solutions, together with a cost-benefit analysis in terms of dollar and/or other resources, will also have to be presented to help the manager make the decision
  - Yet another type of report might require the researcher to identify the problem and provide the final solution as well.
    - Such a report has to be very comprehensive, following the format of a full-fledged study
  - A fifth kind of research report is the very scholarly publication presenting the findings of a basic study that one usually finds published in academic journals

### The audience for the written report

- The organization of a report, its lengths, focus on details, data presentation, and illustrations will be a function of the audience for whom it's intended
- The letter of transmittal of the report will clearly indicate to whom the report is being sent
- An executive summary placed at the beginning will offer busy executives just the right amount of vital details, in less than 3 pages
- The letter, organization, and presentation model of the report will, among other things, depend at least in part of the target audience
- Sometimes, the findings of a study may be unpalatable to the executive, or may reflect poorly on management.
  - While there is no need to suppress the unpalatable findings, they can be presented in a nonjudgmental, non-fault-finding or finger-pointing manner, using objective data and facts that forcefully lead to, and convince the managers of the correctness of, the conclusions drawn
  - If this isn't done, the report will read defensively, the recommendations will not be accepted, and the problem will remain unsolved
  - Tact and diplomacy combined with honesty and objectivity are essential in report writing and presentation

#### Characteristics of a well-written report

- Certain basic features are integral to all written reports:
  - Clarity, conciseness, coherence, the right emphasis on important aspects, meaningful organization of paragraphs, smooth transition from one topic to the next, apt choice of words, and specificity
- The report should, to the extent possible, be free of technical or statistical jargon unless it happens to be of a technical or statistical nature
- Care should be taken to eliminate grammatical and spelling errors
- Any assumptions made by the researcher should be clearly stated in the report, and facts, rather than opinions, provided
- The importance of the appearance of the report and its readability cannot be overemphasized

### Contents of the research report

- Should bear a title that indicates what the study is about
- Should have at the beginning a table of contents the research proposal, a copy of the authorization to conduct the study, and an executive summary or a synopsis
- All reports should have an introductory section detailing the purpose of the study, giving some background of what it relates to and stating the problem studied, setting the stage for what the reader should expect in the rest of the report
- The body of the report should contain details regarding the framework of the study, hypotheses, sampling design, data collection methods, analysis of data, and the results obtained
- The final part of the report should present the findings and draw conclusions
  - If recommendations have been called for, they will be included, with a cost-benefit analysis provided with respect to each

- $\circ$   $\;$  Such information clarifies the net advantages of implementing each of the recommendations
- The details provided in the report should be such as to convince the reader of the thoroughness of the study, and induce confidence in accepting the results and the recommendations made
- Every professional report should also point out the limitations of the study
- Good descriptions and lucid explanations, smooth and easy flow of materials, recommendations that flow logically from the results of data analysis, and an explicit statement of any limitations to the study, provide scientific authenticity to the report
- The conclusion drawn from the findings should indicate a clear solution to the problem
- The report can be organized in parts, sections, or chapters and should be tailored to meet the needs of the situation
- Appendices, if any, should be attached to the report

### Integral parts of the report

- The title page
  - The title of the report should succinctly indicate what the study is all about
  - The title page will indicate the name of the sponsor of the study, the names of the researchers and their affiliations, and the date of the final report

#### Table of contents

- Usually lists the important headings and subheadings in the report with page references
- A separate list of tables and figures should also be listed in the table of contents
- The research proposal and the authorization letter
  - A copy of the letter of authorization from the sponsor of the study approving the investigation and detailing its scope will be attached at the beginning of the report along with the research proposal

### • The executive summary or synopsis

- A brief account of the research study that provides an overview, and highlights the following important information related to it:
  - The problem statement
  - Sampling design
  - Data collection methods used
  - Results of data analysis
  - Findings
  - Recommendations, with suggestion for their implementation
- Will be brief, usually 3 pages or less in length

### • The introductory section

- Starts with a statement of the problem under investigation
- The research objective, together with background information on why and how the study was initiated, is also stated
- The body of the report

- The details of the interviews conducted, the literature survey, the theoretical framework, and the hypotheses are furnished
- The design details, such as sampling and data collection methods, as well as the nature and type of study, the time horizon, the field setting, and the unit of analysis, are described
- The details of the types of data analysis done to test the hypotheses, and the findings therefrom, will be provided next
- Tabular and pictorial depictions of the results of data analysis will find a place here

### • The final part of the report

- Will contain the conclusions drawn from the findings
- In most cases a list of recommendations for implementation will follow
- Frequently, a cost-benefit analysis will also be provided
- Any limitations to the study will find place here
- A brief summation paragraph will also be provided at the end

#### Acknowledgments

• Help received from others is next acknowledged

### • References

- Starting on a fresh page, a list of the references cited in the literature review and in other places in the report will be given
- Appendix
  - The appropriate place for the organization chart, newspaper clippings or other materials that substantiate the text of the report, detailed verbatim narrations of interviews with members, and whatever else might help the reader follow the text
  - It should also contain a copy of the questionnaire administered to the respondents

### Oral presentation

- Usually organizations require about a 20-minute oral presentation of the research project, followed by a question and answer session
- Different stimuli have to be creatively provided to the audience to consistently sustain their interest throughout the presentation
- Sliders, overheads, charts, graphs, handouts, all in large, bold print, and preferably in multiple colors, help the presenter to sustain the interest of the audience
- Factors irrelevant to the written report, such as dress, mannerisms, gestures, voice modulation, and the like, take on added importance in oral presentations

### Deciding on the content

- It's important to determine what the presenter would like the listener to walk away with, and then organize the presentation accordingly
- Depending on the type of audience, it may become necessary to put more stress on the data analytic aspects
- The time and attention devoted to the various components of the study will require adjustment, depending on the audience

#### Visual aids

- Provide a captivating sensory stimulus that sustains the attention of the audience
- The selection of specific visual modes of presentation will depend, among other things, on the size of the room, the availability of a good screen for projection, and the cost constraints of developing sophisticated visuals

#### The presenter

- An effective presentation is also a function of how unstressed the presenter is
- The speaker should establish eye contact with the audience, speak audibly and understandably, and be sensitive to the nonverbal reactions of the audience
- Strict adherence to the time frame and concentration on the points of interest to the audience are critical aspects of presentation
- One should also not minimize the importance of the impression created on the audience by dress, posture, bearing, and the confidence with which one carries oneself

#### The presentation

• The opening remarks set the stage for riveting the attention of the audience

#### Handling questions

- Concentrated and continuous research on the research topic over a considerable period of time indisputably makes the presenter more knowledgeable about the project than anyone else in the audience
- It's important to be non-defensive when questions are posed that seemingly find fault with some aspect of the research
- The question and answer session, when handled well, leaves the audience with a sense of involvement and satisfaction
- Reporting has to be done in an honest and straightforward manner
  - It's unethical to fail to report findings that are unpalatable to the sponsors or that reflect poorly on management

#### Surveys

- Surveys are information collection methods used to describe, compare, or explain individual and societal knowledge, feelings, values, preferences, and behavior
- A survey can be a self-administered questionnaire that someone fills out alone or with assistance, or a survey can be an interview done in person or on the telephone
- Some surveys are on paper or online, and the respondent can complete them privately at home or in a central location
- Surveys are by no means the only source of information for making decisions, nor are they necessarily the most relevant. Some other sources of information:
  - Observations or eyewitness reports
  - Performance tests that require a person to perform a task
  - Written tests of ability or knowledge
  - Reviews that rely on existing documentation
- Sometimes surveys are combined with other sources of information
  - This is particularly true for evaluations and research
- All surveys consist of:
  - Questions and responses
  - Sampling and design
  - Data processing or management and analysis
  - o Pilot testing
  - Response rate
- Information from surveys is obtained by asking questions
  - Questions are sometimes referred to as *items*
  - The questions may have forced-response choices
- Surveys are data collection methods used to obtain information from and about people
  - The sample is the number and characteristic of people in the survey
  - The design refers to how often the survey takes place, whether the participants are selection at random or not, and how many separate groups are included
- Regardless of a survey's design or size, a plan must be made on how to analyze the survey's data
- All surveys must be pilot tested
  - A **pilot test** is a tryout, and its purpose is to help produce a survey form that is usable and that will provide you with the information needed
  - $\circ$   $\,$  Self-administered questionnaires depend heavily on the clarity of their language
  - $\circ$   $\,$  Whenever possible the environment the survey will be taken in should be duplicated
- The get decent results:
  - Decide on the type of survey
  - Select the survey's content and writing questions and try out the form

- Decide who should participate
- Administer the survey
- Process the data
- Analyze and interpret the results
- $\circ$   $\;$  Report the results orally or in writing
- A reliable surveys results in consistent information
- A valid survey produces accurate information
- Reliable and valid surveys are obtained by making sure the definitions and models you use to select questions are grounded in theory or experience
- Focus groups and pilot tests help you decide which type to use and whether you have done a good job of designing the survey
- A well designed, easy to use survey always contributes the reliability and validity
- The results will be useful if they are valid and if the survey method is one that users accept as correct
- Costs refer to the financial burden of developing and administering each type of survey
- The costs associated with written questionnaires include paper, reproduction, and incentives
- The costs of face-to-face and telephone interviews include phone charges or the costs of purchasing a telephone system
- Online surveys require extensive development and testing
- Online surveys are self-administered questionnaires
- At least four potential concerns accompany most web surveys:
  - Multiple addresses
  - Privacy and confidentiality
  - Identity of respondents
  - Technological differences
- Cell phone surveys can be more expensive
  - They are also more complicated by the fact that they are attached to individuals rather than geography
  - Almost always require incentives for participation because the respondent otherwise pays for the call directly or indirectly
- Some people have become suspicious of surveys
  - They fear the information they provide will be used inappropriately
  - Many techniques exist for protecting each person's privacy:
    - The surveyor needs to reassure potential respondents that these techniques have been incorporated into each survey
    - All completed surveys should be kept in locked files, and only a limited number of staff should have access to them
    - The surveys should be shredded when done using
    - Separate identifying information from survey responses by assigning codes to individuals and using the codes to link them to their responses
    - The consent form gives potential respondents sufficient written information to decide whether to complete a survey
    - If passwords need to be used, ensure the password setup is secure

- Decide in advance whether to use SSL and how to explain security choices to respondents
- Sensitive identifiable data need to be protected in transit by using either an SSL protocol or a SFT protocol
- All databases storing sensitive and identifiable information must be protected
- The informed consent model with children requires parents or legal guardians to provide permission for:
  - Their children to take the survey
  - The surveyor to contact the children
- When going through the assent process, make sure children know:
  - They are being asked to participate in a research study or in a local effort to understand more about a problem
  - The purpose of the survey
  - How much time is involved in participating
  - What will happen to them if they agree to participate
  - The foreseeable risks and/or discomforts and any benefits they may experience
  - That they can ask their parents or surveyor any questions they have about participating
  - That their participation is voluntary and that they may stop at any time
  - That their father/mother/guardian knows they have been asked to be part of the study
- When doing international surveys make sure the questionnaire reflects not only the language of the originating country but also the standards, expectation, values, and preoccupations
- Ethical research in international settings means that culturally appropriate procedures must be followed to protect survey participants:
  - Be sure the survey is available in the original language and in translation
  - If the survey includes children, very the age at which participants have the ability to consent or assent to participation in surveys
  - If local customs and regulations are such that active parental permission would be culturally inappropriate, the surveyor should compile proof that such permission is not culturally appropriate
  - If the surveyor is eligible for a waiver of active parental permission, they should be prepared to provide parents with a letter describing the study
  - The surveyor also needs to get letters of agreement from the appropriate officials that state that the survey itself has been reviewed and is acceptable

# Find 2

## The survey form

- To select the content of a survey, you have to define your terms and clarify what you need and can get from asking people about their views
  - o Define the terms
  - o Select your information needs or hypotheses
  - Make sure you can get the information
  - Do not ask for information unless you can act on it
- Multiple-choice questions are efficient and often more reliable than other questions
  - $\circ$   $\;$  This efficiency comes from being easy to use and score  $\;$
  - Their reliability is enhanced because of the uniform data they provide
- Open-ended questions can offer insight into why people believe the things they do
  - Interpreting them can be difficult, unless accompanied by an elaborate classification system
- Choose open-ended questions when you want to give the respondents the opportunity to express opinions in their own words and you have the interest and resources to interpret the findings
- Choose closed questions for their relative ease of scoring, analysis, and interpretation
  - Can be difficult to prepare because they require that all respondents interpret them the same way and that all relevant choices are included, mutually exclusive, and sensible
- One open-ended set of questions that is particularly appropriate for getting at satisfaction requires collection info about what people like best (LB) and what they like least (LB). The LB/LL technique:
  - Ask respondents to list what is good and what is bad
  - Once you have all the responses, the next step is to categorize and code them
  - $\circ$   $\,$  When you are satisfied about reliability, the next step is to count the number of responses for each code
    - First you must decide on your response categories and assign each one a code, then:
      - Put the codes in rank order
      - Calculate the percentage of students assigned to each code
      - Count the number of responses assigned to each code
      - Calculate the percentage of responses assigned to each code
      - Calculate the cumulative percentage of responses by adding the percentages together
    - For each response category, you should look for 2 points on the x axis: the percentage of participants and the cumulative percentage of responses

- The cumulative percentages of responses are joined with a square
- Some of the points representing percentages of participants are joined with a diamond
- Multiple-choice, closed survey questions consist of a stem, which presents a problem, followed by several alternative choices or solutions
  - Each question should be meaningful to respondents
  - Use standard language rules
  - Make questions concrete
  - Avoid biased words and phrases
  - Check your own biases
  - Use caution when asking for personal information
  - o Each question should have just one thought
- Yes-and-no responses are simple to use and score
- A rating scale requires the respondent to place the item being rated in any one of an ordered series of categories or at some point along a continuum
  - A numerical value is assigned to the ordered categories or points
  - There are 3 types of rating or measurement scales:
    - Categorical. These are sometimes called nominal response scales. They require people to affirm or name the groups to which they belong
    - Ordinal. These scales require that respondents place answers in order of importance
    - Numerical scales. Numerical scales take two forms:
      - Discrete: produces a precise number
      - Continuous: produces a number that falls on a continuum
- The distinctions among categorical, ordinal, and continuous scales are important because they determine the kinds of statistical treatments you can use
- When raters use ordinal scales, they select one of a limited number of categories that have some order to them
  - Ordinal scales are easy to use and interpret
  - Responses to Likert items can be analyzed as ordinal data
    - When regarded as ordinal data they can be collated into bar charts and summarized by the median or more, not mean, and variation summarized by the range or analyzed using nonparametric tests
  - If the surveyor regards the scale as continuous, then means and standard deviations are appropriate statistics
  - Treating the data as continuous can occur if the surveyor decides to compare average ratings
  - Data from Likert scales are sometimes reduced to the categorical level by combining all agree and disagree responses into 2 categories of accept and reject
  - Forced-choice questions are often useful when you want to divert the respondent from taking the path of least resistance by choosing a middle category. Forcing respondents to choose may annoy them and may not uncover the truth.
  - o Some ordinal scales rely on relative judgments or ranks

- o Ranking involves comparing one factor to another
- Comparative scale is another form ordinal scales can take: this one asks the respondent to contract a single specific object in terms of a general situation
- A checklist provides respondents with a series of answers. They may choose just one or more answers depending on the instructions
  - Checklists help remind respondents of some things they might have forgotten
  - o Somewhat difficult to format and interpret responses to checklists
  - Tips for creating written and online surveys for children and adolescents:
    - Keep questions short

٠

- With very young children be specific and avoid abstractions
- Try out all questions before you use them
- Ask for responses to positive statements
- Make certain that children understand that there are no wrong or right answers
- $\circ$   $\,$  Do not force children to answer a question if they don't want to
- Reinforce your respect for the respondent's privacy
- Online surveys are almost always self-administered questionnaires
  - Rules are exactly the same as for other self-administered questionnaires
    - The survey must be designed to get the information you need from the people who are most likely to give the most accurate answers
    - The questions must be grammatically correct with a reasonable range of options for responses
  - Online surveys often contain a progress bar
  - Should explain how to go back to change answers and then move forward
- Survey questions and response choices should be written at a 6<sup>th</sup> to 8<sup>th</sup> grade reading level
- The methods used to produce an additive scale require sophisticated survey construction skills because you have to prove conclusively that high scorers are in actuality different from low scores with respect to each and every item
- 2<sup>nd</sup> use of the term scale refers to one question or a collection of questions whose scores are meaningful
- Differential scales distinguish among people in terms of whether they agree or disagree with theory or research
- A summated scale aligns people according to how their responses add up

## Practical concerns

- The lengths of a survey depend on what you need to know and how many questions are necessary so that the resulting answers are credible. Another consideration is the respondents
- All surveys should be preceded by an introduction and the first set of questions should be related to the topic described in it
- A questionnaire's appearance is important
  - o Confused or irritated respondents can result in a loss of data
  - A poorly designed interview will reduce the efficiency of interviewers
  - $\circ~$  Do: put just one question on a line. Leave plenty of space for responses
  - Don't: squeeze several questions together. Don't abbreviate questions
- Be careful with branching question
  - Skip patterns can be confusing to people
  - Online surveys are effective for branching
- Self-administered questionnaires require a great deal of preparation and monitoring to get a reasonable response rate
- Interviewers should fit in as well as possible with respondents
- The key to a good telephone or face-to-face interview in training
- Tips on conducting interviews:
  - Make a brief introductory statement that will describe who is conducting the interview, tell why the interview is being conducted, explain why the respondent is being called, and indicate whether answers will be kept confidential
  - Try to impress on the person being interviewed the importance of the interview and of the answers
  - Prepare yourself to be flexible
  - Interview people alone
  - If using a printed interview survey, be sure to ask questions as they appear in the interview schedule
  - Interviewers should follow all instructions given at the training session and described on the interview form
  - $\circ$   $\,$  Make certain interviewers understand the informed consent process, and that they adhere to it
- Monitor the quality of interviews to get the most accurate data possible
- To prevent problems:
  - o Establish a hotline
  - o Provide written scripts for the interviewer
  - o Make sure you give out extra copies of all supplementary materials
  - Prepare an easy-to-read handout describing the survey
  - Provide a schedule and calendar so that interviewers can keep track of their progress
  - Consider providing interviewers with visual aids
  - Consider the possibility that some interviewers may need to be retrained and make plans to do so

- In reviewing a published survey, you should ask the following questions about 3 types of reliability, test-retest, equivalence, and internal consistency
  - Does the survey have test-retest reliability?
  - Are alternative forms equivalent?
  - How internally consistent are the survey questions in measuring the characteristics, attitudes, or qualities that they are supposed to measure?
- A survey is valid if the information it provides is an accurate reflection of respondents' knowledge, attitudes, values, and behavior.
  - Does the survey have predictive validity? You can validate a survey by proving that it predicts an individual's ability to perform a given task or behave in a certain way
  - Does the survey have concurrent validity? You can validate a survey by comparing it against a known and accepted measure
  - Does the survey have content validity? A survey can be validated by proving that its items or questions accurately represent the characteristics or attitudes they are intended to measure
  - Does the survey have construct validity? Surveys can be validated by demonstrating that they measure a construct such as hostility or satisfaction
- One way to make sure that you're using a reliable and valid survey is to rely on one that has been carefully tested by other surveyors and researchers in a scientific study
  - Select one or more online digital libraries that are relevant to your survey
  - $\circ$   $\,$  Specify the topics or variables you want your survey to cover
  - o Learn how to search the appropriate electronic libraries
  - Make certain that you can administer the survey the way it was intended or that you have the resources to adapt it
  - Check on the characteristics of the survey respondents
  - Select the survey that has the best available evidence of reliability and validity
  - Try out the survey with a small group of people to see if it's useable and useful
- Here are some basic rules for pilot testing a survey:
  - Try to anticipate the actual circumstances in which the survey will be conducted and make plans to handle them
  - You can start by trying out selected portions of the survey in an informal fashion
  - $\circ$  Choose respondents similar to the ones who will eventually complete the survey
  - Enlist as many people in the pilot as seems reasonable without wasting your resources
  - For reliability, focus on the clarity of the questions and the general format of the survey
  - To help bolster validity, you should ensure that all relevant topics have been included in the survey
  - Test your ability to get a range of responses

## Sampling

•

If you decide to sample, you must ask: How many people should be included?

- Some other questions you should answer:
  - How guickly are data needed?
  - What type of survey is planned?
  - What are your resources?
  - How credible will your findings be?
  - How familiar are you with sampling methods?
- 3 of the most commonly used random sampling methods are:
  - Simple random sampling
  - Stratified random sampling
  - Simple random cluster sampling
  - Stratified random cluster sampling
- A simple random sample is one in which each person has an equal chance of being selected from a population
  - The advantages of simple random sampling are: 0
    - It is the simplest of all random sampling methods
      - Aids are available to assist you
  - A major disadvantage of simple random sampling is that it cannot be used to divide respondents into subgroups or strata
- In stratified random sampling, you first subdivide the population into ٠ subgroups or strata and select a given number or proportion of respondents from each stratum to get a sample
  - Increasing sample size may be easier than implementing a stratified random sample
- Simple random cluster sampling is used primarily for administrative convenience, not to improve sampling precision
  - Cluster sampling is analogous to random sampling, except that groups rather than individuals are assigned randomly
  - This method presupposes that the population is organized into natural or predefined clusters or groups
  - The advantages of simple random cluster sampling are:
    - It can be used when selecting individuals randomly is inconvenient or unethical
    - It simplifies survey administration
  - The disadvantages of simple random cluster sampling is that it 0 requires complex statistical methods to reconcile sampling units and analytic units
- Complex sampling strategies require an understanding of sampling statistics
- In systematic sampling, you pick a number, say 5, and select every 5<sup>th</sup> name on a list that represents the population
  - There is a danger in systematic sampling:
    - Lists of people are sometimes arranged so that certain patterns can be uncovered; if you use on of these lists, your sample will be subject to a bias imposed by the pattern

- A convenience sample is one that you get because people who are willing to complete the survey are also available when you need them
- Because of bias, convenience samples are unconvincing unless you prove otherwise
- These are standard convenience sampling techniques:
  - Snowball samples
    - Previously identified members of a group identify other members
  - Quota samples
    - The group is divided into subgroups in specific proportions. This is similar to stratified sampling
  - Purposive samples
    - Also commonly called a judgmental sample, is one that is selected based on the knowledge of a population and the purpose of the study
    - The respondents are selected because of some characteristic
  - Focus groups
    - 10 to 20 people are brought together to answer specific questions
  - Panels
    - People are recruited to participate in surveys on an as-needed basis
- The trick is to select a sample that is representative of all voters who interest you
- Larger samples tend to reduce sampling errors when the samples are randomly selected. The statistic used to describe sampling error is called the *standard error of the mean*
- Suppose you want to compare 2 groups:
  - First, divide the population in two
  - Then, use statistical calculations to find out whether each group's sample size is large enough to pick up a difference, if one is present
  - If the population is large, you may want to select a sample and then assign persons to the 2 groups
  - o If you select a sample at random, you have random sampling
  - If you assign people to groups at random, you have random assignment
- The dependent variables are the variables for which estimates are to be made or inferences drawn
- Identify subgroups
  - The subgroups refer to the groups whose survey results must be obtained in sufficient numbers for accurate conclusions
- Identify survey type and data collection needs
  - The dependent variables tell you the content of the survey
- Check the survey's resources and schedule
  - A survey with many subgroups and measures will be more complex and costly than those with few subgroups
- Calculate sample size
- 5 questions to ask when determining sample size:

- What is the null hypothesis?
  - The null hypothesis (H<sub>0</sub>) is a statement that no difference exists between the average or mean scores of 2 groups
- What is the desired level of significance (a level) related to the null hypothesis involving the mean in the population  $(m_o)$ ?
  - The level of significance, when chosen before the test is performed, is called the *alpha value* 
    - The alpha gives the probability of rejecting the null hypothesis when it is actually true
    - Tradition keeps the alpha value small, .05, .01, or .001, to avoid rejecting a null hypothesis when it's true (and no difference exists between group means)
  - The p value is the probability that an observed result (or result of a statistical test) is due to chance (rather than to participation in a program)
    - It's calculated *after* the statistical test
- If the *p* value is less than alpha, then the null is rejected
  What chance should there be of detecting an actual difference?
  - Power is the ability to detect a difference of a given size if the difference really exists
    - It's calculated as 1 ß
    - It's defined as the probability of rejecting the null hypothesis when it's false or accepting the alternative hypothesis when it's true
    - You want high power
- What differences between means are important?
  - What is a meaningful m<sub>1</sub> m<sub>2</sub>?
  - Suppose the survey uses the Goals and Satisfaction Scale (GASS)
    - This hypothetical scale has 50 points
    - If the scale is valid, you will have access to published scoring rules and data describing what the scores mean
- What is a good estimate of the standard deviation s in the population?
  - The standard deviation (s) is a common measure of dispersion or spread of data about the mean
- Estimates of the standard deviation can come from previously done surveys
- Another way to estimate the standard deviation in your group is to conduct a small pilot test using about 25 to 50 people. You can also have experts give you estimates on the highest and lowest values or scores as the basis for calculating the standard deviation
- You can use a statistical approach to distinguish different effect sizes
  The effect size is the difference divided by the standard deviation
- The response rate is the number of persons who respond (numerator) divided by the number of eligible respondents (denominator)
  - All surveys hope for a high response rate. No single rate is considered the standard though
- Tips for improving response rate:
  - Know your respondents

- Use trained personnel to recruit respondents and conduct surveys
- Identify a larger number of eligible respondents than you need in case you don't get the sample size desired
- Use surveys only when you are fairly certain that respondents are interested in the topic
- Keep survey responses confidential or anonymous
- Send reminders to non-responders
- Provide options for completing the survey
- Provide gift or cash incentives
- Be realistic about the eligibility criteria
- Anticipate in advance the proportion of respondents who may not be able to participate because of survey circumstances or by chance
- $\circ$   $\;$  Tell respondents how you will use the survey's data  $\;$
- $\circ$   $\,$  Be precise in describing how privacy is safeguarded  $\,$
- Surveys rarely obtain information from everyone within the population of interest
  - To make up for the discrepancy, and to ensure that the survey results are representative of the population, surveyors often assign weights or values to each case
  - The 2 most common types of weights are:
    - Design weights
      - Used to compensate for oversampling or under-sampling of specific cases or for disproportionate stratification
    - Post-stratification or nonresponse weights
      - Used to compensate for the fact that persons with certain characteristics are not equally as likely to respond
  - There are many respondent characteristics that are likely to be related to the tendency to respond, including age, education, race/ethnicity, sex, and place of residence
  - It's usually more difficult to come up with post-stratification than design weights because you need supplementary information about the population. This means identifying and analyzing very large databases to get estimates of the demographic variables that interest you
- Margin of error and confidence level are 2 terms typically associated with polls
  - $\circ$   $\;$  A poll is a survey of people's opinions at a particular point in time  $\;$

#### Survey design

- Surveys may take place on their own but surveys are often included as part of a larger study, and that study may include other data collection methods as well
- Survey data can be used to describe the status of things, show change, and make comparisons
- The survey's design refers to the way in which environment is controller or organized
- The environment refers to the number and characteristics of respondents and how often and when they will be surveyed
- The more environmental control you have, the more accurate your results will be
- The environmental variables over which surveyors have the most control are:
  - When the survey is to be given
  - $\circ$  How often
  - $\circ \quad \text{The sample size} \quad$
  - The number of groups
  - How the groups were assigned
- Cross-sectional survey designs
  - With this design, data are collected at a single point in time
  - $\circ$   $\;$  Think of a cross-sectional survey as a snapshot of a group of people or organizations
  - Cross-sectional surveys have important uses
    - They describe things as they are so that people can plan for change
    - If they are unhappy with the picture that a cross-sectional survey reveals, they can change it
    - Cross-sectional surveys are also relatively easy to do
    - They are limited, however, in that if things change rapidly, the survey information will possibly become outdated
  - Longitudinal surveys
    - Longitudinal survey design refers to the collection of survey data over time
    - They are sometimes called *time-series designs*
    - A cohort design refers to surveying a particular group more than just once
    - Cohorts provide the surveyor with data on whether effects last over time
      - The problem with cohort designs is that, unless the respondents are motivated, they may drop out and the remainder may not be typical of the entire group
      - The group may be atypical in that it's willing and has the time to participate in the program and complete the surveys

- With these designs, people are divided into 2 or more groups and their survey results are compared
  - Experimental designs in some fields are sometimes divided into quasiand true experimental design
  - In quasi-experimental designs, creation and assignment to groups are usually deliberate and not random
  - In true experimental designs, and individual's assignment to groups is a chance result
  - $\circ$   $\;$  True experiments are the more powerful of the two
  - This true experimental and longitudinal design is among the most sophisticated and will enable you to make sound inferences
- Factorial designs enable researchers to evaluate different methods for offering a new program, treatment, or practice to see which combination works best
  - In a 2 x 2 design, there are 4 study groups
  - With this design, the researchers can study main effects or interactive effects
  - The outcome in this study is always the response rate
  - Factorial designs may include many factors and many levels
  - It's the number of levels that describes the name of the design
  - Survey participants in factorial designs may be selected either at random or by convenience
- Randomized controlled trials are characterized by random selection and assignment
  - If you want to know whether an experiment has results, the randomized controlled trial is the way to go
  - o Nonrandomized designs are easier to implement
- 2 lesser-known survey designs, which both offer some control over the survey's environment by making use of special comparison groups, are:
  - Normative survey design
    - A norm is a standard for comparing groups
    - Norms that are historical come from existing data or large databases such as those maintained in government registries
    - Can be less expensive and time-consuming than other comparison designs because the surveyor doesn't have to find or create a group to act as the control
    - These designs depend on the availability of appropriate databases
  - Case control design
    - A case control design is one in which groups of individuals are selected because they have (the case) or do not have (the control) the condition being studied and the groups are compared with respect to existing or past health care conditions, attitudes, habits, beliefs, or demographic factors judged to be of relevance to the causes of the condition
    - Generally used by researchers who are testing a specific hypothesis
    - Needs 2 groups a case and the control

- Ideally, the 2 groups will differ only in that the case has the characteristics or condition being studied and the control doesn't
- Most often case control designs mean selecting a control that is like the case in ways that are strongly suspected to affect the condition
- Major weakness of the case control design is that the 2 survey groups may not be alike at all no matter how selected or matches because of impossibility in controlling for all characteristics that may affect the condition
- Valid survey information comes from reliable and valid survey instruments (*measurement validity*) and from the context in which the survey takes place (*design validity*)
  - The 2 types of validity differ in that *measurement validity* refers to the characteristics of the survey instrument, whereas *design validity* refers to the selection of survey respondents and the timing and frequency of administration
- Threats to external validity are most often the consequence of the way in which respondents are selected and assigned to study groups
- When you choose and implement a survey design, you almost always automatically inherit threats to internal validity
  - The one exception is the randomized controlled trials in which respondents are randomly selected and assigned without knowing whether they are in an experimental or a control group
  - Perfect randomized trials require vast resources
- 5 potential threats to the validity of qualitative surveys:
  - The dynamics of the survey-respondent relationship can influence who is selected and who responds to the survey, and it can affect the truthfulness of the responses
  - The surveyors' preference for certain theories, designs, or issues can influence the choice of survey topics and their explanation of the findings
  - $\circ\;$  There can be bias due to how and which events, places, and people are sampled
  - Financial and human resources can affect the availability and reliability of informants
  - The surveyors' personal and professional qualities, including their experience and relationship to the community and individual people, can affect all components of the survey including how it is formatted; the sampling methods used and the questions asked; who collects, analyzes, and reports the findings; when the findings are released, to whom, and in what form

## Presenting the survey results

- The advantage of using the survey form to give results is to let the reader or listener see the questions and response choices
  - This approach shows the survey's raw data
  - It provides the data without any analysis or interpretation
  - $\circ$   $\,$  Online survey systems provide information on a real-time basis
- Survey reports nearly always have tables
  - Their purpose is to describe respondents and show relationships and changes
- Some table preparation rules
  - Tables display columns and rows of numbers, percentages, scores, and statistical test results
  - $\circ$   $\;$  Each table should have a title that summarizes its purpose and content  $\;$
  - When the source of a table's data is not immediately obvious, it should be given
  - When you use a term that may be confusing, define it. Set of definitions with asterisks or superscripts
  - $\circ$   $\,$  Columns are the independent variables, such as group and timing of the survey
  - o Select a table format and use it consistently
  - Present data is some logical order
  - Include the sample size and differentiate among numbers, percentages, and other statistics
- Pie diagrams show visually the proportion of the whole that each response category occupies
- Bar graphs are commonly used to display survey data because they provide an overview of several kinds of information at one glance
  - Bar graphs should always have a title, a legend or key to the bars, and any other explanations needed to keep the results honest
- Line graphs are drawings that allow you to show changes and compare groups
- Use pictures or diagrams to get your point across
- A fairly typical conclusion to survey's activities is a written report of its purposes, methods, results, conclusions, and recommendations
- 3 aims should guide your writing:
  - Be comprehensive
    - Organize carefully
    - Write as clearly as you can
- Organizing the report
  - Consider including all these:
    - Abstract
      The
      - The abstract is usually about 150 to 250 words
        - Consider using a structured abstract
    - Summary

- The summary is a distillation of all the report's key components (objectives, methods, findings)
  - It should be about 3 pages long and take 5 minutes to read
  - You can provide more detail than you do in the abstract and add information on who conducted the survey
  - You can also include one or two important tables of figures
- Table of contents
  - The table of contents should list al major sections of the report with the associated page numbers
- List of tables and figures
  - List each table and figure and give it complete name and the page number on which it appears
- Glossary of terms
  - All technical terms, and ambiguous concepts can be included in the glossary
- Road map: what to look for in the report and where to find it
  - Some reports are complex, and the reader needs a road map to follow them
- Statement of purposes or objectives
  - The purposes are the survey's objectives
- Outcomes
- Methods
  - The type of survey and the limitations that resulted because you chose that type of survey
  - Survey questions asked
  - Survey logistics
  - Survey construction
  - Sampling and response rate
  - Survey research design and its limitations
  - Analysis
  - Ethics
- Results
  - Usually these are given as statistics
- Conclusions
- Recommendations
- References
  - Put the references in alphabetical order or list them in the order in which they appear in your text
- Index
  - Include an index only in long reports
- Appendix
  - Include any relevant material that is too cumbersome to be contained in the report
- Acknowledgments
- Here are some tips on how to write clearly:

- Use the active voice whenever possible
- o Don't sprinkle sentences with prepositional phrases
- Try a readability formula
- Posters are an important way to present survey results
  - A poster is a summary of your work that is designed to be read and understood without oral explanation
  - A poster should present an overview of the study including a description of the survey or surveys
  - Include the standard sections of a journal article
  - A poster should have a main title that's readable from 7 meters away (25ft)