

Core Course Paper – 1: Methodology of Education Research**Chapter 1: Concept of Educational Research****1.1: Concept of Educational Research: Meaning, Characteristics and contribution to knowledge****➔ What is Educational Research?**

Generally, educational research is defined as research that investigates the behaviour of students, teachers, administrators, parents and other members of the community who interact with educational institutions. The word behaviour is taken broadly to mean such phenomena as learning, attitudes, aptitudes, abilities, interests, practices, processes, emotions and so forth.

An area that has been the focus of educational research for decades is "learning". Since education is fundamentally concerned with the development of individuals, the central purpose educational research is to find ways to improve student learning. For example, a classroom teacher experimenting with alternative ways of explaining laws of physics and a full-time researcher comparing the effectiveness of different reading methods in early literacy programmes are both engaged in a search for ways to improve student learning, and, in this sense, are both engaged in educational research.

➔ Definition

Educational research is a careful, systematic investigation into any aspect of education. From the French word "recherche" which means to travel through or survey.

The nature of educational research is analogous with the nature of research itself, which is systematic, reliable and valid to find the "truth", investigates knowledge, and solves problems (William Wiersma, 1991). Moreover, educational research process involves steps to collect the information in order to investigate problems and knowledge. However, the educational research is more complex because it can use various approaches and strategies to solve problems in educational setting. It also can involve many disciplines such as anthropology, sociology, behaviour, and history. In addition, educational research is important because of contributing knowledge development, practical improvement, and policy information (John W.Creswell, 2005). Therefore, educators can use those research findings to improve their competences and teaching and learning process.

Furthermore, the characteristics of educational research are a part of its nature. According to Gary Anderson (1998), there are ten characteristics of educational research. I tried to classify those into three categories, which are the purpose of research, the procedures of research, and the role of researcher. The purposes of research are to solve the problems, investigate knowledge, and establish the principles in educational phenomena. In short, it focuses on solving the problems and developing knowledge. Furthermore, procedure is an important characteristic of educational research, which involves collecting data with accurate observation, objective interpretation, and verification. Finally, researchers need to be experts and familiar with their field of study, using the data to develop solutions and increase knowledge. The researchers also need to be patient and careful to use every step of research's procedures to achieve the purpose of research.

➔ Purposes for Studying Educational Research

1. To orient students to the nature of educational research: its purposes, forms, and importance.
2. To provide information which helps students become more intelligent consumers of educational research: where to locate it, how to understand it and critique it.
3. To provide information on the fundamentals of doing educational research such as selecting a problem, using available tools, organizing a project, etc.

→ Characteristics of Educational Research:

Research is a way of thinking and to qualify as a research it needs to have certain characteristics such as follows (adaptation of Leedy, 1993, Borg & Borg, 1983)

→ Research begins with a question in the mind of the researcher.

You need only to look around and everywhere you see phenomena which will arouse your curiosity. For example, why are children in this school unable to read? Why are girls performing better than boys? These are situations in which the meaning of which you do not comprehend. By asking relevant questions we create an inquisitive environment which is the prerequisite for research. Research arises from a question that is intelligently asked with regards to a phenomenon that the researcher observes and is puzzling him or her.

→ Research requires a plan.

One does not discover the truth or explanations about a phenomenon without serious and meticulous planning. Research is not just by looking-up something in the hope of coming across the solution to your problem. Rather it entails a definite plan, direction and design.

→ Research demands a clear statement of the problem.

Successful research begins with a clear, simple statement of the problem. The statement of the problem should be stated precisely and grammatically complete, must set forth what it seeks to discover and enables one to see what one is attempting to research

→ Research deals with the main problem through subproblems.

Divide the main problem into appropriate subproblems, all of which when resolved will result in the solution of the main research problem.

→ Research seeks direction through appropriate hypotheses

Having stated the problem and the related subproblems, the subproblems are then each viewed through logical constructs called hypotheses. A hypothesis is a logical supposition, a reasonable guess, an educated conjecture which may give direction to thinking with respect to the problem, and thus, aid in solving it.

→ Research deals with facts and their meaning.

Having defined the problem, the subproblems and hypothesis, the next step is to collect whatever facts pertinent to the problem. Organise the data collected into a form that is potentially meaningful.

1.2: Types of Educational Research

There is not general agreement on the types of educational research which exist. The types that will be presented in this primer are as follows:

- **Ethnographic** - attempts to describe group behavior and interactions in social settings. It relies on qualitative techniques especially observation and careful recording of events and social interactions.
- **Historical** - attempts to describe and explain conditions of the past. It generally relies on qualitative data such as written documents and oral histories.
- **Descriptive** - attempts to describe and explain conditions of the present. It relies on qualitative and quantitative data gathered from written documents, personal interviews, test results, surveys, etc.
- **Correlational** - attempts to explore relationships or make predictions. It relies on quantitative data such as test scores, grade point averages, attitudinal instruments, etc. which can be correlated and shown that some relationship exists between or among them.
- **Action and Evaluation Research** - attempts to determine the value of a product, procedure, or program in a particular (e.g., school, district) setting with the goal of improving same. Action and evaluation research does not attempt to generalize results for a broader population.

- **Causal Comparative** - attempts to explore cause and effect relationships where causes already exist and cannot be manipulated. It relies on both qualitative and quantitative data such as written documents, interviews, test scores, etc.
- **Experimental** - attempts to explore cause and effect relationships where causes can be manipulated to produce different kinds of effects. It relies mostly on quantitative data such as test scores and measures of performance.

	Purpose	Examples
Qualitative	To provide rich narrative descriptions with words of phenomena that enhance understanding.	Observations of school renewal teams to understand the role of parents.
Quantitative	To describe phenomena numerically to answer specific questions or hypotheses.	The relationship between amount of homework and student achievement.
Basic	To increase knowledge and understanding of phenomena.	Understand how feedback affects motivation or learning styles of adolescents.
Applied	To solve practical educational problems.	Determine best approach to train teachers to use portfolios for assessment.
Evaluation	To make a decision about a program or activity.	Decide whether to keep or phase out a prekindergarten program.
Action	To improve practice in a school or classroom.	Determine which grouping procedure results in the highest achievement for all students.
Non-Experimental	To describe and predict phenomena without manipulating factors that influence the phenomena.	Determine the relationship between socioeconomic status and attitudes.
Experimental	To determine the causal relationship between two or more phenomena by direct manipulation of factors that influence student performance or behavior.	Determine which of two approaches to teaching science results in the highest student achievement.

1.3: Areas of Educational Research: *to be taken from external sources*

1.4: Steps of Research Problems:

No unique style, format is declared and no rigidity is found to determine the steps of educational research. However, there are some common character found almost on each research process and those could be considered as the steps of educational research.

- | | | |
|---|--------------------------------------|--|
| 1. Topic selection | 7. Proposal writing | 13. Data analysis with appropriate techniques |
| 2. Literature review | 8. Considering ethical points | 14. Writing the individual findings and then compare |
| 3. Identification of limitations in previous research | 9. Tools development | 15. Recommendation |
| 4. Rationale | 10. Data collection | 16. Suggestion for further studies |
| 5. Objective of the study | 11. Validity and reliability of data | |
| 6. Appropriate methodology | 12. Data collecting and finalizing | |

Chapter 2: RESOURCES FOR RESEARCH

2.1 Purpose of literature review:

A literature review is an evaluative report of studies found in the literature related to your selected area. The review should describe, summarize, evaluate and clarify this literature. It should give a theoretical basis for the research and help you determine the nature of your own research. Select a limited number of works that are central to your area rather than trying to collect a large number of works that are not as closely connected to your topic area.

A literature review goes beyond the search for information and includes the identification and articulation of relationships between the literature and your field of research. While the form of the literature review may vary with different types of studies, the basic purposes remain constant:

- Provide a context for the research
- Justify the research
- Ensure the research hasn't been done before (or that it is not just a "replication study")
- Show where the research fits into the existing body of knowledge
- Enable the researcher to learn from previous theory on the subject
- Illustrate how the subject has been studied previously
- Highlight flaws in previous research
- Outline gaps in previous research
- Show that the work is adding to the understanding and knowledge of the field
- Help refine, refocus or even change the topic

2.2 Major steps in literature Review:

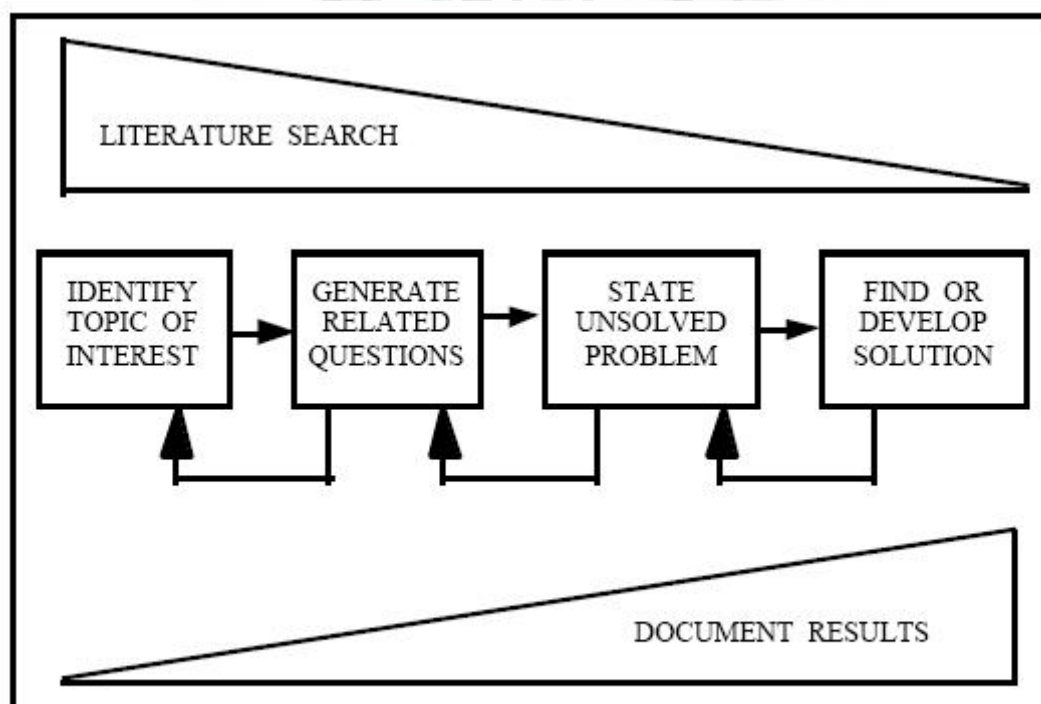


FIGURE 1. Flow Diagram of Research Process.

2.2.1 The literature review process

Table 2. The literature review process

- Selecting a review topic
- Searching the literature
- Gathering, reading and analysing the literature
- Writing the review
- References

1 Selecting a review topic

Selecting a review topic can be a daunting task for students and novice reviewers (Timmins and McCabe, 2005). A common error for novices is to select a review title that is all encompassing, such as 'pressure ulcers' or 'pain'. Although this may be a useful initial strategy for determining how much literature is available, subjects such as these generate a considerable amount of data making a review infeasible. Therefore, it is advisable to refine this further so that the final amount of information generated is manageable. For example, to focus the topic of interest, consider what aspects of pressure ulcers or pain are of particular significance. Is there a specific element of this topic that is of interest, such as prevention or management? Identifying what exactly is of interest and why can help refine the topic (Hendry and Farley, 1998). Talking to others, such as clinical specialists, or reading around a topic can also help to identify what areas of the subject the reviewer is interested in and may help indicate how much information exists on the topic (Timmins and McCabe, 2005).

Having sufficient literature is also important, particularly when the review is an academic assignment. These academic exercises usually have short deadlines, so having enough literature is key from the perspective of being able to do the review and submit it on time. Attempting to change the topic close to the deadline for submission is usually a recipe for disaster so select an area that will hold your interest and ensure that there is enough data to meet your needs.

Literature reviews that are part of academic coursework usually have strictly enforced word limits and it is important to adhere to that limit. Topics that are too broad will result in reviews that are either too long or too superficial. As a rule of thumb, it is better to start with a narrow and focused topic, and if necessary broaden the scope of the review as you progress. It is much more difficult to cut content successfully, especially if time is short.

2 Searching the literature

Having selected a topic the next step is to identify, in a structured way, the appropriate and related information. A systematic approach is considered most likely to generate a review that will be beneficial in informing practice (Hek and Langton, 2000).

3 Gathering, reading and analysing the literature

At this point of the process, what has been determined as appropriate literature will have been gathered. While the focus of the literature may vary depending on the overall purpose, there are several useful strategies for the analysis and synthesis stages that will help the construction and writing of the review. Initially, it is advisable to undertake a first read of the articles that have been collected to get a sense of what they are about. Most published articles contain a summary or abstract at the beginning of the paper, which will assist with this process and enable the decision as to whether it is worthy of further reading or inclusion. At this point, it may also be of benefit to undertake an initial classification and grouping of the articles by type of source.

Table 5. Defining the types of sources for a review

Source	Definition
Primary source	Usually a report by the original researchers of a study
Secondary source	Description or summary by somebody other than the original researcher, e.g. a review article
Conceptual/theoretical	Papers concerned with description or analysis of theories or concepts associated with the topic
Anecdotal/opinion/clinical	Views or opinions about the subject that are not research, review or theoretical in nature. Clinical may be case studies or reports from clinical settings

From: Colling (2003)

4 Writing the review

Once the appraisal of the literature is completed consideration must be given to how the review will be structured and written. The key to a good academic paper is the ability to present the findings in such a way that it demonstrates your knowledge in a clear and consistent way.

The basis of good writing is to avoid long and confusing words and keep jargon to a minimum. Sentences should be kept as short as possible with one clear message and spelling and grammar should be accurate and consistent with the form of English being used.

Many universities provide facilities for developing and improving writing skills and it is a good idea to try to attend such a course. Study skills books, such as that of Ely and Scott (2007), offer some good tips for writing competently.

The organization of material in an objective manner and the structure of the review are crucial to its comprehensiveness. To some extent, the structure will depend on the purpose of the review. For example, systematic reviews have a clear structure that must be followed and that will dictate for the most part how the writing should be undertaken.

However, for most students or practitioners a review is either part of a coursework assignment, research proposal or research dissertation, and as such, there is some freedom in how the writing is structured. Nonetheless, it is important to be logical and there are some key elements that need to be included in all literature reviews. Primarily, the written report should include an introduction, body and conclusion (Burns and Grove, 2007). The length of literature reviews vary and word limits and assignment criteria must be considered in the overall construction. If it is a stand alone review, an abstract may also be necessary. An abstract is a short summary of the findings of the review and is normally undertaken last (Hendry and Farley, 1998).

5 References:

The literature review should conclude with a full bibliographical list of all the books, journal articles, reports and other media, which were referred to in the work. Regardless of whether the review is part of a course of study for publication, it is an essential part of the process that all sourced material is acknowledged. This means that every citation in the text must appear in the reference/bibliography and vice versa. Omissions or errors in referencing are very common and students often lose vital marks in assignment

because of it. A useful strategy is to create a separate file for references and each time a publication is cited, it can be added to this list immediately.

Some universities offer their students access to referencing systems, such as Endnote, and while they may initially appear difficult to learn they are worth the effort later in terms of ensuring the reference list is accurate. Remember, the reference list may be a useful source of literature for others who are interested in studying this topic (Coughlan et al, 2007), and, therefore, every effort should be made to ensure it is accurate.

2.3 Types of Sources:

Sources are generally described as primary, secondary, or tertiary [1]. **Primary**. Primary sources are “materials that you are directly writing about, the raw materials of your own research [1].” **Secondary**. Secondary sources are “books and articles in which other researchers report the results of their research based on (their) primary data or sources [1].” **Tertiary**. Tertiary sources are “books and articles based on secondary sources, on the research of others [1].” Tertiary sources synthesize and explain the work of others and might be useful early in your research, but they are generally weak support for your own arguments [1].

Primary. Examples of primary sources are data sets, computer runs, computer programs, scale models, drawings, and engineering notebooks. A well-kept engineering notebook can provide valuable information for later documentation of test conditions and assumptions, materials used, observations as well as measurements, and unusual occurrences that prompted further testing.

Secondary. Examples of secondary sources include conferences, proceedings, journals, and books. Journal articles are often the most current source of information on a topic of study that is new or subject to rapid change. Lists of references at the end of each journal article can provide leads to further sources. Engineering journals are typically field-specific. For a selected list of current journals in agricultural, chemical, civil, computer, electrical, environmental, industrial, and mechanical engineering, ask at your facility or university library for specific guides.

Tertiary. Examples of tertiary sources include dictionaries, encyclopedias, guides, and handbooks. “Dictionaries and encyclopedias are excellent starting points for research. They can provide general background information to help narrow or broaden the focus of a topic, define unfamiliar terms, and offer bibliographies of other sources. Some works include an index, which will provide excellent access to a subject [4].” Guides and handbooks cover topics such as tables, formulas, engineering fundamentals, measures and units of conversion, mathematics, statistics, and numerical calculations; these sources are especially useful during the writing phase of your research.

2.4 Library Skills: Reading Skills and Note Taking Skills

A. Reading Skills

1. Critical Reading Skills

2. Pre Reading Strategies

3. KWL Skills

4. SQ3R Skills

1. Critical reading

Summarize, review and study your reading assignment, whether book, chapter, handout, article, whitepaper, etc.:

Characteristics of Critical Readers

- They are honest with themselves
- They resist manipulation
- They overcome confusion
- They ask questions
- They base judgments on evidence
- They look for connections between subjects
- They are intellectually independent

Ask yourself the following questions as you read:

- What is the topic of the book or reading?
What issues are addressed?
- What conclusion does the author reach about the issue(s)?
- What are the author's reasons for his or her statements or belief?
Is the author using facts, theory, or faith?

Facts can be proven

Theory is to be proved and should not be confused with fact

Opinions may or may not be based on sound reasoning

Faith is not subject to proof by its nature

- Has the author used neutral words or emotional words?
Critical readers look beyond the language to see if the reasons are clear
- Be aware of why you do, or do not, accept arguments of the author

2. Pre Reading Strategies

What you bring to the printed page

will affect how you understand what you read,

and may be what is most important in understanding what you read

Organize yourself before you read

➔ **Strategies to activate your *prior knowledge*:**

➔ **Brainstorming:**

Examine the title of the selection you are about to read

List all the information that comes to mind about this title

Use these pieces of information to recall and understand the material

Use this knowledge to reframe or reorder what you know, or to note what you disagree with, for further research

→ Group discussions:

Group discussions in and out of class will help you to discover what you bring to your reading, what your fellow students bring, as well as shared experiences

If you find they have new background information, ask for more information from them

→ Concept or mind mapping:

This is a type of brainstorming where you place the title/subject as the main idea, then develop a "mind map" around it. It can be effective either in a group or by yourself

→ Pre-questions:

Often chapters in texts provide organizing questions.

You can also write out a series of questions you expect to be answered when reading:

Examples:

Definition:

What is....? Where does ... fit? What group does ... belong to?

Characteristics:

How would I describe...? What does ... look like? What are its parts?

Examples

What is a good example of...?

What are similar examples that share attributes but differ in some way?

Experience

What experience have I had with? What can I imagine about ...?

→ Visual Aids:

Pictures and other visual material can activate your prior knowledge.

Use the Internet to search for pictures related to your title/topic to give you visual images of what you are about to read.

→ Advance Organizers:

Relate new reading material to something you already know, to your background or experiences. Ask your teacher for assistance in developing these.

→ Additional Prereading Strategies:**→ Overviews:**

Discussing information about the selection or assignment prior to reading must take place.

This may take the form of class discussions, printed previews, photographs, outlines, or films. Spend enough time before the students begin the assignment to ensure understanding of it.

→ Vocabulary Previews:

Unfamiliar key words need to be taught to students before reading so that new words, background information, and comprehension can improve together.

List all words in the assignment that may be important for students to understand. Arrange words to show the relationships to the learning task. Add words students probably already understand to connect relationships between what is known and the unknown. Share information with students. Verbally quiz them on the information before assigned reading begins.

→ **Structural Organizers:** Before reading an assignment, basic frameworks which are included in the text should be pointed out such as cause-effect or problem-solution. It can be beneficial to call attention to specific plans of paragraph or text organization such as signal words, main idea sentences, highlighted phrases, headings and subtitles. A review of skimming techniques might also be appropriate as these various areas are covered.

→ **A Purpose for Reading:** When students have a purpose for reading a selection, they find that purpose not only directs their reading towards a goal, but helps to focus their attention. Purposes may come from teacher directed questions, questions from class discussions or brainstorming, or from the individual student. Along with the question, it is a good idea to pose predictions of the outcome and problems which need to be solved. These may be generated by the student or the teacher, but the teacher should use these to guide students in the needed direction for the assigned selection.

→ **Author Consideration:** Depending upon the content area, a discussion of the author of the particular work can be helpful to the understanding of it. What is the author trying to say? What is his point of view and his reason for writing the particular work?

3. KWL: This strategy consists of three steps for students to use with expository text:

What do I **K**now? What do I **W**ant to learn? What did I **L**earn?

KWL reading method

KWL is a reading strategy formed from its letters:

Survey! Question! Read! Recite! Review!

KWL is intended to be an exercise for a study group or class that can guide you in reading and understanding a text.

You can adapt it to working alone, but discussions definitely help.

It is composed of only three stages that reflect a worksheet of three columns with the three letters:

What we K now	what we W ant to know	what we L earned

K stands for Know

This first stage may surprise you:

Think first about, then list, what you know about the topic before reading!

This advanced organizer provides you with a background to the new material, building a scaffold to support it.

Think of it as a pre-reading inventory.

- **Brainstorm!**
Before looking at the text, think of keywords, terms, or phrases about the topic, either in your class or a study group.
- Record these in the **K** column of your chart until you cannot think of more.
- Engage your group in a discussion about what you wrote in the **K** column.
- Organize the entries into general categories.

W stands for Will or Want

The second stage is to list a series of questions of what you want to know more of the subject, based upon what you listed in **K**.

- Preview the texts table of contents, headings, pictures, charts etc.
Discuss what you want to learn
- List some thoughts on what you want, or expect to learn, generally or specifically.
Think in terms of what you will learn, or what do you want to learn about this.
- Turn all sentences into questions before writing them down.
They will help you focus your attention during reading.
- List the questions by importance.

L stands for Learned

The final stage is to answer your questions, as well as to list what new information you have learned. Either while reading or after you have finished.

- List out what you learn as you read, either by section, or after the whole work, whichever is comfortable for you.
- Check it against the **W** column, what you wanted to learn
- Create symbols to indicate main ideas, surprising ideas, questionable ideas, and those you don't understand!

Expand this exercise beyond K W L:

Add an H!

Stands for **HOW** you can learn more.

- Pose new questions about the topic
- How can I learn more or answer questions not answered in my worksheet
These include other sources of information, including: organizations, experts, tutors, websites, librarians, etc.

5 W's and an H

Another reading strategy is to answer the questions that form the basis of good journalism:

Who What When Where Why and How

- Who are the main characters?
- What does the author say happened?
- Where did the action occur?
- When did it happen or what is the span of time?
- Why did this happen?
- How did it happen?

I keep six honest serving-men

(They taught me all I knew); Their names are What and Why and When And How and Where and Who.

I send them over land and sea, I send them east and west; But after they have worked for me, I give them all a rest.

4. SQ3R reading method

SQ3R is a reading strategy formed from its letters:

Survey! Question! Read! Recite! Review!

SQ3R will help you build a framework to understand your reading assignment.

→ **Before you read, Survey the chapter:**

- the title, headings, and subheadings
- captions under pictures, charts, graphs or maps
- review questions or teacher-made study guides
- introductory and concluding paragraphs
- summary

→ **Question while you are surveying:**

- Turn the title, headings, and/or subheadings into questions
- Read questions at the end of the chapters or after each subheading
- Ask yourself,
"What did my instructor say about this chapter or subject when it was assigned?"
- Ask yourself,
"What do I already know about this subject?"

Note: If it is helpful to you, write out these questions for consideration.
This variation is called SQW3R

→ **When you begin to Read:**

- Look for answers to the questions you first raised
- Answer questions at the beginning or end of chapters or study guides
- Reread captions under pictures, graphs, etc.
- Note all the underlined, italicized, bold printed words or phrases
- Study graphic aids
- Reduce your speed for difficult passages
- Stop and reread parts which are not clear
- Read only a section at a time and recite after each section

→ **Recite after you've read a section:**

- Orally ask yourself questions about what you have just read, or summarize, in your own words, what you read
- Take notes from the text but write the information in your own words
- Underline or highlight important points you've just read
- Reciting:
The more senses you use the more likely you are to remember what you read Triple strength learning: Seeing, saying, hearing
Quadruple strength learning: Seeing, saying, hearing, writing!!!

Review: an ongoing process**Day One**

- After you have read and recited the entire chapter, write questions in the margins for those points you have highlighted or underlined.
- If you took notes while reciting, write questions for the notes you have taken in the left hand margins of your notebook.
- Complete the form for a [critical reading review](#)

Day Two

- Page through the text and/or your notebook to re-acquaint yourself with the important points.
- Cover the right hand column of your text/note-book and orally ask yourself the questions in the left hand margins.
- Orally recite or write the answers from memory.
- Develop mnemonic devices for material which need to be memorized. Make flash cards for those questions which give you difficulty.

Days Three, Four and Five

- Alternate between your flash cards and notes and test yourself (orally or in writing) on the questions you formulated.
- Make additional flash cards if necessary.

Weekend

- Using the text and notebook, make a Table of Contents - list all the topics and sub-topics you need to know from the chapter.
- From the Table of Contents, make a Study Sheet/ Spatial Map.
- Recite the information orally and in your own words as you put the Study Sheet/Map together.
- As you have consolidated all the information you need for this chapter, periodically review the Sheet/Map so that at test time you will not have to cram.

B. Note Taking Skills

Both *Brusaw, Charles T. and Gerald J. Alred and Walter E. Oliu, 1993. Handbook of Technical Writing, Fourth Edition. New York: St. Martins Press.*

and
Booth, Wayne C. and Gregory G. Colomb and Joseph M. Williams, 1995. The Craft of Research. Chicago: The University of Chicago Press. have good advice about notetaking.

Brusaw, Charles T. and Gerald J. Alred and Walter E. Oliu, states that the challenge is to condense others' work without distorting it.

These seven-step process includes:

- i) do not write everything down,
- ii) create your own shortcuts and shorthand,
- iii) use numbers for numerical terms,
- iv) leave out vowels when you can,
- v) record all vital names, dates, and definitions,
- vi) mark items that need further examination, and
- vii) check accuracy before returning or filing the source.

Booth, Wayne C. and Gregory G. Colomb and Joseph M. Williams, 1995. The Craft of Research. Chicago: The University of Chicago Press.:

recommends writing and summarizing as you find sources and has a three-page section called "Quick Tip: Speedy Reading."

It recommends a five-step process:

- i) become familiar with the geography of the source,
- ii) locate the point of the argument,
- iii) identify key subpoints,
- iv) identify key themes, and
- v) skim paragraphs.

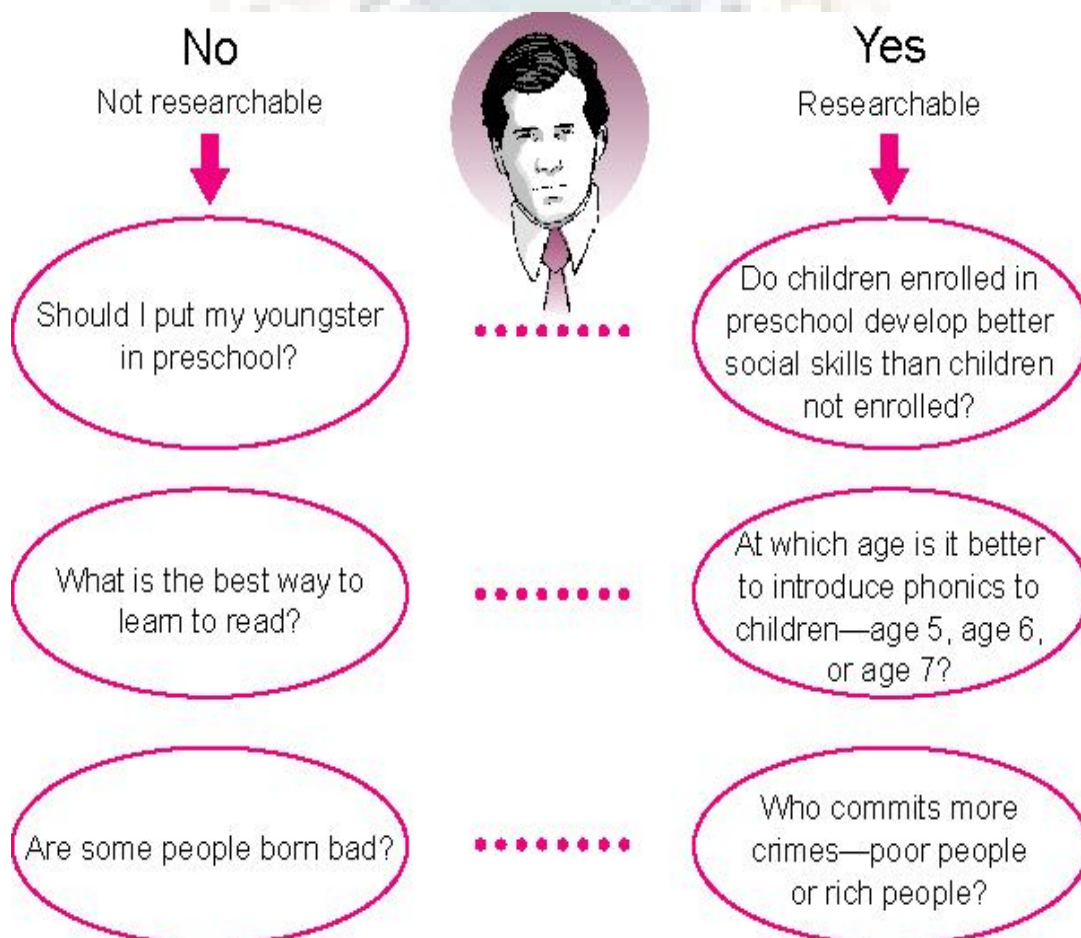
The book suggests that not all five steps are needed all the time. The main point is speed and efficiency, and focusing on material relevant to your own research without spending time on material that is at best only marginally related. While taking notes and preparing condensed summaries of the work of others, you must be ever mindful of the requirement to eventually cite all "borrowed" work in your final paper. All of the sources mentioned contain discussions about direct quotations, summarizing, and plagiarism. Readers and researchers are advised to locate and carefully read about these topics from any available source in order to avoid trouble while writing.

2.5 Internet Search: to be taken from external resources

Chapter 3: SELECTION OF THE RESEARCH PROBLEM

3.1 Standards of Selection: Novelty, Uniqueness, Originality and Researcher Skills

- The topic should be *important (significant)*
 - Writing a thesis or dissertation is an exercise to learn how to conduct research.
 - However, graduate students can learn the research process on an important topic just as easy as learning the research process on a piddle topic!
- Consider the *feasibility* of the project.
 - How much *time* do you have available
 - Do you really want to do a longitudinal study that will take 3 years to complete for a MS thesis?
 - How *difficult* is it. Are data available?
 - How much will it *cost*?
- Make sure the topic is *ethical* to study.
- You should have a *personal interest* in the topic.
By the time you are done, you may really be tired of the topic
- The “*newness*” of the topic may hold you interest longer, however there is some value in repeating previous research
- Make sure the research question is *clear*.



3.2 SOURCES OF THE SELECTION OF PROCESS

Problems can be conceptualized at a number of levels. At one level we are all searching for the answer to the question “why do organisms behave the way they do?” Questions with more specificity could proceed from; why do humans behave the way they do, to how does reinforcement affect behavior, to how does reinforcement affect studying, to how does reinforcement affect studying for tests in university males.

Beginners tend to start with relatively specific research problems focused on the face value of the question, but eventually develop a broad research question with great generality. For example, what started as “how can I help my roommate study more?” evolves into “what controls studying in people?” At the beginning, the roommate’s behavior is at issue for itself. Later the person and the behavior are seen as arbitrary instances of a much more important and challenging question. Career long research problems tend to emerge following several years of specific research topics, and require many specific research studies to solve. This section details some of the sources for an initial, relatively specific, research problem. It is intended to help you come up with research which is manageable, enjoyable, and productive.

In addition to not knowing what unresolved problems remain, is missing the more fundamental broader issue underlying any specific behavior change. When looking at the world, try to see each functional relationship as only an instance of a more general class of relationships.

A. Research Problem from Expert

The simplest source of a problem to solve is to have it given to you as a class assignment, as a directed research project, or as a task while you are an apprentice in someone's lab. You are told what problem to research and how to do it. This is probably an ideal way to assure that your first research topic is a good one.

Example: Students in Experimental Psychology were assigned the task of finding out if social attention made their roommate study more. They were told to measure the amount of time their roommate studied on days during which they expressed interest in their roommate's course material as compared to days when they refrained from talking about academic topics.

B. Research Problem from Folklore

Common beliefs, common sense, or proverbs could be right but on the other hand, they could also be wrong. You must verify that they are true before considering them as a source of knowledge. It is possible that some unverified beliefs have the roots of a better idea and therefore would be a worthy research topic. It is critical to note, however, that the task of research is not to simply validate or invalidate common sense but rather to come to understand nature.

Example: It's commonly believed that studying within the two hours preceding a test will decrease test scores. To research this belief a randomly selected half of a class was told to study immediately before taking a test while the other half was prohibited from studying before the test. This research was intended to determine whether or not studying immediately before a test decreased the points earned.

C. Research Problem From Insight

Sometimes people research an issue simply because it occurred to them and it seemed important. The systematic development of the idea is lacking. This is “intuitive” or good guess research. It is risky because you may not be able to get other researchers to understand why the research is important. It is fun because you get to do what interests you at the moment. Alternatively, it could be the application of a general rule of thumb or guessing that a new problem is actually a well-understood function in disguise.

Example: While feeling especially competent after explaining course material to three friends you realize that orally presenting material may help test performance. You conducted a study in which material was orally presented before the test on a random half of the occasions. The research was based on your insightful realization that oral presentation may increase test performance.

D. Research Problem from Informal Discussion

This is a research problem that some discussion group feels is interesting. Discussion among friends can often spark our interest in a problem or provides us with the reinforcers for pursuing a question.

Example: After telling a group of friends about your success with oral presentations on test taking, the group talks about it for awhile and becomes interested in the possibility of the subject becoming confused as well as doing better as a result of feedback from the listeners. The group provides you with the idea and the excitement to do research on how students can affect the accuracy of a teacher's understanding.

E. Research Problem from Knowledge of Techniques and Apparatus

This is the selection of a research topic based on your special knowledge outside the field. A technique or apparatus with which you are familiar can offer the potential for a major advance in the field of psychology. Sometimes we realize that we can apply a new technique or apparatus to an area to which it has not yet been applied. Because we are specially qualified to succeed, solving the problem can be especially gratifying.

Example: You may know about microelectronics and be good at detailed work. You find out that many researchers are anxious to discover the migration patterns of butterflies so you mount an integrated circuit transmitter on a butterfly and thereby trace the behavior of the free ranging butterfly.

F. Research Problem from Reading the Literature

These are research problems which capture your interest while reading. While reading you will often wonder why, or will disagree, or will realize that you have a better idea than the original author.

Example: While you were reading about jet lag and its effects on sleep the first night, you realize that the author failed to control for light cycle. You try stretching either the light period or stretching the dark period to make up the phase shift. You implement this by changing the cabin illumination period on various trans-Atlantic flights, and monitoring the passengers sleep for the next three days.

3.3 NARROWING THE SCOPE OF PROBLEM

Whether it's assigned to you or chosen by you, narrowing down your topic is the first step in writing a research paper. Without a focused and specific question, you might produce a response that is vague and too general to adequately address it in the space and time provided. There are numerous strategies available to help you narrow down your topic and focus your research accordingly.

Instructions

1. Conduct research. Use the Internet, or go to the local library to browse journals and books in your topic area. Look for sub-topics or current debates that you could use to narrow down your focus. Alternatively, search periodicals for recent events in the field, such as discoveries or breakthroughs.
2. Brainstorm the topic. Write down your research topic in the center of a sheet of paper. Write down as many related words and ideas as you can. Brainstorming crime, for example, will produce ideas like murder, theft, victims and prison. When you've finished, look at your answers and categorize them as best as you can. Collate words and ideas that are related, such as types of crime, prison sentences, effects on the victim or society in general. You can then investigate one or more of these narrower, more focused topics.
3. Focus on one aspect of the topic. This could be psychological, economic, social, cultural or political. For a research paper on eating disorders, for example, you could focus on the cost implications of treatment, or its emotional effects on patients.

4. Consider time and place. Another way to break down a topic is to focus on a particular time period, like the Middle Ages, the 19th century or even the 1980s. A research paper on women's health, for example, can be narrowed to women's sexual health in the 1700s. You can also narrow the topic further with geographical boundaries, such as women's sexual health in the 1700s in England.
5. Focus on a social group. Narrow your topic area down to include a specific gender, age and cultural or ethnic background. A paper on eating disorders, for example, could be narrowed down to focus on children, teenagers, African-American women or Chinese men. You can also compare and contrast between two or more groups throughout your research.

How to Narrow Your Topic

Example: I'm thinking of doing a paper on "fashion." This topic could develop in many different ways.

Hint: Ask Yourself Questions About Your Topic:

- What do you know about it? What don't you know?
- What aspects of your topic interest you: historical, sociological, psychological, etc.?
- What time period do you want to cover?
- On what geographic region do you want to focus?
- What kind of information do you need?
 - A brief summary or a lengthy explanation?
 - Periodical articles, books, essays, encyclopedia articles?
 - Statistics?

Sample Topic Narrowing Chart:

General Topic:	fashion
Time span:	1920s
Place:	US; urban; big cities (not rural)
Person or group:	youth; college age
Event or Aspects:	sexual attitudes; behavior; sociological

3.4 DRAFTING A RESEARCH PROPOSAL: *to be taken from sir*

3.6 DEFINING THE RELATED TERMS OF RESEARCH PROBLEM: *to be taken from sir*

Chapter 4: VARIABLES AND HYPOTHESES

4.1 MEANING AND TYPES OF VARIABLES

Constant - a characteristic or condition that is the same for all individuals in a study.

Variable - a characteristic that takes on different values/conditions for different individuals.

Types of Variables

Independent Variable - a variable that affects the dependent variable under study and is included in the research design so that its effects can be determined. (Also known as a predictor variable in certain types of research.)

Levels of the Variable - describes how many different values or categories an independent has in a research design.

Dependent Variable - a variable being affected or assumed to be affected by an independent variable. (Variable used to measure the effects of independent variables. Also known as an outcome variable in certain types of research.)

Organismic Variable - a preexisting characteristic of an individual that cannot be randomly assigned to that individual (e.g. gender). Serve as control variables only when effects are known/predetermined.

Intervening Variable - a variable whose existence is inferred, but which cannot be manipulated or directly measured. Also known as nuisance variables, mediator variables, or confounding variables.

Control Variable - an independent variable not of primary interest whose effects are determined by the researcher. (May be included in the research design to help explain variation in results.)

Moderator Variable - a variable that may or may not be controlled, but has an effect on the research situation.

! when controlled - control variable (effects are known)

! when uncontrolled - intervening variable (effects unknown)

VARIABLES AS THEY ARE MEASURED

Nominal - measurement scale in which numbers are used as names of categories; i.e., categorizes without order. (Frequency data)

Ordinal - measurement scale that categorizes AND indicates relative amount or rank-order of a characteristic. (Ordered data)

Interval - measurement scale that categorizes, indicates relative amount or rank-order of a characteristic, and has units of equal distance between consecutive points on the scale. (Score data)

Ratio - measurement scale that categorizes, indicates relative amount or rank-order of a characteristic, has units of equal distance between consecutive points on the scale, and compares terms as ratios of one to another (i.e. has a true zero point). Rarely used in social science research.

NOTE - measurement scales are hierarchical, with each higher level of measurement provided greater information and permitting more stringent statistical analyses to be performed.

OPERATIONAL DEFINITION - a definition expressed in terms of the processes or operations and conditions that are being used to measure the characteristic under study.

4.2 MEANING AND TYPES OF HYPOTHESES

Research problem statement - by itself provides general direction for the study--it does not include all the specific information.

Hypothesis - a conjecture or proposition about the solution to a problem, the relationship of two/more variables, or the nature of some phenomenon (i.e. an educated guess based on available fact).

A good hypothesis should:

1. state an expected relationship between two or more variables
2. be based on either theory or evidence (and worthy of testing)
3. be testable
4. be as brief as possible consistent with clarity
5. be stated in declarative form
6. be operational by eliminating ambiguity in the variables or proposed relationships

Types and Forms of Hypotheses

Research (Substantive) Hypothesis - simple declarative statement of the hypothesis guiding the research.

Example: Members of the general population are, on average, less intelligent than USD students.

Statistical Hypothesis:

- (1) a statement of the hypothesis given in statistical terms.
- (2) a statement about one or more parameters that are measures of the population under study.
- (3) a translation of the research hypothesis into a statistically meaningful relationship.

Example: The mean intelligence of members of the general population is lower than the mean intelligence of USD students.

Null Hypothesis - a statistical hypothesis stated specifically for testing (which reflects the no difference situation).

Example: There is no significant difference in mean intelligence between members of the general population and USD students.

Alternative Hypothesis - an alternative to the null hypothesis that reflects a significant difference situation.

Example: There is a significant difference in mean intelligence between members of the general population and USD students.

Directional Hypothesis - a hypothesis that implies the direction of results.

Example: The mean intelligence of members of the general population is significantly lower than that of USD students.

Nondirectional Hypothesis - a hypothesis that does not imply the direction of results.

Foreshadowed Problems - (in ethnographic research) statements of specific research problems that provide a focus for the research. They identify factors for the researcher to consider without specifying anticipated results.

Research Questions - relatively narrow, specific delineations of what the proposed research will address. Questions emerge from the researcher's topic of interest plus information gathered during the literature review.

- Emerge from either theory or evidence (literature review)
- Operationalize the research problem by identifying variables and/or the relationships among variables
- Be answerable

NOTE: While hypotheses generally require Yes-No responses (prove-disprove), research questions often require descriptive or correlational information.

4.3 HYPOTHESES CONSTRUCTION: BASES, INVOLVING VARIABLES

(to be taken from sir or external resrouces)

4.4 TESTING A HYPOTHESIS

A **statistical hypothesis test** is a method of making decisions using data, whether from a [controlled experiment](#) or an [observational study](#) (not controlled). In [statistics](#), a result is called [statistically significant](#) if it is unlikely to have occurred by [chance](#) alone, according to a pre-determined threshold probability, the [significance level](#). The phrase "*test of significance*" was coined by [Ronald Fisher](#): "Critical tests of this kind may be called tests of significance, and when such tests are available we may discover whether a second sample is or is not significantly different from the first."^[1]

Hypothesis testing is sometimes called **confirmatory data analysis**, in contrast to [exploratory data analysis](#). In [frequency probability](#), these decisions are almost always made using [null-hypothesis](#) tests (i.e., tests that answer the question *Assuming that the null hypothesis is true, what is the probability of observing a value for the test statistic that is at least as extreme as the value that was actually observed?*)^[2] One use of hypothesis testing is deciding whether experimental results contain enough information to cast doubt on conventional wisdom.

A result that was found to be statistically significant is also called a **positive result**; conversely, a result that is not unlikely under the null hypothesis is called a **negative result** or a **null result**.

Statistical hypothesis testing is a key technique of [frequentist statistical inference](#). The Bayesian approach to hypothesis testing is to base rejection of the hypothesis on the [posterior probability](#).^[3] Other approaches to reaching a decision based on data are available via [decision theory](#) and [optimal decisions](#).

The *critical region* of a [hypothesis test](#) is the set of all outcomes which cause the [null hypothesis](#) to be rejected in favor of the [alternative hypothesis](#). The critical region is usually denoted by the letter *C*.

Types of Tests

Tests of hypothesis can be carried out on one or two samples. One sample tests are used to test if the population parameter (μ) is different from a specified value. Two sample tests are used to detect the difference between the parameters of two populations (μ_1 and μ_2).

Two sample tests can further be classified as unpaired or paired two sample tests. While in unpaired two sample tests the sample data are not related, in paired two sample tests the sample data are paired according to some identifiable characteristic. For example, when testing hypothesis about the effect of a treatment on (say) a landfill, we would like to pair the data taken at different points before and after implementation of the treatment.

Both one sample and two sample tests can be classified as :

One tailed test : Here the alternate hypothesis H_A is one-sided and we test whether the test statistic falls in the critical region on only one side of the distribution.

1. One sample test: For example, we are measuring the concentration of a lake and we need to know if the mean concentration of the lake is greater than a specified value of 10mg/L.
Hence, $H_0: \mu \leq 10 \text{ mg/L}$, vs, $H_A: \mu > 10 \text{ mg/L}$.
2. Two sample test: In Table1, cases 2 and 3 are illustrations of two sample, one tailed tests. In case 2 we want to test whether the population mean of the first sample is lesser than that of the second sample.
Hence, $H_0: \mu_1 \geq \mu_2$, vs, $H_A: \mu_1 < \mu_2$.

Two tailed test : Here the alternate hypothesis H_A is formulated to test for difference in either direction, i.e., for either an increase or a decrease in the random variable. Hence the test statistic is tested for occurrence within either of the two critical regions on the two extremes of the distribution.

1. One sample test: For the lake example we need to know if the mean concentration of the lake is the same as or different from a specified value of 10 mg/L.
Hence, $H_0: \mu \leq 10 \text{ mg/L}$, vs, $H_A: \mu \neq 10 \text{ mg/L}$.
2. Two sample test: In Table 1, case 1 is an illustration of a two sample two tailed test. In case 1 we want to test whether the population mean of the first sample (μ_1) is the same as or different from the mean of the second sample (μ_2).
Hence $H_0: \mu_1 = \mu_2$, vs, $H_A: \mu_1 \neq \mu_2$.

Given the same level of significance the two tailed test is more conservative, i.e., it is more rigorous than the one-tailed test because the rejection point is farther out in the tail. It is more difficult to reject H_0 with a two-tailed test than with a one-tailed test.

Steps in Hypothesis Testing

1

Identify the null hypothesis H_0 and the alternate hypothesis H_A .

2

Choose α . The value should be small, usually less than 10%. It is important to consider the consequences of both types of errors.

3

Select the test statistic and determine its value from the sample data. This value is called the observed value of the test statistic. Remember that a t statistic is usually appropriate for a small number of samples; for larger number of samples, a z statistic can work well if data are normally distributed.

4

Compare the observed value of the statistic to the critical value obtained for the chosen α .

5

Make a decision.

If the test statistic falls in the critical region:

Reject H_0 in favour of H_A .

If the test statistic does not fall in the critical region:

Conclude that there is not enough evidence to reject H_0 .

Chapter 5: TOOLS OF RESEARCH

5.1: Psychological Test: (a) Types and its uses (b) General Information about Construction Procedure

(A) Types & Uses

Psychological tests can be grouped into several broad categories. Personality tests measure personal qualities, sometimes referred to as traits. Achievement tests measure what a person has learned. Aptitude tests are designed to predict future behaviour, such as success in school or job performance. Intelligence tests measure verbal and/or nonverbal skills related to academic success. Interest inventories are used to help individuals make effective career choices.

Psychological tests are usually administered and interpreted by a psychologist because studies in psychopathology, along with academic courses and supervision in psychological testing, are an integral part of the doctoral degree in clinical psychology. A counsellor who has had the appropriate academic courses and supervision may administer occupational tests or achievement and aptitude tests, but most counselors have not received the training to administer personality tests. Academic courses and supervision in psychological testing are usually not a part of a psychiatrist's medical training, so most psychiatrists can ethically administer only some specific clinical tests that are straight-forward check-lists of symptoms.

Of course, ethics is one thing, and the desire to make money is another thing. Therefore you will often find individuals offering to do all kinds of psychological testing—often on the Internet—even when they lack the training to administer and interpret such tests.

Psychological tests fall into several categories:

1. Achievement and aptitude tests are usually seen in educational or employment settings, and they attempt to measure either how much you know about a certain topic (i.e., your achieved knowledge), such as mathematics or spelling, or how much of a capacity you have (i.e., your aptitude) to master material in a particular area, such as mechanical relationships.

Intelligence tests attempt to measure your intelligence, or your basic ability to understand the world around you, assimilate its functioning, and apply this knowledge to enhance the quality of your life. Or, as Alfred Whitehead said about intelligence, “it enables the individual to profit by error without being slaughtered by it.”

Intelligence, therefore, is a measure of a potential, not a measure of what you’ve learned (as in an achievement test), and so it is supposed to be independent of culture. The trick is to design a test that can actually be culture-free; most intelligence tests fail in this area to some extent for one reason or another.

2. Neuropsychological tests attempt to measure deficits in cognitive functioning (i.e., your ability to think, speak, reason, etc.) that may result from some sort of brain damage, such as a stroke or a brain injury.

3. Occupational tests attempt to match your interests with the interests of persons in known careers. The logic here is that if the things that interest you in life match up with, say, the things that interest most school teachers, then you might make a good school teacher yourself.

4. Personality tests attempt to measure your basic personality style and are most used in research or forensic settings to help with clinical diagnoses. Two of the most well-known personality tests are the Minnesota Multiphasic Personality Inventory (MMPI), or the revised MMPI-2, composed of several hundred “yes or no” questions, and the Rorschach (the “inkblot test”), composed of several cards of inkblots—you simply give a description of the images and feelings you experience in looking at the blots.

Personality tests are either objective or projective.

Objective Tests

Objective tests present specific questions or statements that are answered by selecting one of a set of alternatives (eg. true or false). Objective tests traditionally use a "paper-and-pencil" format which is simple to score reliably. Although many objective tests ask general questions about preferences and behaviours, situational tests solicit responses to specific scenarios.

The MMPI - The Minnesota Multiphasic Personality Inventory is the leading objective personality test. Its hundreds of true-false items cover a broad range of behaviours. A major advantage of the MMPI is the incorporation of validity scales designed to detect possible response bias, such as trying to present oneself in a socially desirable way.

Projective Techniques

Projective personality tests use ambiguous stimuli into which the test taker presumably projects meaning. This indirect type of assessment is believed by many to more effectively identify a person's real or underlying personality.

a. Scoring Projective Techniques

Because the test taker is free to respond in any way, rather than being required to select an answer from a set of alternatives, projective tests can be difficult to score.

To ensure reliability, projective tests must be accompanied by a specific set of scoring criteria. Projective tests are more reliable and valid when scoring focuses on the way the questions are answered (structure of responses) rather than the content of the answers.

Two leading projective tests are the Rorschach and the Thematic Apperception Test (TAT).

b. The Rorschach Test

In the Rorschach, individuals are asked to describe in detail their impressions of a series of inkblots. Scoring involves analysis of both the structure and content of responses.

c. The Thematic Apperception Test (TAT)

In the TAT, individuals construct stories to describe a series of pictures. TAT analysis traditionally focuses on the role played by the main character in each story.

5. Specific clinical tests attempt to measure specific clinical matters, such as your current level of anxiety or depression.

Functions and Uses of Psychological testing

1. Detection of specific Behavior

Psychological test is used to measure and to detect the abilities of a person.

2. Individual Differences

A psychological test is used to measure the individual differences, that is different between abilities of different persons and the performance of the same person at different time.

3. To diagnose by the Psychological Test

The psychological tests are usually used in clinical psychology. In clinical psychology a test's function is to diagnose mental disorders. So tests are used in mental hospitals and coaching and guidance centers for the assessment and diagnosis of mental disorders. Major tests are MMPI, (Minnesota Multiphasic Personality Inventory) RISB, (Roter Incomplete Sentences Blank) Bender Gestalt Test, and RPM, (Revon Progressive Matrices) etc.

4. Legal Classification

A psychological test helps in classifying a number of people into different categories. For example normal and abnormal, criminal and innocent, intellectual and mental retarded, able and disable etc.

Methods of Legal Classification

Selection

The person who express certain level of performance on a test are selected and others are rejected.

Screening

An ordinary test refers to a quick survey to located individuals who may need or be eligible for special treatment.

Certification

At the end of certain training program a test is used to recommended that the objectives of training program has been achieved and the person has the acquired the desired skill to perform in the relevant field.

Placement

Placement is a sorting process that provides different level of serving for different persons.

5. Promoting Self Understanding

A psychological test provide standardized information about the abilities, capabilities, aptitudes, potential competencies interest, trait and states of a person which helps in understanding one's personality and planing future prospective.

6. Program Evaluation

An effectiveness of a particular program is assessed by the applications of some kind of test. This function is usually perform by an achievement test.

7. Scientific Inquiry or Research

Some experts use tests for research purpose which provide information about the mental level and personality of the subject.

8. Military Selection

A closely related application of psychological testing is to be found in the selection and classification of military personnel. From simple beginnings in the World War-I, the scope and variety of psychological tests employed in military situations underwent a phenomenal increase during World War-II. Subsequently research on test development has been continuing on a large scale in all branches of the armed services.

9. Industry

In industry and business tests are helpful in selection and classifying personnel for placement in jobs that range from the simpler semiskilled to the highly skilled, from the selection of filling clerks and sales-person to top management for any of these positions, however test results are only one source of information, though an important one.

Industrial and Organizational Psychology (also known as I-O psychology, industrial-organizational psychology, work psychology, organizational psychology, work and organizational psychology, industrial psychology, occupational psychology, personnel psychology or talent assessment) applies psychology to organizations and the workplace. In January 2010, the Society for Industrial and Organizational Psychology (SIOP) announced that, as a result of a membership vote, it would retain its name and not change it to the Society for Organizational Psychology (TSOP) to eliminate the word "Industrial". "Industrial-organizational psychologists contribute to an organization's success by improving the performance and well-being of its people. An I-O psychologist researches and identifies how behaviors and attitudes can be improved through hiring practices, training programs, and feedback systems."

I-O psychology is represented by Division 15 of the American Psychological Association.

Common research and practice areas for I-O psychologists include:

- Job performance
- Job analysis
- Personnel recruitment and selection
- Performance appraisal/management
- Individual assessment (knowledge, skills, and ability testing, personality assessment, work sample tests, assessment centers)
- Psychometrics
- Compensation
- Training and training evaluation/Development
- Employment law
- Work motivation
- Job attitudes (e.g., job satisfaction, commitment, organizational citizenship, and retaliation)
- Occupational health and safety
- Work/life balance
- Human factors and decision making
- Organizational culture/climate
- Organizational surveys
- Leadership and executive coaching
- Ethics
- Diversity
- Job design
- Human resources
- Organizational development (OD)
- Organizational research methods
- Technology in the workplace
- Group/team performance
- Employee safety and health

I-O psychologists are trained in the “scientist-practitioner” model. The training enables I-O psychologists to employ scientific principles and research-based designs to generate knowledge. They use what they have learned in applied settings to help clients address workplace needs. I-O psychologists are employed as professors, researchers, and consultants. They also work within organizations, often as part of a human resources department where they coordinate hiring and organizational development initiatives from an evidence-based perspective.

Industrial psychology is a relatively new branch of psychology that was created for corporations and organizations that needed more structure. Industrial psychology is able to provide this structure by evaluating employee behavior for the good of the company. It is often referred to as organizational psychology because of its emphasis on analyzing individuals who work for various organizations.

Essentially, industrial psychologists study the behavior of employees in a work setting. Although industrial psychology didn't begin until the 1920's, the discipline has evolved rapidly and revolutionized the workplace within the last century. Because the workplace is a social system, the application of industrial psychology is useful in understanding its complexity.

10. Education

Psychological tests especially those of general intelligence and of specific aptitudes have very extensive use in educational classification, selection and planning from the 1st grade (and sometimes earlier) through the university. Prior to World War-II schools and colleges were the largest user of psychological tests.

"Now at present schools are among the largest test users".

Educational psychology is the study of how humans learn in educational settings, the effectiveness of educational interventions, the psychology of teaching, and the social psychology of schools as organizations. Educational psychology is concerned with how students learn and develop, often focusing on subgroups such as gifted children and those subject to specific disabilities. Although the terms "educational psychology" and "school psychology" are often used interchangeably, researchers and theorists are likely to be identified in the US and Canada as educational psychologists, whereas practitioners in schools or school-related settings are identified as school psychologists. This distinction is however not made in the UK, where the generic term for practitioners is *"educational psychologist."*

Educational psychology can in part be understood through its relationship with other disciplines. It is informed primarily by psychology, bearing a relationship to that discipline analogous to the relationship between medicine and biology. Educational psychology in turn informs a wide range of specialties within educational studies, including instructional design, educational technology, curriculum development, organizational learning, special education and classroom management. Educational psychology both draws from and contributes to cognitive science and the learning sciences. In universities, departments of educational psychology are usually housed within faculties of education, possibly accounting for the lack of representation of educational psychology content in introductory psychology textbooks.

11. Mental Hospitals

In clinical or mental hospitals psychological tests are used primarily for individual diagnoses of factors associated with personal problems of learning, behavior attitudes or specific interpersonal relations.

Another areas are JAILS, CORRECTIONAL CENTERS and TRAINING INSTITUTIONS etc.

(B) General Information about construction of Psychological Test:

Here are the basic steps to constructing a useful psychological test:

1) Determine the trait, ability, emotional state, disorder, interests, or attitude that you want to assess.

Psychological tests can be created that measure --

Abilities, such as musical skill, writing skill, intelligence, or reading comprehension.

Personality Traits, such as extroversion, creativity, or deviousness,

Disorders, such as anxiety, depression, psychotic thought disorder,

Emotions, such as happiness and anger,

Attitudes, such as authoritarianism or prejudice,

Interests, such as career-related interests.

2) Decide how you want to measure the construct you selected. In general, the best measures sample the behavior of interest. For instances, if you want to determine how aggressive a person is, the best measure would be to provide a frustrating situation, and see whether the person reacts aggressively. It's not always practical or ethical to directly measure constructs, so instead, tests rely on a person's self-report of their behavior.

A number of other factors need to be considered. Should the test be written, or should it be administered orally? Should the responses be discrete (a rating scale, or Yes/No answers), or should it allow open-ended answers that can be reliably rated? Should the responses be oral, written, or nonverbal?

3) Does the construct that you want to measure have only one dimension, or can it be broken down into several dimensions? For instance, intelligence is usually considered multi-dimensional, consisting of several different verbal abilities and nonverbal abilities.

4) Once you've made decisions about the factors above, you can begin creating your test items. If the items are measuring a particular area of knowledge, then you will review textbooks or consult subject-matter experts in that area. If you are measuring a personality trait or emotional state, then the items should be consistent with a theory or agreed upon description of what you are measuring. It's generally best for several experts to generate items.

5) After generating items, it often makes sense to have experts rate the quality of the items, and to retain only the items with the highest ratings. The experts can also suggest revisions. If your items measure depression, the experts should be mental health professionals. If your items measure business skill, your experts should be business executives and managers.

6) Your test is then ready to be tested on a sample of people. Your sample should be a good cross-section of the people that you will want to compare test-takers to. After you administer your test to a sample of people:

-Determine the correlation between each item and the sum of the other items. If your test has subscales, do this separately for each subscale. Eliminate items that do not correlate well with the rest of their scale.

-Eliminate items that are too easy or too hard. If almost everyone agrees with an item or gets the correct answer, it is not a useful item.

-This procedure will maximize the test's internal consistency, one measure of reliability. You should calculate coefficient alpha. This statistic measures the degree to which a test scale measures a single construct, and the degree to which the test items are all measuring the same ability or trait. Alpha has a theoretical maximum of +1.00. A good test alpha is greater than .70.

7) The final test should be cross-validated on a new sample. During cross-validation, you can demonstrate test validity:

-You should be able to show that your test scores correlate with what they are supposed to correlate with.

For instance, a test of math skill should yield higher scores for students with higher math grades. A test of depression should yield higher scores for people who have been diagnosed with Major Depression.

-Factor analysis can be used to demonstrate that the test subscales group together (inter-correlate) in the way that theory would predict.

8) When the test is cross-validated, you can also calculate normative data. You can calculate the mean (average) score for test-takers, and calculate the standard deviation to determine how spread out the scores are around the mean. These statistics are extremely useful, because now any individual's score can be compared to the scores of people in general.

If your test has subscales, you will find the mean and standard deviation for each subscale. It is also often useful to find separate normative data for different groups of potential test takers. Many tests have norms according to gender, ethnic group, and age.

5.2: Questionnaire: (a) types, format and basic guidelines for constructing questions and questionnaire, (b) advantages and disadvantages of questionnaires

A **questionnaire** is a [research](#) instrument consisting of a series of [questions](#) and other prompts for the purpose of gathering information from respondents. Although they are often designed for [statistical](#) analysis of the responses, this is not always the case. The questionnaire was invented by [Sir Francis Galton](#).^[citation needed]

Questionnaires have advantages over some other types of [surveys](#) in that they are cheap, do not require as much effort from the questioner as verbal or telephone surveys, and often have standardized answers that make it simple to compile data. However, such standardized answers may frustrate users. Questionnaires are also sharply limited by the fact that respondents must be able to read the questions and respond to them. Thus, for some demographic groups conducting a survey by questionnaire may not be practical.

As a type of survey, questionnaires also have many of the same problems relating to question construction and wording that exist in other types of [opinion polls](#).

(A) Types, Format and basic guidelines for constructing questions and questionnaire:

Types of Questionnaire:

There are basically two types of questionnaire viz. Open and Closed.

1. Open – ended questionnaire:

These simply ask a question and leave an adequate space for an unstructured response. The replies may give important insights but they are very difficult to analyze in any systematic or statistical way. An example of Open ended questionnaire is give as under:

The Internal Revenue Service is currently soliciting information from individuals who have recently filed return(s) claiming that Internal Revenue Code §§ 3401 and 3121 allow for an exemption of income reported on Forms W-2 and 1099. Further, any withheld Social Security Tax and Medicare Tax may be claimed as a refund.

A review of your Federal income tax return(s) for taxable year(s) 2004 shows that you made such a claim on this (these) return(s).

We are requesting that you complete the attached questionnaire and return it with your response to the Letter 3176.

1. Did you advise your employer that the Form W-2 provided to you was incorrect? _____
2. Have you filed for exemption from withholding? _____
3. Where did you learn that IRC 3401, 3121 "and others" allows for an exemption of wages? _____
4. Did you engage anyone to assist you in preparing your returns? If so, did you pay for these services? _____
5. Have you ever filed a return for other tax periods that was deemed frivolous by the IRS? If yes, why did you change your position? _____
6. Have you contacted the IRS to provide copies of all payer documents reported? If yes, what was the purpose of this request? _____
7. Where did you get the document(s) attached to your return(s)? _____
8. Did you pay for these materials? If so, how much? _____
9. Have you purchased or otherwise obtained any manuals, videotapes, audiotapes, compact discs or other informational packages/materials to assist you in preparing your Federal income tax returns? If so, please list the materials you have purchased or obtained. _____
10. Have you attended any seminars or courses relating to the preparation of your Federal income tax returns? If so, briefly describe the seminar. _____

Sample of Open ended Questionnaire

Open ended questionnaire has following advantages:

1. This method is better to find opinions and general experiences of the respondents.
2. They are also useful for exploring the solutions for a problem or a process.
3. This type of questionnaire provide ample opportunity for the respondents to discuss the reasons for their ideas.
4. This type of questionnaire bring out the creativity of the user.

2. Closed Ended Questionnaire:

A Closed ended questionnaire is useful for elaborating factual information. This would demand that the analyst has a good understanding of the area under review so that simple effective questions can be formulated. Each question will have a clear purpose and would be posed in an unambiguous way, allowing the respondents to give a definite answer. An example of closed ended questionnaire is given below:

ABOUT RESOURCES - (Continued)		<Disagree Agree>				
26	The people in my team co-operate well with each other	1	2	3	4	5
27	My manager/team leader helps my team get the resources we need	1	2	3	4	5
28	We have sufficient skills in my team to work efficiently	1	2	3	4	5
29	Priorities or objectives are stable enough to let me get my work done	1	2	3	4	5
30	My factory/department makes good use of new technology	1	2	3	4	5

ABOUT PERSONAL ACHIEVEMENT & DEVELOPMENT - Please tell us here how well you think the company develops and uses the skills and knowledge of its employees						
31	I am satisfied with the opportunities I get to use my skills	1	2	3	4	5
32	I am given clear targets	1	2	3	4	5
33	I agree and accept the targets I am expected to achieve	1	2	3	4	5
34	I am trusted to achieve my targets on my own initiative	1	2	3	4	5
35	I am encouraged to take on responsibilities beyond my core job	1	2	3	4	5
36	I am satisfied with the sense of personal achievement my job provides	1	2	3	4	5
37	I am satisfied with the variety of tasks in my work	1	2	3	4	5
38	I am given the chance to learn new skills	1	2	3	4	5
39	I think my performance is fairly evaluated	1	2	3	4	5
40	I am encouraged to achieve more and develop my potential	1	2	3	4	5
41	I am given opportunities for personal growth, development and progression	1	2	3	4	5
42	I have been provided with the training I need to do my job well	1	2	3	4	5
43	My manager/team leader recognises good performance	1	2	3	4	5
44	I feel valued by my manager/team leader	1	2	3	4	5
45	The company promotes the most competent people	1	2	3	4	5

➔ Types of Questions in Questionnaire:

The following is a list of the different types of questions in questionnaire design:

1. Open Format Questions

Open format questions are those questions that give your audience an opportunity to express their opinions. In these types of questions, there are no predetermined set of responses and the person is free to answer however he/she chooses. By including open format questions in your questionnaire, you can get true, insightful and even unexpected suggestions. Qualitative questions fall under the category of open format questions. An ideal questionnaire would include an open format question at the end of the questionnaire that would ask the respondent about suggestions for changes or improvements.

Example of an Open Format Question

State your opinion about the quality of ABC Corporation's products and services?

.....

.....

2. Closed Format Questions

Closed format questions are questions that include multiple choice answers. Multiple choice questions fall under the category of closed format questions. These multiple choices could either be in even numbers or in odd numbers. By including closed format questions in your questionnaire design, you can easily calculate statistical data and percentages. Preliminary analysis can also be performed with ease. Closed format

questions can be asked to different groups at different intervals. This can enable you to efficiently track opinion over time.

Example of Closed Format Question

Which are the gadgets that you cannot live without?
-circle those applicable:

<input checked="" type="radio"/> Cell Phone	<input type="radio"/> Laptop	<input type="radio"/> PDA	<input checked="" type="radio"/> iPOD	<input type="radio"/> Digital Camera	<input type="radio"/> XBOX 360
---	------------------------------	---------------------------	---------------------------------------	--------------------------------------	--------------------------------

3. Leading Questions

Leading questions are questions that force your audience for a particular type of answer. In a leading question, all the answers would be equally likely. An example of a leading question would be a question that would have choices such as, fair, good, great, poor, superb, excellent etc. By asking a question and then giving answers such as these, you will be able to get an opinion from your audience.

Example of Leading Format Question

How would you rate the products of XYZ Corporation?

<input checked="" type="radio"/> 1 Fair	<input type="radio"/> 2 Good	<input type="radio"/> 3 Excellent	<input type="radio"/> 4 Superb
--	---------------------------------	--------------------------------------	-----------------------------------

4. Importance Questions

In importance questions, the respondents are usually asked to rate the importance of a particular issue, on a rating scale of 1-5. These questions can help you grasp what are the things that hold importance to your respondents. Importance questions can also help you make business critical decisions.

Example of Importance Format Question

Cost-effective services are:

<input type="radio"/> 1 Extremely Important	<input checked="" type="radio"/> 2 Very Important	<input type="radio"/> 3 Somewhat Important	<input type="radio"/> 4 Not Very Important	<input type="radio"/> 5 Not at all important
--	--	---	---	---

5. Likert Questions

Likert questions can help you ascertain how strongly your respondent agrees with a particular statement. Likert questions can also help you assess how your customers feel towards a certain issue, product or service.

Example of Likert Format Question

ABC Corporation's Products have to improve on quality?

1	2	3	4	5
Strongly Agree	Agree	Neither Agree or Disagree	Disagree	Strongly Disagree

6. Dichotomous Questions

Dichotomous questions are simple questions that ask respondents to just answer yes or no. One major drawback of a dichotomous question is that it cannot analyze any of the answers between yes and no.

Example of Dichotomous Format Question

Do you like the Products of ABC Corporation?

Yes No

7. Bipolar Questions

Bipolar questions are questions that have two extreme answers. The respondent is asked to mark his/her responses between the two opposite ends of the scale.

Example of Bipolar Format Question

How would you describe the services of XYZ Corporation?

Efficient	X	Inefficient
Fast	X	Slow
Reliable	X	Unreliable

8. Rating Scale Questions

In rating scale questions, the respondent is asked to rate a particular issue on a scale that ranges from poor to good. Rating scale questions usually have an even number of choices, so that respondents are not given the choice of an middle option.

Example of Rating Scale Format Question

How would you rate the services of ABC Corporation?

1 Good	2 Fair	3 Poor	4 Very Poor
-----------	-----------	-----------	----------------

9. Buying Propensity Questions

Buying propensity questions are questions that try to assess the future intentions of customers. These questions ask respondents if they want to buy a particular product, what requirements they want to be addressed and whether they would buy such a product in the future.

Example of Buying Propensity Format Question

If mobile phones had an in-built mp3 Player, would you prefer to buy it?

1 Definitely	2 Probably	3 Probably Not	4 Not Sure	4 Definitely Not
-----------------	---------------	-------------------	---------------	---------------------

Guidelines for creating better questionnaires

by Pamela Narins Manager, Market Research

Why is good question ordering and layout important? Because the cost of poorly designed questionnaires is the same as the cost of poorly written questions or a bad sample: the results you get can be biased, uninterpretable, insufficient or misleading.

This column concentrates on 15 guidelines for better survey design. As with much of survey research, these guidelines follow common sense rather than hard-and-fast principles. They are easy to forget, however, because they are so simple. But, keeping these guidelines in mind will greatly improve your chances of success.

Remember, responding to a survey should be an interesting, stress-free experience. If respondents become bored, confused or irritated, the results you get back may be of little use.

Think about the order and layout of the survey as if you had to fill it out. Is the survey attractive and professional looking? Is it easy to read? Are the questions interesting? Are there many questions on the same subject or are the topics varied?

There are some differences in guide-lines between self- administered surveys and interviewer-administered surveys. For now, simply keep the following differences in mind.

Self-administered surveys require respondents to be both interested in and capable of filling in the survey. In other words, the questionnaire not only has to motivate recipients in some way through its importance, appearance and questions, but be easy enough to complete to minimize fatigue, boredom or confusion.

Interviewer-administered surveys must be easy for the interviewer to read aloud, and must have instructions that reduce the chance of making a mistake. Additionally, questions must flow in such a way that the interest of the respondent, who cannot see the survey document, is kept up.

General guidelines for question ordering

1. The first several questions should:
 - be easy for respondents to understand
 - be important to the study's purpose
 - engage the attention and interest of your respondent

Do not begin with an open-ended question or one which respondents might feel has a "right" answer.

2. It is important to get your respondent interested in the survey at the very beginning. In cases where the topic is already of interest or importance to the respondent, **start with general questions, then funnel to more specific ones.** If the topic is of low importance to respondents, start with specific questions. This gives respondents a frame of reference; then ask broader, more general questions.
3. **Group questions in sections,** and position sections or questions in a logical order.
4. **Introduce new sections with a sentence or phrase** so that participants have a chance to switch mental gears.
5. **Place questions about sensitive issues** such as income, sexual habits or drug abuse **toward the end of the document, or section.** This helps avoid alienating, taxing or in other ways worrying participants.
6. **Consider lists of similar items carefully.** For example, you might ask about product satisfaction using 10 different attributes (price, availability, delivery, color, fit and so on). The first items on a list such as this usually get treated differently by respondents than the last items do. This is called position bias.

There are several ways to minimize position bias in a self-administered survey. One good way is to print several versions of the questionnaire, each with a different ordering of the items on the list. This way is often impractical, however, because of the high cost involved in printing several different versions of your questionnaire.

You can also randomly or alphabetically order list items and indicate in the instructions how they are ordered. This reduces the likelihood that respondents will see the first items as most important.

For interviewer-administered surveys, have your interviewers modify the order in which the attributes are presented to the respondent.

7. **Put demographic questions at the end of the questionnaire,** if possible. There are at least two reasons for this. First, some demographic questions such as age and income can be sensitive and should be placed at the end, as discussed in guideline number 5.

Second, it is better to **keep respondents' minds on the purpose of the survey at the beginning,** while you have their attention. Demographic questions rarely require much thought, so wait until the end when respondents might be tired.

8. **Try to minimize the number of times the respondent or the interviewer has to follow a skip rule.** For example "If the answer to question 3 is 'None of the Above,' skip to question 6." When there are too many skips or when skips become too complicated, you run the risk of introducing error and confusion.

General guidelines for questionnaire layout

1. **Create professional, attractive and uncluttered questionnaires;** fonts should be large enough to avoid eye strain; instructions for completing the survey should be easy to understand. If your survey has multiple pages, it should be bound in a booklet so that pages cannot get mixed up. And, if respondents are to see the survey, it should have an attractive cover.
2. **Make the survey easy to complete;** the check boxes or lines easy to see; and the numbers to be circled far enough apart so the respondent or interviewer will not inadvertently circle two numbers. If you are using scannable forms, where bubbles need to be filled in completely and with a specific writing utensil such as a number 2 pencil, make sure the instructions are clear and easy to find.
3. **Number your questions clearly.** This will lessen the chance, particularly in longer surveys, of respondents or interviewers getting lost.
4. **Start with a brief introduction describing the survey's purpose,** the topics being covered and how the results will be used. Also, mention any incentive for completing the survey, such as a drawing entry, the opportunity to have a copy of the results and so on.
5. If you are conducting a telephone survey, do NOT have the interviewer say "How are you?" in the introduction. This will irritate respondents who, at such an early stage in the interview, have been given no reason to want to talk to the interviewer, much less tell them how they are. **Keep the introduction short, polite and to the point.** Normally, the only question that should be asked in the introduction is if respondents are willing to participate.
6. For interviewer-administered surveys, make any interviewer instructions (that is, anything that is not supposed to be read to the respondent), easy to distinguish. **Put interviewer instructions in UPPER CASE, color or italics.**
7. **Leave plenty of room for respondents to write answers to open-ended questions.** Do not supply lines because this could constrain any comments.

(B) Advantages and Disadvantages of Questionnaire:

Advantages of Questionnaire:

1. Questionnaire are relatively cheap, particularly when there is a scattered group of users and operators.
2. Unlike the interview, which generally questions one person at a time, a questionnaire can be administered to larger numbers of individuals simultaneously.
3. Responses from questionnaire can be tabulated and analyzed more quickly.
4. The respondents feel greater confidence in anonymity of a questionnaire than in that of an interview. In an interview, the analyst usually knows the user staff by name, job function or other identification. With questionnaire, respondents give opinions without fear that the answer will not be connected to their names, if so desired.
5. The questionnaire places less pressure on users for immediate responses. Respondents have time to think about the questions, refer to documents and documentation and do calculations to provide more accurate data.

Disadvantages of questionnaire:

1. The number of respondents is often low.
2. All the questions given in the questionnaire are usually not answered completely, and there is no immediate opportunity to clarify a vague or incomplete answer to any question.

3. It is also not possible for the system analyst to observe and analyze the respondents body language.
4. Lack of direct contact may mean that questions are interpreted in different ways.
5. Good questionnaire are difficult to prepare.
6. Cost of developing and distributing questionnaire is also very high.

5.3: Interview: (a) types of interview such as individual & group and structured & unstructured, (b) construction of interview / interview schedule

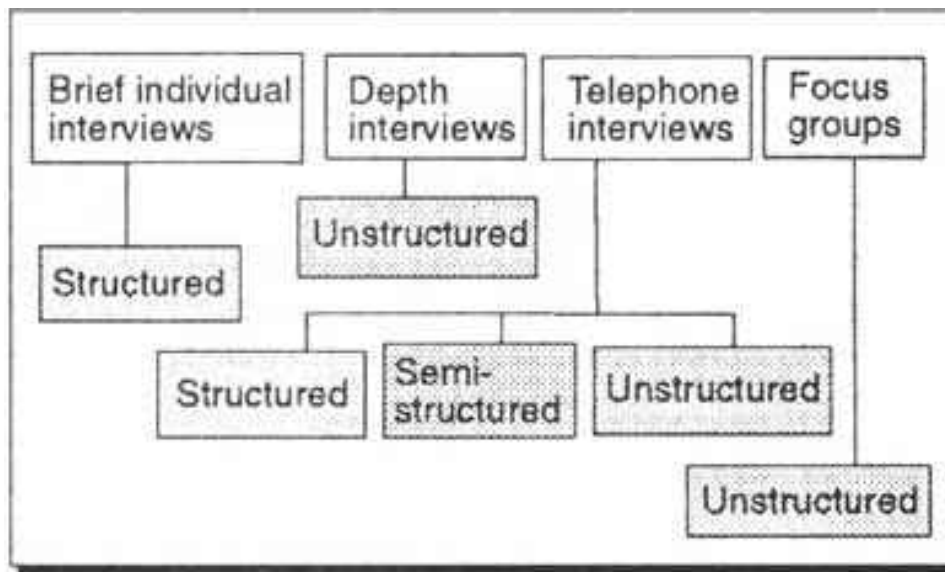
(A) Types of Interviews:

Analyst uses the interview to collect information from individuals or from groups. The respondents are the user of the investigated system. Some analyst always prefer the interview method compare to other fact-finding techniques, it is not always best that to collect the data about the application. Because if the number of the user is very high then this process is time consuming, so at that time other fact-finding technique should be adopted.

In this method respondents and system analyst converse during the interview. Interviews provide analyst with opportunity for gathering information from respondents who have been chosen for their knowledge of the system under study. This method is frequently the best source of qualitative information.

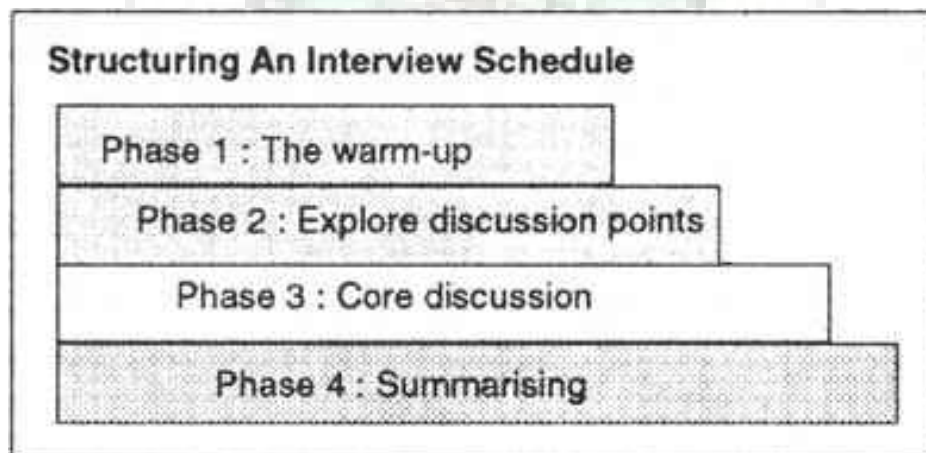
This method of fact finding can be especially helpful for gathering information from individuals who do not communicate effectively in writing or who may not have the time to complete questionnaires. It allows the analyst to discover area of misunderstanding, unrealistic expectations, and even indication of resistance to the proposed system. Interview can be structure or unstructured. Unstructured interviews use a questions and answers format when the analyst wants to get the general information about the system. Structured interviews use standardized questions in either an open response or in close response and the answer must be in the pre formatted.

Structured Interview	Unstructured Interview
Advantages: →Ensure uniform wording of questions for all respondents →Easy to administer and evaluate →Result in shorter interviews →More objective evaluation of both interviewer and respondent. →Limited interview training needed. Disadvantages: →Cost of preparation is high →Respondent has to give the answer into required and pre define format. →Analyst has to prepare before interviewing.	Advantages: →Interviewer has greater flexibility in wording question to suit respondent →Interviewers can pursue the areas those arise at the time of interview. →It may produce the information about the areas that were overlooked or not thought at all. Disadvantages: →It may waste the time of both While taking the interview, the biases may appear. →Some information will be gathered which is of no use. →It takes extra time to collect the actual facts.



(B) Construction of Interview / Interview Schedule

The interview schedule has at least four distinct sections: the warm-up, exploration of discussion points, the core discussion section and a summary.



The warm-up: This section has the objective of creating an atmosphere conducive to an open and free-flowing discussion. One technique that can be used to break down the initial bashfulness among group members who, in most instances, are strangers to one another is to divide them into pairs and exchange simple facts about themselves (e.g. their names, details of the families, place of work, interests etc.). Each group member is then asked to introduce their neighbour to the rest of the group.

The warm-up phase of the session then moves on to encourage the group members to engage in a free-ranging discussion around the topic upon which the discussion will eventually focus. For example, a municipal authority considering establishing a new fruit and vegetable wholesale market positioned outside a congested city centre would ultimately wish to determine what innovative facilities might attract traders to use the new market which is less convenient to them in terms of location. During the warm-up phase the moderator will direct the discussion in such a way as to obtain general information on how participants

currently behave with respect to the topic, issue or phenomenon under investigation. The emphasis is upon a **description** of current behaviour and attitudes. For instance, the traders would be asked to describe their own modes of operation within the wholesale market as well as those of fellow traders.

Exploration of discussion points: In this phase the discussion moves on to the participants' attitudes, opinions and experiences of existing products, services (or in this case facilities) and on to what they like and dislike about those products/services. With reference to the wholesale markets example, at this stage traders would be invited to comment on the advantages and disadvantages of the facilities within which they currently operate.

Core discussion: This part of the group discussion focuses directly upon the principal purpose of the research. The flow of the session moves on to the participants' perceptions of new concepts, possible developments or innovations. The wholesale traders, for instance, would be guided towards discussing peri-urban wholesale markets and the kinds of facilities which might attract traders like themselves. A common approach is to follow a sequence of first exploring the ideas which participants generate themselves and then to solicit participants' reactions to ideas preconceived by researchers, or their clients, about possible future developments.

Summary: The final phase of the focus groups session allows participants to reflect upon the foregoing discussion and to add any views or information on the topic that they may have previously forgotten or otherwise have omitted. A common tactic is to conclude the session by inviting the group, as well as its individual members, to "advise the manufacturer" (or whoever) on the issue at hand.

5.4: Rating Scales: Types, format and basic guidelines for constructing scale

A **rating scale** is a set of categories designed to elicit information about a quantitative or a qualitative attribute. In the social sciences, common examples are the Likert scale and 1-10 rating scales in which a person selects the number which is considered to reflect the perceived quality of a product.

A rating scale is an instrument that requires the rater to assign the rated object that have numerals assigned to them.

Types of Rating Scales

All the rating scales can be classified into one of the following four classifications:-

1. Some data are measured at the **ordinal level**. Numbers indicate the relative position of items, but not the magnitude of difference. One example is a Likert scale:

Statement: I could not live without my computer.

Response options:

1. Strongly disagree
 2. Disagree
 3. Agree
 4. Strongly agree
2. Some data are measured at the **interval level**. Numbers indicate the magnitude of difference between items, but there is no absolute zero point. Examples are attitude scales and opinion scales.
 3. Some data are measured at the **ratio level**. Numbers indicate magnitude of difference and there is a fixed zero point. Ratios can be calculated. Examples include age, income, price, costs, sales revenue, sales volume and market share.

More than one rating scale is required to measure an attitude or perception due to the requirement for statistical comparisons between the categories in the **polytomous Rasch model** for ordered categories (Andrich, 1978). In terms of Classical test theory, more than one question is required to obtain an index of internal reliability such as **Cronbach's alpha** (Cronbach, 1951), which is a basic criterion for assessing the effectiveness of a rating scale and, more generally, a psychometric instrument.

Scaling is the branch of measurement that involves the construction of an instrument that associates qualitative constructs with quantitative metric units. Scaling evolved out of efforts in psychology and education to measure "unmeasurable" constructs like authoritarianism and self esteem. In many ways, scaling remains one of the most arcane and misunderstood aspects of social research measurement. And, it attempts to do one of the most difficult of research tasks -- measure abstract concepts.

Most people don't even understand what scaling is. The basic idea of scaling is described in **General Issues in Scaling**, including the important distinction between a scale and a response format. Scales are generally divided into two broad categories: unidimensional and multidimensional. The unidimensional scaling methods were developed in the first half of the twentieth century and are generally named after their inventor. We'll look at three types of unidimensional scaling methods here:

- Thurstone or Equal-Appearing Interval Scaling
- Likert or "Summative" Scaling
- Guttman or "Cumulative" Scaling

In the late 1950s and early 1960s, measurement theorists developed more advanced techniques for creating multidimensional scales. Although these techniques are not considered here, you may want to look at the method of [concept mapping](#) that relies on that approach to see the power of these multivariate methods.

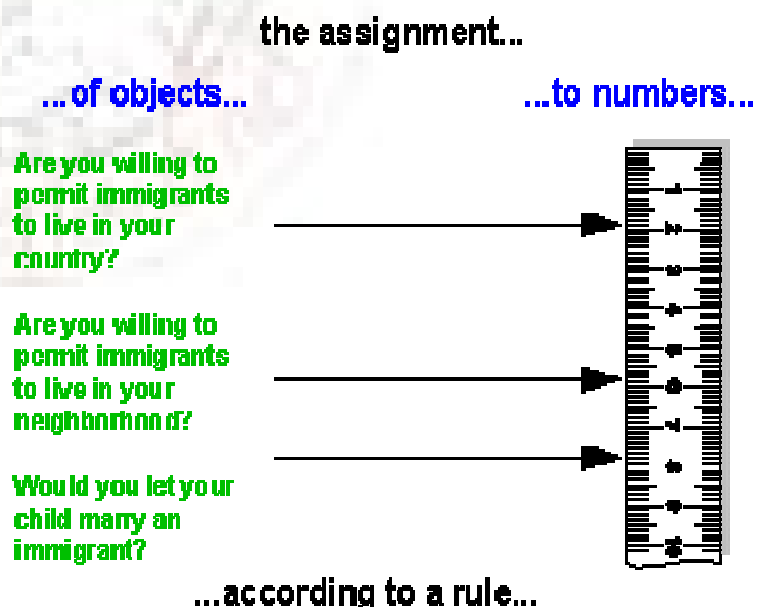
General Issues in Scaling

S.S. Stevens came up with what I think is the simplest and most straightforward definition of scaling. He said:

Scaling is the assignment of objects to numbers according to a rule.

But what does that mean? In most scaling, the objects are text statements, usually statements of attitude or belief. The figure shows an example.

There are three statements describing attitudes towards immigration. To scale these statements, we have to assign numbers to them. Usually, we would like the result to be on at least an interval scale as indicated by the ruler in the figure. And what does "according to a rule" mean? If you look at the statements, you can see that as you read down, the attitude towards immigration becomes more restrictive -- if a person agrees with a statement on the list, it's likely that they will also agree with all of the statements higher on the list. In this case, the "rule" is a *cumulative* one. So what is scaling? It's how we get numbers that can be meaningfully assigned to objects -- it's a set of procedures. We'll present several different approaches below.



But first, I have to clear up one of my pet peeves. People often confuse the idea of a scale and a response scale. A response scale is the way you collect responses from people on an instrument. You might use a dichotomous response scale like Agree/Disagree, True/False, or Yes/No. Or, you might use an interval response scale like a 1-to-5 or 1-to-7 rating. But, if all you are doing is attaching a response scale to an object or statement, you can't call that scaling. As you will see, scaling involves procedures that you do independent of the respondent so that you can come up with a numerical value for the object. In true scaling research, you use a scaling procedure to develop your instrument (scale) and you also use a response scale to collect the responses from participants. But just assigning a 1-to-5 response scale for an item is **not** scaling! The differences are illustrated in the table below.

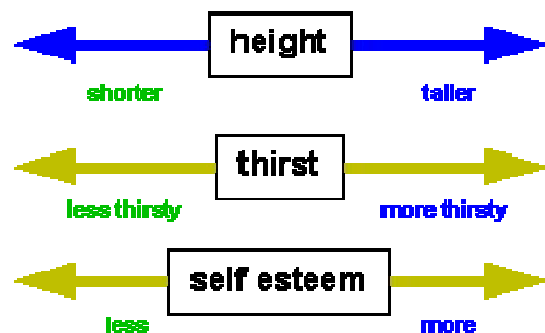
Scale	Response Scale
results from a process	is used to collect the response for an item
each item on scale has a scale value	item not associated with a scale value
refers to a set of items	used for a single item

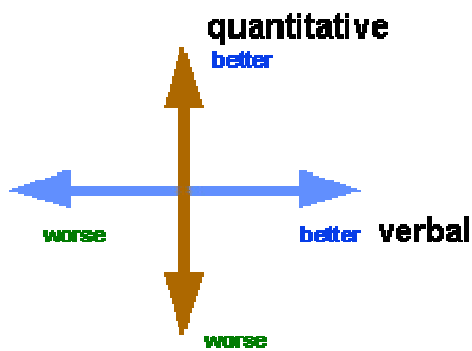
Purposes of Scaling

Why do we do scaling? Why not just create text statements or questions and use response formats to collect the answers? First, sometimes we do scaling to test a hypothesis. We might want to know whether the construct or concept is a single dimensional or multidimensional one (more about dimensionality later). Sometimes, we do scaling as part of exploratory research. We want to know what dimensions underlie a set of ratings. For instance, if you create a set of questions, you can use scaling to determine how well they "hang together" and whether they measure one concept or multiple concepts. But probably the most common reason for doing scaling is for scoring purposes. When a participant gives their responses to a set of items, we often would like to assign a single number that represents that person's overall attitude or belief. For the figure above, we would like to be able to give a single number that describes a person's attitudes towards immigration, for example.

Dimensionality

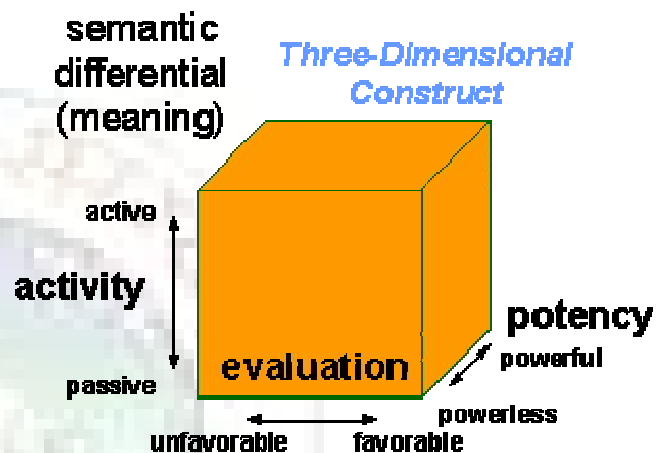
A scale can have any number of dimensions in it. Most scales that we develop have only a few dimensions. What's a dimension? Think of a dimension as a number line. If we want to measure a construct, we have to decide whether the construct can be measured well with one number line or whether it may need more. For instance, height is a concept that is unidimensional or one-dimensional. We can measure the concept of height very well with only a single number line (e.g., a ruler). Weight is also unidimensional -- we can measure it with a scale. Thirst might also be considered a unidimensional concept -- you are either more or less thirsty at any given time. It's easy to see that height and weight are unidimensional. But what about a concept like self esteem? If you think you can measure a person's self esteem well with a single ruler that goes from low to high, then you probably have a unidimensional construct.





What would a two-dimensional concept be? Many models of intelligence or achievement postulate two major dimensions -- mathematical and verbal ability. In this type of two-dimensional model, a person can be said to possess two types of achievement. Some people will be high in verbal skills and lower in math. For others, it will be the reverse. But, if a concept is truly two-dimensional, it is not possible to depict a person's level on it using only a single number line. In other words, in order to describe achievement you would need to locate a person as a point in two dimensional (x,y) space.

OK, let's push this one step further: how about a three-dimensional concept? Psychologists who study the idea of meaning theorized that the meaning of a term could be well described in three dimensions. Put in other terms, any objects can be distinguished or differentiated from each other along three dimensions. They labeled these three dimensions *activity*, *evaluation*, and *potency*. They called this general theory of meaning the **semantic differential**. Their theory essentially states that you can rate any object along those three dimensions. For instance, think of the idea of "ballet." If you like the ballet, you would probably rate it high on activity, favorable on evaluation, and powerful on potency. On the other hand, think about the concept of a "book" like a novel. You might rate it low on activity (it's passive), favorable on evaluation (assuming you like it), and about average on potency. Now, think of the idea of "going to the dentist." Most people would rate it low on activity (it's a passive activity), unfavorable on evaluation, and powerless on potency (there are few routine activities that make you feel as powerless!). The theorists who came up with the idea of the semantic differential thought that the meaning of any concepts could be described well by rating the concept on these three dimensions. In other words, in order to describe the meaning of an object you have to locate it as a dot somewhere within the cube (three-dimensional space).



Unidimensional or Multidimensional?

What are the advantages of using a unidimensional model? Unidimensional concepts are generally easier to understand. You have either more or less of it, and that's all. You're either taller or shorter, heavier or lighter. It's also important to understand what a unidimensional scale is as a foundation for comprehending the more complex multidimensional concepts. But the best reason to use unidimensional scaling is because you believe the concept you are measuring really is unidimensional in reality. As you've seen, many familiar concepts (height, weight, temperature) are actually unidimensional. But, if the concept you are studying is in fact multidimensional in nature, a unidimensional scale or number line won't describe it well. If you try to measure academic achievement on a single dimension, you would place every person on a single line ranging from low to high achievers. But how do you score someone who is a high math achiever and terrible verbally, or vice versa? A unidimensional scale can't capture that type of achievement.

The Major Unidimensional Scale Types

There are three major types of unidimensional scaling methods. They are similar in that they each measure the concept of interest on a number line. But they differ considerably in how they arrive at scale values for different items. The three methods are Thurstone or Equal-Appearing Interval Scaling, Likert or "Summative" Scaling, and Guttman or "Cumulative" Scaling.

5.5: Measurement of Attitude: Thurstone and Likert technique

ATTITUDE MEASUREMENT

THURSTONE SCALE:

In psychology, the **Thurstone scale** was the first formal technique for measuring an attitude. It was developed by Louis Leon Thurstone in 1928, as a means of measuring attitudes towards religion. It is made up of statements about a particular issue, and each statement has a numerical value indicating how favorable or unfavorable it is judged to be. People check each of the statements to which they agree, and a mean score is computed, indicating their attitude.

The Thurstone attitude measurement procedure is generally more consistent with empirical characteristics of disagree-agree responses. It is constructed by the method of equal-appearing intervals, in which a large pool of candidate statements about an attitude object, ranging from strongly negative (*Abortion is never justified*) through neutral (*There are arguments both for and against abortion*) to strongly positive (*Abortion is every woman's right*), are sorted by a group of judges into eleven categories. They are assumed to appear equally spaced on the attitude continuum, according to how favorable the statements are towards the attitude object. Items that yield the highest level of agreement among the judges as to their scale position, and that collectively represent an adequate range of contents and scale positions, are then selected for the final scale. Respondents to the scale endorse just those items with which they agree, and an individual respondent's score is calculated as the mean (or occasionally median) of the items endorsed, such scores being assumed to lie on an interval scale of measurement. It is also called an equal-appearing interval scale.

Thurstone was one of the first and most productive scaling theorists. He actually invented three different methods for developing a unidimensional scale: the **method of equal-appearing intervals**; the **method of successive intervals**; and, the **method of paired comparisons**. The three methods differed in how the scale values for items were constructed, but in all three cases, the resulting scale was rated the same way by respondents.

GUTTMAN SCALE

A type of attitude scale which is named after the US (later Israeli) psychologist Louis H. Guttman (1916–87). The items of which can be arranged in a hierarchical order such that agreement with any particular item implies probable agreement with all those below it in the hierarchy, as would apply to the following (non-attitudinal) items: *I am over 5 feet tall*; *I am over 5 feet 6 inches tall*; *I am over 6 feet tall*; *I am over 6 feet 6 inches tall*. It is constructed by the method of scalogram analysis in which a large pool of candidate statements about an attitude object (such as *Euthanasia is morally wrong*; *Euthanasia should be legalized*, and so on) are administered to a group of respondents who mark just those items with which they agree, and from these responses a set of items is selected that can be arranged into a hierarchy with as few errors (deviations from a perfect linear hierarchy) as possible, a satisfactory Guttman scale having relatively few errors as indexed by a reproducibility (Rep) of at least 0.90, implying that the scale is unidimensional, measuring only one major attitude variable. The process of constructing a Guttman scale is called *scalogram analysis*. A simple example of a Guttman scale is a social distance scale

In statistical surveys conducted by means of structured interviews or questionnaires, a subset of the survey items having binary (e.g., YES or NO) answers forms a **Guttman scale** if they can be ranked in some order so that, for a rational respondent, the response pattern can be captured by a single index on that ordered scale. In other words, on a Guttman scale, items are arranged in an order so that an individual who agrees with a particular item also agrees with items of lower rank-order. For example, a series of items could be

(1) "I am willing to be near ice cream"; (2) "I am willing to smell ice cream"; (3) "I am willing to eat ice cream"; and (4) "I love to eat ice cream". Agreement with any one item implies agreement with the lower-order items.

The concept of Guttman scale likewise applies to series of items in other kinds of tests, such as achievement tests, that have binary outcomes. For example, a test of math achievement might order questions based on their difficulty and instruct the examinee to begin in the middle. The assumption is if the examinee can successfully answer items of that difficulty (e.g., summing two 3-digit numbers), s/he would be able to answer the earlier questions (e.g., summing two 2-digit numbers). Some achievement tests are organized in a Guttman scale to reduce the duration of the test.

By designing surveys and tests such that they contain Guttman scales, researchers can simplify the analysis of the outcome of surveys, and increase the robustness. Guttman scales also make it possible to detect and discard randomized answer patterns, as may be given by uncooperative respondents. The Guttman scale is used mostly when researchers want to design short questionnaires with good discriminating ability. The Guttman model works best for constructs that are hierarchical and highly structured such as social distance, organizational hierarchies, and evolutionary stages.

Guttman scaling is also sometimes known as **cumulative scaling** or **scalogram analysis**. The purpose of Guttman scaling is to establish a one-dimensional continuum for a concept you wish to measure. What does that mean? Essentially, we would like a set of items or statements so that a respondent who agrees with any specific question in the list will also agree with all previous questions. Put more formally, we would like to be able to predict item responses perfectly knowing only the total score for the respondent. For example, imagine a ten-item cumulative scale. If the respondent scores a four, it should mean that he/she agreed with the first four statements. If the respondent scores an eight, it should mean they agreed with the first eight. The object is to find a set of items that perfectly matches this pattern. In practice, we would seldom expect to find this cumulative pattern perfectly. So, we use scalogram analysis to examine how closely a set of items corresponds with this idea of cumulateness.

LIKERT SCALE

The Likert scale was introduced as a scale of attitudes in Likert's "A Technique for the Measurement of Attitudes," It was a bipolar scale running from one extreme through a neutral point to the opposite extreme. The Likert technique presents a set of attitude statements. Subjects are asked to express agreement or disagreement of a five-point scale. Each degree of agreement is given a numerical value from one to five. Thus a total numerical value can be calculated from all the responses.

A **Likert scale** is a psychometric scale commonly used in questionnaires, and is the most widely used scale in survey research. When responding to a Likert questionnaire item, respondents specify their level of agreement to a statement. The scale is named after Rensis Likert, who published a report describing its use.

An important distinction must be made between a *Likert Scale* and a *Likert item*. The Likert Scale is the sum of responses on several Likert items. Because Likert items are often accompanied by a visual analog scale (e.g., a horizontal line, on which a subject indicates his or her response by circling or checking tick-marks), the items are sometimes called scales themselves. This is the source of much confusion; it is better, therefore, to reserve the word 'Likert scale' to apply to the summated scale, and 'Likert item' to refer to an individual item.

A Likert item is simply a statement which the respondent is asked to evaluate according to any kind of subjective or objective criteria; generally the level of agreement or disagreement is measured. Often five ordered response levels are used, although many psychometricians advocate using seven or nine levels; a recent empirical study^[3] found that data from 5-level, 7-level and 10-level items showed very similar characteristics in terms of mean, variance, skewness and kurtosis after a simple transformation was applied.

The format of a typical five-level Likert item is:

1. Strongly disagree
2. Disagree
3. Neither agree nor disagree
4. Agree
5. Strongly agree

Likert scaling is a bipolar scaling method, measuring either positive or negative response to a statement. Sometimes a four-point scale is used; this is a forced choice method since the middle option of "Neither agree nor disagree" is not available.

Likert scales may be subject to distortion from several causes. Respondents may avoid using extreme response categories (*central tendency bias*); agree with statements as presented (*acquiescence bias*); or try to portray themselves or their organization in a more favorable light (*social desirability bias*). Designing a scale with balanced keying (an equal number of positive and negative statements) can obviate the problem of acquiescence bias, since acquiescence on positively keyed items will balance acquiescence on negatively keyed items, but central tendency and social desirability are somewhat more problematic.

Scoring and analysis

After the questionnaire is completed, each item may be analyzed separately or in some cases item responses may be summed to create a score for a group of items. Hence, Likert scales are often called summative scales.

Whether individual Likert items can be considered as interval-level data, or whether they should be considered merely ordered-categorical data is the subject of disagreement. Many regard such items only as ordinal data, because, especially when using only five levels, one cannot assume that respondents perceive all pairs of adjacent levels as equidistant. On the other hand, often (as in the example above) the wording of response levels clearly implies a *symmetry* of response levels about a middle category; at the very least, such an item would fall between ordinal- and interval-level measurement; to treat it as merely ordinal would lose information. Further, if the item is accompanied by a visual analog scale, where equal spacing of response levels is clearly indicated, the argument for treating it as interval-level data is even stronger.

When treated as ordinal data, Likert responses can be collated into bar charts, central tendency summarised by the median or the mode (but not the mean), dispersion summarised by the range across quartiles (but not the standard deviation), or analyzed using non-parametric tests, e.g. Chi-square test, Mann-Whitney test, Wilcoxon signed-rank test, or Kruskal-Wallis test.^[4]

Responses to several Likert questions may be summed, providing that all questions use the same Likert scale and that the scale is a defensible approximation to an interval scale, in which case they may be treated as interval data measuring a latent variable. If the summed responses fulfil these assumptions, parametric statistical tests such as the analysis of variance can be applied. These can be applied only when the components are more than 5.[Clarify: More than 5 Likert questions or Likert questions of more than 5 levels]

Data from Likert scales are sometimes reduced to the nominal level by combining all agree and disagree responses into two categories of "accept" and "reject".

5.6: Concept of some other tools: format and use of (a) Checklist, (b) Opinionnaire (c) Observation Schedule

(A) Checklist:

A **checklist** is a type of informational job aid used to reduce failure by compensating for potential limits of human memory and attention. It helps to ensure consistency and completeness in carrying out a task. A basic example is the "to do list." A more advanced checklist would be a schedule, which lays out tasks to be done according to time of day or other factors.

Applications

- Aviation checklists aid in aviation safety to ensure that critical items are not forgotten
- use in medical practice to ensure that clinical practice guidelines are followed. An example is the Surgical Safety Checklist developed for the World Health Organization by Dr. Atul Gawande.

- used in quality assurance of software engineering, to check process compliance, code standardization and error prevention, and others.
- often used in industry in operations procedures.
- used in civil litigation to deal with the complexity of discovery and motions practice. An example is the open-source litigation checklist.
- can aid in mitigating claims of negligence in public liability claims by providing evidence of a risk management system being in place.
- an ornithological checklist, a list of birds with standardized names that helps ornithologists communicate with the public without the use of scientific names in Latin.
- a popular tool for tracking sports card collections. Randomly inserted in packs, checklist cards provide information on the contents of sports card set.

Format

Checklists are often presented as lists with small checkboxes down the left hand side of the page. A small tick or checkmark is drawn in the box after the item has been completed.

Other formats are also sometime used. Aviation checklists generally consist of a system and an action divided by a dashed line, and lack a checkbox as they are often read aloud and are usually intended to be reused.

B) Opinionnaire:

A form containing a list of statements, each of which the members of a selected group are asked to endorse or reject; the purpose being to gather information for a survey

Opinionnaire = Opinion + Questionnaire

A questionnaire designed to elicit views on matters of opinion from which generalizations may be abstracted.

Example:

Fears and Phobias Opinionnaire

Read each of the following statements. Write A on your paper if you agree with a statement or D if you disagree with a statement. We will discuss these in class. We will revisit these statements after we complete the unit to see if you have changed your opinions. **(Get the questionnaire from your classroom teacher).**

1. People basically enjoy being afraid.
2. Being afraid is therapeutic and cathartic.
3. Halloween is one of my favorite holidays.
4. A little fear goes a long way.
5. The bloodier and gorier a movie is, the more I like it.
6. People with stage fright are faking it trying to get more attention.
7. People with vivid imaginations should not read scary books or go to scary movies.
8. The only thing we have to fear is fear itself.
9. Fear is construed as the flipside of pleasure in many instances.
10. I believe in haunted houses and in ghosts.
11. More people are afraid of critters and snakes than of the unknown.
12. The thought of UFOs frightens me.
13. Fatal attractions and stalkers are more frightening than snakes and critters.

14. Being stressed by fears is essential for human survival.
15. One's imagination turns normal fears into terror or into phobias.
16. Phobias can incapacitate a person.
17. "The Blair Witch Project" was not a scary movie.
18. Always fear your English teacher.
19. There will never be any new ideas for scary movies, only "revisiting" the old ideas in new ways.
20. I think horror movies are funny and entertaining.
21. I think reading a scary book is more frightening than seeing a scary movie.
22. Nothing frightens me.
23. Scary movies cause nightmares.
24. Stephen King is my favorite author.
25. I loved to read R. L. Stine when I was small.

(C) Observation Schedule:

Observation schedules are one of many essential analytical devices that scientists can use to turn multifaceted and complex visual observations into usable research data.

1. Identification

- An observation schedule is an analytical form, or coding sheet, filled out by researchers during structured observation. It carefully specifies beforehand the categories of behaviors or events under scrutiny and under what circumstances they should be assigned to those categories. Observations are then fragmented, or coded, into these more manageable pieces of information, which are later aggregated into usable, quantifiable data.

Types

- Observation schedules are utilized primarily in the fields of education, psychology, speech and language therapy, learning and behavioral therapy and market research. Schedules can range from exceedingly complex multiple-page examinations to simple tally sheets. Types of observation schedules include event sampling, time sampling, interval recording, rating scales and duration recording.

Examples

- One of the most widely known and sophisticated observation schedules is the Autism Diagnostic Observation Schedule (ADOS), which systematically tests for telltale signs of autism in its subjects. Other notable examples include the Modified-Classroom Observation Schedule to Measure Intentional Communication (M-COSMIC) and the Flanders Interaction Analysis Categories.

Autism Diagnostic Observation Schedule (ADOS)

The **Autism Diagnostic Observation Schedule** is an instrument for diagnosing and assessing Autism. It was created by Catherine Lord, Ph.D., Michael Rutter, M.D., FRS, Pamela C. DiLavore, Ph.D., and Susan Risi, Ph.D. in 1989 and became commercially available in 2001 through the WPS, or Western Psychological Services. The protocol consists of a series of structured and semi-structured tasks that involve social interaction between the examiner and the subject. The examiner observes and identifies segments of the subject's behavior and assigns these to predetermined observational categories. Categorized observations are subsequently combined to produce quantitative scores for analysis. Research-determined cut-offs identify the potential diagnosis of autism or related autism spectrum disorders, allowing a standardized assessment of autistic symptoms. The Autism Diagnostic Interview-Revised (ADI-R), a companion instrument, is a structured interview conducted with the parents of the referred individual and covers the

subject's full developmental history. The ADOS should not be used for formal diagnosis with individuals who are blind, deaf, or otherwise seriously impaired by sensory or motor disorders, such as cerebral palsy or muscular dystrophy.

Method

The ADOS generally takes from 30 to 60 minutes to administer. During this time the examiner provides a series of opportunities for the subject to show social and communication behaviors relevant to the diagnosis of autism. Each subject is administered activities from just one of the four modules. The selection of an appropriate module is based on the developmental and language level of the referred individual. The only developmental level not served by the ADOS is that for adolescents and adults who are nonverbal.^[1] A revision, the ADOS-2, is currently in development with a release goal date in early 2011. It will include improved algorithms for Modules 1 to 3 and a new Toddler Module that facilitates assessment in children ages 12 to 20 months.

Modules

Module 1 is used with children who use little or no phrase speech. Subjects that do use phrase speech but do not speak fluently are administered Module 2. Since these modules both require the subject to move around the room, the ability to walk is generally taken as a minimum developmental requirement to use of the instrument as a whole. Module 3 is for younger subjects who are verbally fluent and Module 4 is used with adolescents and adults who are verbally fluent. Some examples of Modules 1 or 2 include response to name, social smile, and free or bubble play. Modules 3 or 4 can include reciprocal play and communication, exhibition of empathy or comments on others' emotions.

5.7: Standardization of research tool: Reliability, Validity and Norms *(refer chapter 2 of Elective Subject: Psychological Testing)*