



Research Methods & Technical Writing

Lesson 5 - Week 5

Measurement and Scaling Techniques

Lecturer: Dr. Msagha J Mbogholi, PhD

Flashback from Lesson 4

- Advantages of sampling include saving time, saving costs and accuracy.
- A sample design is a definite plan for obtaining a sample from a given population. It refers to the technique or the procedure the researcher would adopt in selecting items for the sample.
- The two causes of wrong interpretation are systematic bias and sampling error.
- There are different types of sample designs based on two factors viz., the representation basis and the element selection technique. On the representation basis, the sample may be probability sampling or it may be non-probability sampling. On element selection basis, the sample may be either unrestricted or restricted. When each sample element is drawn individually from the population at large, then the sample so drawn is known as 'unrestricted sample', whereas all other forms of sampling are covered under the term 'restricted sampling'.
- Non-probability sampling designs include convenience sampling, purposive sampling, quota sampling, and referral /snowball sampling
- Probability sampling designs include simple random sampling, stratified sampling, systematic sampling, cluster sampling and multi stage sampling

Content

- Introduction
- Measurement scales
- Sources of error in measurement
- Tests of sound measurement
- Technique of developing measurement tools
- Scaling



Part 1

Introduction

Introduction

- We use the term measurement in our daily lives. What does the term measurement actually mean? We even use it in daily life as a form of comparison in language; “does pizza inn measure up to pizza hut?” (I’d love to know the answer to that one).
- The Mirriam-webster dictionary defines a measure as “something (as a yardstick or cup) used in measuring”; the paradox is you can’t define ‘measure’, ‘measuring’, ‘measurement’ without using one of the other terms in the same sentence. Interesting isn’t it?
- What we do know that there must be some kind of instrument involved when we wish to measure something, and that instrument has been calibrated to some standard. Think about weight, height, and other physical concepts; these are easy to measure since we talk of kgs, cms, meters, pounds, ounces, and so on. How about abstract concepts, how do we measure these? Let’s start with a controversial concept like intelligence, how do we measure a person’s intelligence? Just try and do one of the free online intelligence tests and share your opinion; is that really your level of intelligence?

Introduction (cont'd)

- How do we measure the ability to stand stress, or the ability to withstand pain, for that matter? Oh, he has 20 units of stress resistance? Or he can withstand 40 units of pain?
- “Technically speaking, measurement is a process of mapping aspects of a domain onto other aspects of a range according to some rule of correspondence. In measuring, we devise some form of scale in the range (in terms of set theory, range may refer to some set) and then transform or map the properties of objects from the domain (in terms of set theory, domain may refer to some other set) onto this scale.” (Kothari, 2004)
- For example let us say we have a postgraduate workshop in the university. We wish to find out the ratio of masters students vs PhD students who attend the workshop for purposes of future planning. In terms of set theory, this process is one of mapping the observed physical properties of those attending the workshop (the domain) on to a class (Masters or postgraduate) classification (the range). The rule of correspondence is: If the object in the domain appears to be a Masters student, assign to “0” and if PhD student assign to “1”. The same can be done for relationship status; a person’s relationship status can be “single”, “married”, “divorced”, “widowed”, “it’s complicated”, “rather not say”, and so on; each of these can be assigned a value; for example, single (1), married (2),.....rather not say (6). Of course it is only natural to assign values from 1 since our minds have been trained to start from the beginning of the decimal numbering system (consciously or subconsciously) but we can also start assigning values from 10 or 52 or 100, since these values are subject to the researcher and for example single isn’t really equal to any number, it’s just an assignment.

Introduction (cont'd)

- The question is why do we even need to assign values to the data? It is because numerical data is what is used in research to give the data some meaning; that is, for analysis purposes. In a later lesson we shall examine these analysis techniques (tests and so on) that give our data meaning.
- Further it is important to understand that this type of data is abstract and not physical. For example in the case of physical measurement we can say that if Ted is 95 kg and Bill is 50 kg, then their difference in weight is $(95 - 50) = 45$ kgs. What of abstract data, can we apply the same formula? If we say single = 1, and "rather not say" = 6, can we say that "rather not say"(6) – "single"(1) = "it's complicated"(5)? What would that even mean in terms of interpretation? Thus it means that this type of data needs different way of analyzing (like I said we shall cover this in a later lesson).
- Clearly then, this means that in research we have to first understand the type of data we are dealing with, then determine a scale for it (either by creating one or using an existing acceptable one) and finally use the scale as a measure, before using it as a research instrument.
- This lesson covers these aspects; we begin by explaining the different types of data (no need to reinvent the wheel, we shall use existing literature for this), then discuss measurements and scaling.



Part 2

Measurement scales

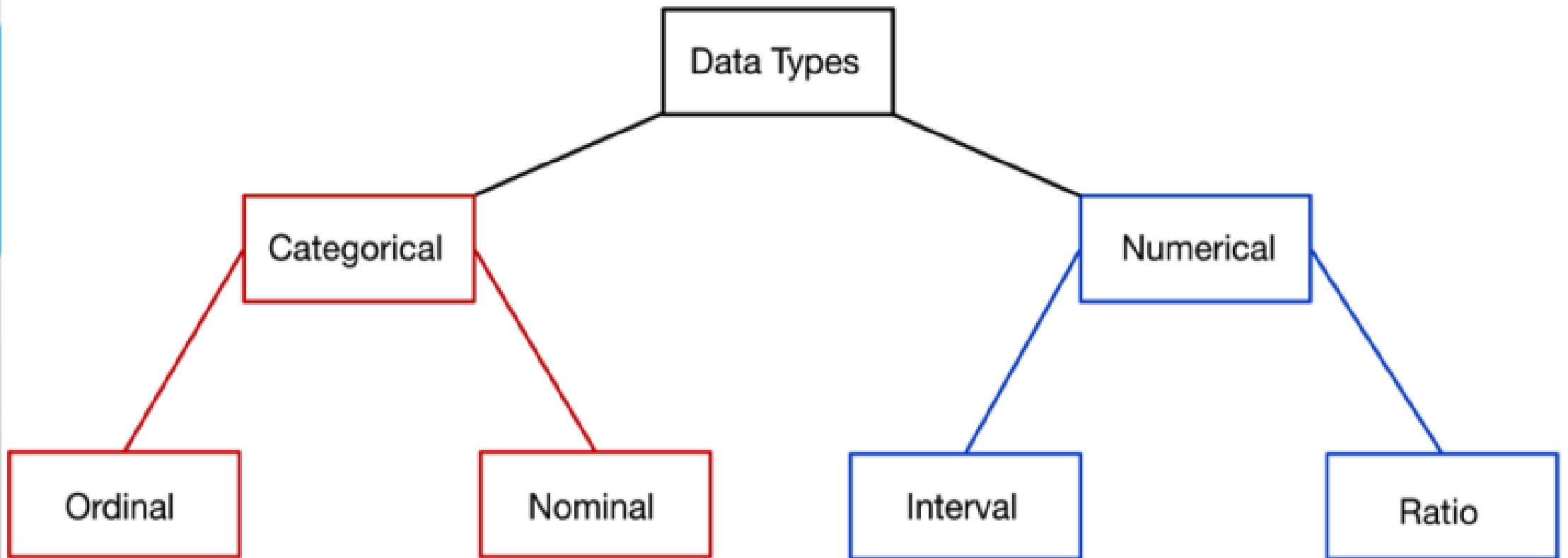


Fig 1 Data types in statistics. (Donges, n.d)

2.1 Categorical data types

- Fig 1 shows the classification of data used in statistics according to types. Data can be classified as being either categorical or numerical. Categorical data can be further classified as being either ordinal or nominal. Numerical data can further be classified as being either interval or ratio.
- The different types of variables used in research are categorized as belonging to one of the above groups. Further the different data types are also referred to as **measurement scales** in literature (Donges, n.d).
- Categorical (qualitative) data: Categorical data represents characteristics. Therefore it can represent things like a person's gender, language, etc. Categorical data can also take on numerical values (Example: 1 for female and 0 for male). Note that those numbers don't have mathematical meaning. (Donges, n.d)
- Nominal data: Nominal values represent discrete units and are used to label variables that have no quantitative value. Just think of them as "labels." Note that nominal data that has no order. Therefore, if you would change the order of its values, the meaning would not change. Fig 2 shows an example of ordinal data. The only logical operation that you can apply to them is equality or inequality which you can also use to group them. The descriptive statistics you can do with nominal data include frequencies, proportions, percentages, and central points. And, to visualize nominal data, you can use a pie chart or a bar chart. (*Types in Statistics - Qualitative and Quantitative Data Types*, n.d.).

2.1 Categorical data types (cont'd)

- A further scenario where nominal data is used in a research setting is as follows (*7 Types of Data in Business (with Examples and Uses)*, 2022):
- “A research organization has gathered polling data for an upcoming election. The researchers categorize the voters by age group. During polling, they also held interviews with passersby to gather voters' rationale for their voting preferences. With the data they've collected, the researchers can draw conclusions not only about the correlation between age and voting preference but also about the motivations that drive each age group. ”
- Ordinal data: Ordinal values represent discrete as well as ordered units. Unlike nominal, here the ordering matters. However, there is no consistency in the relative distance between the adjacent categories. And, similar to nominal data, ordinal data also don't have a meaningful zero. The descriptive statistics that you can do with ordinal data include frequencies, proportions, percentages, central points, percentiles, median, mode, and the interquartile range. Here the visualization methods that can be used are the same as nominal data. (*Types in Statistics - Qualitative and Quantitative Data Types*, n.d). Fig 3 shows an example of ordinal data; note that the difference between Elementary and High School is different from the difference between High School and College. This is the main limitation of ordinal data, the differences between the values is not really known. Because of that, ordinal scales are usually used to measure non-numeric features like happiness, customer satisfaction and so on. (Donges, n.d)

What is your Gender?

Female

Male

What languages do you speak?

Englisch

French

German

Spanish

Fig 2. Nominal data (Donges, n.d)

What Is Your Educational Background?

1 - Elementary

2 - High School

3 - Undgraduate

4 - Graduate

Fig 3. Ordinal data (Donges, n.d)

2.2 Numerical data types

- Numerical data are also known as quantitative data types. They may be discrete or continuous:
- Discrete data - Discrete data have distinct and separate values. Therefore, they are data with fixed points and can't take any measures in between. So all counted data are discrete data. Some examples of discrete data include shoe sizes, number of students in class, number of languages an individual speaks, etc. (*Types in Statistics - Qualitative and Quantitative Data Types*, n.d)
- Continuous - Continuous data, on the other hand, represent an endless range of possible values within a specified range. It can be divided into finer parts to be measured but not counted. Continuous data examples include temperature range, height, weight, etc. (*Types in Statistics - Qualitative and Quantitative Data Types*, n.d). Both height and weight can be described by using intervals on a number line; for example, height can be 184.3 cm, and weight can be 88.5 kgs.

2.2 Numerical data types

- Interval data: Interval values represent ordered units that have the same difference. Therefore we speak of interval data when we have a variable that contains numeric values that are ordered and where we know the exact differences between the values. (Donges, n.d) A common example used is the temperature scale which we are all familiar with, time when measured using the 12 hour clock, and so on. The bigger issue with interval data is that it does not have a true zero; for example there is no such thing as no temperature or no time. You can compare the data with interval data and add/subtract the values but cannot multiply or divide as it doesn't have a meaningful zero. The descriptive statistics you can apply for interval data include central point, range, and spread. (*Types in Statistics - Qualitative and Quantitative Data Types*, n.d).

2.2 Numerical data types

- An example of interval data in a research setting is as follows (*7 Types of Data in Business (with Examples and Uses)*, 2022):
- *A company sends out a survey to its customers in an effort to gather demographic information. One of the questions asks about customers' income, presenting choices as categories based on increments:*
 - *Below \$20,000*
 - *\$20,000 to \$40,000*
 - *\$40,000 to \$60,000*
 - *\$60,000 to \$80,000*
 - *\$80,000 to \$100,000*
 - *Above \$100,000*
- *The company might find that the \$40,000 to \$60,000 range appears most frequently in the responses. Seeing that, it can decide to shift its marketing to appeal more directly to that category of consumers.*
- Ratio data: Ratio values are also ordered units that have the same difference. Ratio values are the same as interval values, with the difference that they do have an absolute zero. Good examples are height, weight, length, etc. (Denges, n.d)



Part 3

Sources of error in measurement

Sources of error

- Sources of error in measurement are attributed to four sources: the respondent, the interviewer, the instrument (i.e., the survey questionnaire), and the mode of data collection. (Bavdaž, 2010). These are described in this part of the lesson:
- Respondent : At times the respondent may be reluctant to express strong negative feelings or it is just possible that he may have very little knowledge but may not admit his ignorance. All this reluctance is likely to result in an interview of 'guesses.' Transient factors like fatigue, boredom, anxiety, etc. may limit the ability of the respondent to respond accurately and fully. (Kothari, 2004)
- Interviewer: The interviewer can distort responses by rewording or reordering questions. His behavior, style and looks may encourage or discourage certain replies from respondents. Careless mechanical processing may distort the findings. Errors may also creep in because of incorrect coding, faulty tabulation and/or statistical calculations, particularly in the data-analysis stage (Kothari, 2004). Further the investigation by Bavdaž (2010) also opined that there is "some evidence that the survey staff involved in the response process could also become a source of measurement error".

Sources of error

- Instrument: Error may arise because of the defective measuring instrument. The use of complex words, beyond the comprehension of the respondent, ambiguous meanings, poor printing, inadequate space for replies, response choice omissions, etc. are a few things that make the measuring instrument defective and may result in measurement errors. Another type of instrument deficiency is the poor sampling of the universe of items of concern. (Kothari, 2004). Further Bavdaž(2010) also confirms this based on his investigation.
- Mode of data collection: Kothari (2004) refers to these as situational factors and adds the following regarding them: Any condition which places a strain on interview can have serious effects on the interviewer-respondent rapport. For instance, if someone else is present, he can distort responses by joining in or merely by being present. If the respondent feels that anonymity is not assured, he may be reluctant to express certain feelings.



Part 4

Tests of sound measurement

3.1 Introduction

- For a measurement instrument to be accepted as being of the required standard to fulfill the needs of the research it must pass three tests; these are the tests of validity, reliability and practicality. What do these three terms mean for the measurement tool?
- “Validity refers to the extent to which a test measures what we actually wish to measure. Reliability has to do with the accuracy and precision of a measurement procedure ... Practicality is concerned with a wide range of factors of economy, convenience, and interpretability ...” (Thorndike and Hagen, as cited by Kothari, 2004).

3.2 Test of validity

- (Mohajan, 2017) observes the following in the paper on validity and reliability of research instruments: “
- Validity is often defined as the extent to which an instrument measures what it asserts to measure [Blumberg et al., 2005]. Validity of a research instrument assesses the extent to which the instrument measures what it is designed to measure (Robson, 2011). It is the degree to which the results are truthful. So that it requires research instrument (questionnaire) to correctly measure the concepts under the study (Pallant, 2011). It encompasses the entire experimental concept, and establishes whether the results obtained meet all of the requirements of the scientific research method.”
- The implication for both qualitative and quantitative research respectively is as follows (Mohajan, 2017):
- “Qualitative research is based on the fact that validity is a matter of trustworthiness, utility, and dependability [Zohrabi, 2013]. Validity of research is an extent at which requirements of scientific research method have been followed during the process of generating research findings. It is a compulsory requirement for all types of studies [Oliver, 2010]. ...(cont'd)

3.2 Test of validity

- (cont'd)...In quantitative research validity is the extent to which any measuring instrument measures what it is intended to measure [Thatcher, 2010]. But, in qualitative research it is when a researcher uses certain procedures to check for the accuracy of the research findings [Creswell, 2014]. It is not a property of the instrument, but of the instrument's scores and their interpretations. It is the best viewed as a hypothesis for which evidence is collected in support of proposed inferences [Messick, 1989]. Lee J. Cronbach and Paul E. Meehl first introduced the issue of validity in quantitative research in the mid 20th century in relation to the establishment of the criteria for assessing psychological tests [Cronbach & Meehl, 1955]. “
- There are three types of validity tests that the instrument should pass:
 - Content validity
 - Criterion-related validity
 - Construct validity

3.2.1 Content validity

- Kothari (2004) describes content validity as follows: “the extent to which a measuring instrument provides adequate coverage of the topic under study. If the instrument contains a representative sample of the universe, the content validity is good. Its determination is primarily judgemental (sic) and intuitive. It can also be determined by using a panel of persons who shall judge how well the measuring instrument meets the standards, but there is no numerical way to express it.”
- Mohajan (2017) describes it as follows: “It is the extent to which the questions on the instrument and the scores from these questions represent all possible questions that could be asked about the content or skill [Creswell, 2005]. It ensures that the questionnaire includes adequate set of items that tap the concept. The more the scale items represent the domain of the concept being measured, the greater the content validity [Shekaran & Bougie, 2010]. With it is the interested (sic) in assessing current performance rather than predicting future performance. It is related to a type of validity in which different elements, skills and behaviors are adequately and effectively measured [DeVellis, 2006; Messick, 1995]. There is no statistical test to determine whether a measure adequately covers a content area, content validity usually depends on the judgment of experts in the field. The unclear and obscure questions can be amended, and the ineffective and nonfunctioning questions can be discarded by the advice of the reviewers...”

3.2.1 Content validity (cont'd)

- To effectively evaluate content validity, L. Crocker and J. Algina suggest the four steps procedures as [Crocker and Algina, 2010, as cited in Mohajan, 2017]:"
 - identify and outline the domain of interest,
 - gather resident domain experts,
 - develop consistent matching methodology, and
 - analyze results from the matching task
- There are two types of content validity, namely face validity and logic validity. However, most literature use the term logical validity and content validity interchangeably.

3.2.1 Content validity (cont'd)

- Mohajan (2017) defines face validity separately: “It is considered as a basic and minimum index of content validity, but it is determined after the test is constructed... The concepts of content evidence and face validity bear superficial resemblance, but they are in fact quite different. Face validity refers to the degree to which a test appears to measure what it claims to measure... It ascertains that the measure appears to be assessing the intended construct under study. It is usually used to describe the appearance of validity without empirical testing... So, it is normally considered to be the weakest form of validity... For example, estimating the speed of a car based on its outward appearance (guesswork) is face validity... If the test is known to have content validity, face validity can be assumed, but face validity does not ensure content validity. The stakeholders can easily assess face validity. Although this is not a very scientific type of validity, it may be an essential component for enlisting motivation of stakeholders...”

3.2.2 Criterion – related validity

- “It is used to predict future or current performance. It correlates test results with another criterion of interest... It deals with relationship between scale scores, and some specific measurable criterion. It tests how the scale differentiates individuals on a criterion it is expected to predict... That is, when we are expecting a future performance based on the scores obtained currently by the measure, correlate the scores obtained with the performance...” (Mohajan, 2017).
- Kothari (2004) avers that the criterion must possess the following qualities: “
 - *Relevance*: (A criterion is relevant if it is defined in terms we judge to be the proper measure.)
 - *Freedom from bias*: (Freedom from bias is attained when the criterion gives each subject an equal opportunity to score well.)
 - *Reliability*: (A reliable criterion is stable or reproducible.)
 - *Availability*: (The information specified by the criterion must be available.)”
- Criterion – related validity is established by concurrent validity and predictive validity.

3.2.2 Criterion – related validity (cont'd)

- Concurrent validity: “It is the degree to which the scores on a test are related to the scores on another, already established as valid, designed to measure the same construct, test administered at the same time or to some other valid criterion available at the same time.... Example, a new simple test is to be used in place of an old troublesome one, which is considered useful; measurements are obtained on both at the same time.” (Mohajan, 2017).
- Predictive validity: “It is a test constructed and developed for the purpose of predicting some form of behavior...It indicates the ability of the measuring instrument to differentiate among individuals with reference to a future criterion...(for example)...Test that are constructed to pick applicants who are most likely to be successful subsequently in their training while rejecting those applicants who are most likely to be failures if given admission.” (Mohajan, 2017).
- Kothari (2004) adds that “Criterion-related validity is expressed as the coefficient of correlation between test scores and some measure of future performance or between test scores and scores on another measure of known validity.” while Mohajan (2017) concludes that “Logically, predictive and concurrent validation are the same, the term concurrent validation is used to indicate that no time elapsed between measures. The higher the correlation between the criterion and the predictor indicates the greater the predictive validity. If the correlation is perfect, that is 1, the prediction is also perfect. Most of the correlations are only modest, somewhere between 0.3 and 0.6.”

3.2.3 Construct validity

- A measure is said to possess construct validity to the degree that it confirms to predicted correlations with other theoretical propositions. Construct validity is the degree to which scores on a test can be accounted for by the explanatory constructs of a sound theory. For determining construct validity, we associate a set of other propositions with the results received from using our measurement instrument. If measurements on our devised scale correlate in a predicted way with these other propositions, we can conclude that there is some construct validity. (Kothari, 2004)
- Construct validity can be divided in two: convergent validity and discriminant validity.
- Convergent validity “refers to the extent to which scores on a measure share a high, medium or low relationship with scores obtained on a different measure intended to assess the similar construct [Messick, 1995]. It is established when the scores obtained with two different instruments measuring the same concept are highly correlated.” (Mohajan, 2017)
- Discriminant validity on the other hand “is established when, based on theory, two variables are predicted to be uncorrelated, and the scores obtained by measuring them are indeed empirically found to be so, that is, to differentiate one group from another. It is the lack of a relationship among measures which theoretically should not be related.”
- Let us describe an example of each based on the definitions, in order to understand them better.

3.2.3 Construct validity (cont'd)

- “Let’s say you were researching depression in college students. In order to measure depression (the construct), you use two measurements: a survey and participant observation. If the scores from your two measurements are close enough (i.e. they converge), this demonstrates that they are measuring the same construct. If they don’t converge, this could indicate they are measuring different constructs (for example, anger and depression or self-worth and depression).” (Glen, 2015). This is an example of convergent validity.
- Surveys that are used to identify potential high school drop-outs would have discriminant validity if the students who graduate score higher on the test than students who leave before graduation (Campbell, 1959, as cited in Mohajan, 2017).

3.3 Test of reliability

- A measuring instrument is reliable if it provides consistent results. Reliable measuring instrument does contribute to validity, but a reliable instrument need not be a valid instrument. For instance, a scale that consistently overweighs objects by five kgs., is a reliable scale, but it does not give a valid measure of weight. But the other way is not true i.e., a valid instrument is always reliable. Accordingly reliability is not as valuable as validity, but it is easier to assess reliability in comparison to validity. (Kothari, 2004)
- In quantitative research, reliability refers to the consistency, stability and repeatability of results, that is, the result of a researcher is considered reliable if consistent results have been obtained in identical situations but different circumstances. But, in qualitative research it is referred to as when a researcher's approach is consistent across different researchers and different projects. (Mohajan, 2017).
- Both Kothari (2004) and Mohajan (2017) state that stability and equivalence are aspects of reliability that should always be mentioned. Stability "is defined as the ability of a measure to remain the same over time despite uncontrolled testing conditions or respondent themselves." (Mohajan, 2017). The *equivalence aspect* considers how much error may get introduced by different investigators or different samples of the items being studied (Kothari, 2004)
- Two methods to test stability are: i) test-retest reliability, and ii) parallel-form reliability. 30

3.3 Test of reliability (cont'd)

- Test-retest reliability: (When) the reliability coefficient is obtained by repetition of the same measure on a second time, (it) is called the test-retest reliability (Graziano and Raulin, 2006, as cited in Mohajan, 2017).
- Parallel-forms reliability: It is a measure of reliability obtained by administering different versions of an assessment tool to the same group of individuals. The scores from the two versions can then be correlated in order to evaluate the consistency of results across alternate versions. If they are highly correlated, then they are known as parallel-form reliability (DeVellis, 2006, as cited in Mohajan, 2017).

3.3 Test of reliability (cont'd)

- Kothari (2004) suggests that reliability can be improved in the following ways:
- By standardizing the conditions under which the measurement takes place i.e., we must ensure that external sources of variation such as boredom, fatigue, etc., are minimized to the extent possible. That will improve stability aspect.
- By carefully designed directions for measurement with no variation from group to group, by using trained and motivated persons to conduct the research and also by broadening the sample of items used. This will improve equivalence aspect.
- Fig 4 shows the structure of reliability and validity. Some parts of the diagram have not been discussed due to time and space constraints; the learner is encouraged to do some further reading on these.

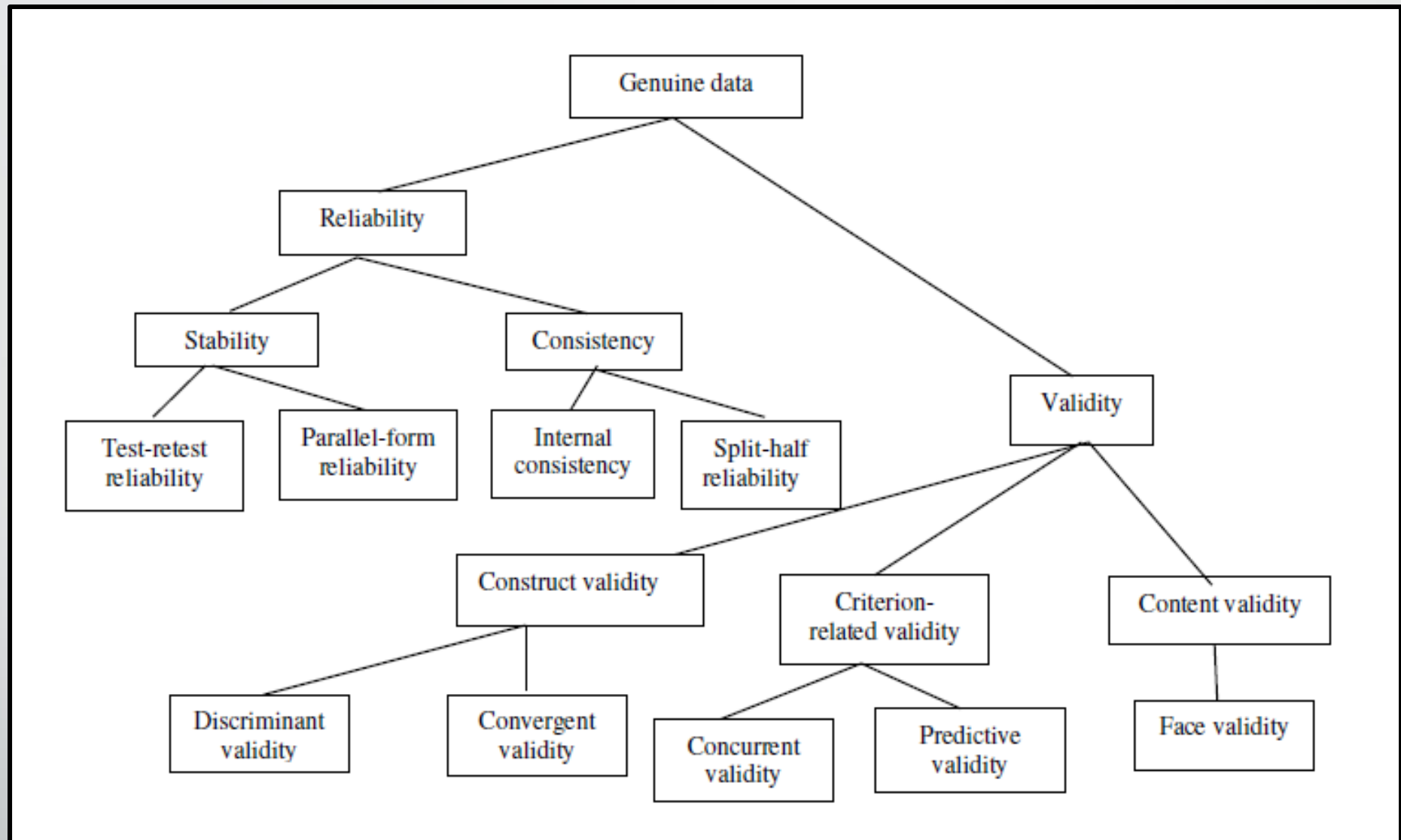


Fig 4. Structure of validity and reliability (Bajpai and Bajpai (2014), as cited in Mohajan, 2017)³³

3.4 Test of practicality

- Kothari (2004) succinctly sums up the test of practicality:
- “The practicality characteristic of a measuring instrument can be judged in terms of economy, convenience and interpretability. From the operational point of view, the measuring instrument ought to be practical i.e., it should be economical, convenient and interpretable. *Economy* consideration suggests that some trade-off is needed between the ideal research project and that which the budget can afford. The length of measuring instrument is an important area where economic pressures are quickly felt. Although more items give greater reliability as stated earlier, but in the interest of limiting the interview or observation time, we have to take only few items for our study purpose. Similarly, data-collection methods to be used are also dependent at times upon economic factors. *Convenience* test suggests that the measuring instrument should be easy to administer. For this purpose one should give due attention to the proper layout of the measuring instrument. For instance, a questionnaire, with clear instructions (illustrated by examples), is certainly more effective and easier to complete than one which lacks these features. *Interpretability* consideration is specially important when persons other than the designers of the test are to interpret the results. The measuring instrument, in order to be interpretable, must be supplemented by (a) detailed instructions for administering the test; (b) scoring keys; (c) evidence about the reliability and (d) guides for using the test and for interpreting results.”



Part 5

Technique of developing measurement tools

5.1 Technique

- The technique of developing measurement tools involves a four-stage process (Kothari, 2004):
- Concept development – in this first stage the investigator purposes to understand the main concepts relating to the study.
- Specification of concept dimensions – in this second stage develops the dimensions related the ones developed in the first stage. This can be done by adopting a more or less intuitive approach or by empirical correlation of the individual dimensions with the total concept and/or the other concepts. For example in a clinic one may think of patients, patient satisfaction, staff satisfaction, customer care, and so on.
- Selection of indicators – after completing the first two stages the investigator must now develop indicators or measures of the dimensions. Indicators are specific questions, scales, or other devices by which respondent's knowledge, opinion, expectation, etc., are measured.
- Formation of index - When we have several dimensions of a concept or different measurements of a dimension, we may need to combine them into a single index. One simple way for getting an overall index is to provide scale values to the responses and then sum up the corresponding scores.

5.2 Scaling

- After developing the measurement tools, the last task that remains is scaling.
- Scaling describes the procedures of assigning numbers to various degrees of opinion, attitude and other concepts. This can be done in two ways viz., (i) making a judgement about some characteristic of an individual and then placing him directly on a scale that has been defined in terms of that characteristic and (ii) constructing questionnaires in such a way that the score of individual's responses assigns him a place on a scale. (Kothari, 2004)
- Scales can be developed using any of the following techniques (Kothari, 2004):
 - Arbitrary approach – scale is developed on ad hoc basis;
 - Consensus approach – subject matter experts are invited to evaluate items chosen for inclusion.
 - Item analysis approach – individual items are developed into a test and given to a group of respondents; how they respond will determine which items to include in the final (those which show some clear discrimination between high and low scores).
 - Cumulative scales – these are chosen on the basis of their conforming to some ranking of items with ascending and descending discriminating power
 - Factor scales - may be constructed on the basis of inter-correlations of items which indicate that a common factor accounts for the relationship between items. This relationship is typically measured through factor analysis method.

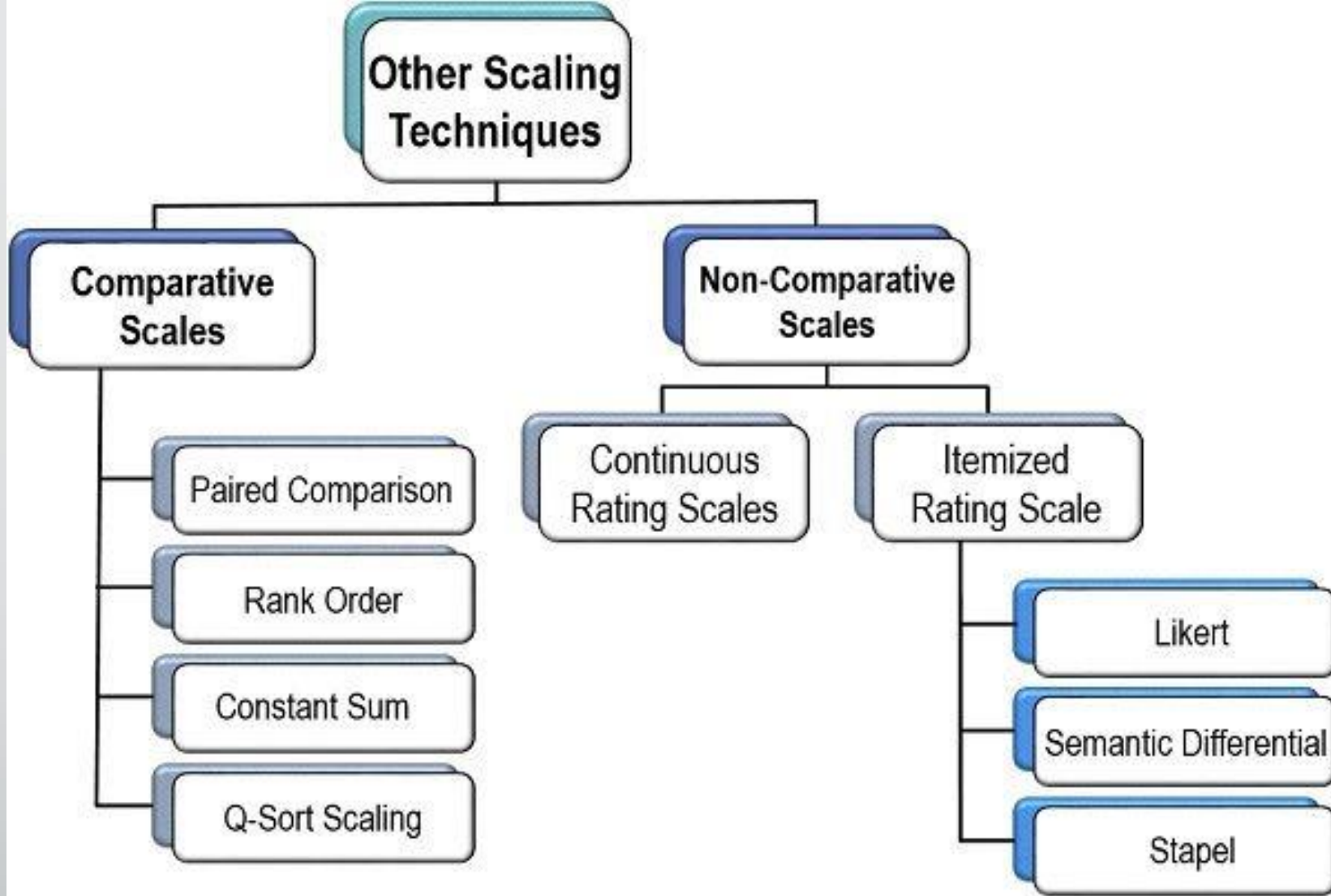


Fig 5. Scaling techniques. Source: <https://theinvestorsbook.com/scaling-techniques.html>, accessed 12/04/23

5.3 Scaling techniques

- Some of the important scaling techniques are depicted in fig 5. Other scaling techniques were discussed earlier in the lesson. Broadly speaking, for comparing two or more variables, a comparative scale is used by the respondents. A non-comparative scale is used to analyze the performance of an individual product or object on different parameters.
- The learner is encouraged to read on these in their own time. Some are commonly used in day to day activities, and there is a high chance that you have already used a few of these. For example the Likert scale is used a lot in research circles, mostly in social sciences.
- Table 1 may act as a guide in determining which scale to use vs one of the approaches described. Notice new scales have been introduced, not seen in fig 5 or anywhere else in the lesson. This goes to show that scales by itself needs a whole lesson (or even course)!

Table 1. Different scales for measuring attitudes of people (Kothari, 2004)

<i>Name of the scale construction approach</i>	<i>Name of the scale developed</i>
Arbitrary approach	Arbitrary scales
Consensus scale approach	Differential scales (such as Thurston Differential scale)
Item analysis approach	Summated scales (such as Likert Scale)
Cumulative scale	approach Cumulative scales (such as Guttman's Scalogram)
Factor analysis approach	Factor scales (such as Osgood's Semantic Differential, Multi-dimensional Scaling, etc.)

Summary

- Measurement scales include categorical (qualitative) or numerical (quantitative) data. Categorical data may be either ordinal or nominal. Numerical data may be either discrete or continuous; further they are classified as being interval or ratio.
- Sources of error in measurement are attributed to four sources: the respondent, the interviewer, the instrument (i.e., the survey questionnaire), and the mode of data collection.
- For a measurement instrument to be accepted as being of the required standard to fulfill the needs of the research it must pass three tests; these are the tests of validity, reliability and practicality.
- The technique of developing measurement tools involves a four-stage process: Concept development, specification of concept dimensions, selection of indicators, and formation of index.
- Scales can be developed using any of the following techniques: arbitrary approach, consensus approach, item analysis approach, cumulative scales and factor scales.
- Scaling techniques can be broadly classified as being either comparative or non-comparative. For comparing two or more variables, a comparative scale is used by the respondents. A non-comparative scale is used to analyze the performance of an individual product or object on different parameters.

References

- *7 Types of Data in Business (With Examples and Uses)*. (2022, December 11). Au.indeed.com; Indeed Editorial team. <https://au.indeed.com/career-advice/career-development/types-of-data>
- Bavdaž, M. (2010). Sources of Measurement Errors in Business Surveys. *Journal of Official Statistics*, 26(1).
- Donges, N. (n.d.). *Data Types in Statistics*. Built In. Retrieved April 12, 2023, from <https://builtin.com/data-science/data-types-statistics>
- Glen, S. (2015, September 8). *Convergent Validity and Discriminant Validity: Definition, Examples*. Statistics How To. <https://www.statisticshowto.com/convergent-validity/>
- Kothari, C. R. (2004). *Research methodology : methods & techniques* (2nd ed.). New Age International (P) Ltd., Publishers, Cop.
- Mohajan, H. (2017). Two Criteria for Good Measurements in Research: Validity and Reliability Two Criteria for Good Measurements in Research: Validity and Reliability. In *MPRA*. IOP Publishing. https://mpra.ub.uni-muenchen.de/83458/1/MPRA_paper_83458.pdf
- *Scaling Techniques - Definition, Types*. (2019, May 14). The Investors Book. <https://theinvestorsbook.com/scaling-techniques.html>