



Batterjee Medical College
For Science and Technology

Scientific Nursing Research SNRNU 401

Measurement and Data Collection

Dr. Nadia Hassan Ali Awad

Associate Professor, nursing program

Objectives by the end of the lecture, the students will be able to:

- **Define** measurement and its importance in nursing and health research.
- **Differentiate** between types of data (nominal, ordinal, interval, ratio).
- **Identify** various methods of data collection used in research.
- **Explain** the concepts of validity and reliability in measurement.
- **Apply** appropriate measurement and data collection techniques to research problems.

Outlines



Introduction to Measurement



Types of Data



Methods of Data Collection



Validity in Measurement



Reliability in Measurement.

Introduction

Measurement is a very important part of the quantitative research process. When quality measurement methods are used in a study, it improves the accuracy or validity of study outcomes or findings.

Definition of measurement

Measurement is the process of assigning numbers or values to individuals' health status, objects, events, or situations using a set of rules.

- **To produce trustworthy data that can be used in statistical analyses.**

For example, we measure a patient's blood pressure (BP) using a measurement method such as a stethoscope, cuff, and sphygmomanometer. Then a number or value is assigned to that patient's BP, such as 120/80 mm Hg.

CONCEPTS OF MEASUREMENT THEORY

Measurement theory guides the development and use of measurement methods or tools in research.

It provides rules that guide how things are measured. These rules allow individuals to be consistent in how they perform measurements; thus, a measurement method used by one person will consistently produce similar results when used by another person.

Basic concepts and rules of measurement theory

1. Directness of measurement

- A. Direct measure
- B. Indirect measure

2. Levels of measurement

- A. Nominal-Level Measurement
- B. Ordinal-Level Measurement
- C. Interval-level measurement
- D. Ratio-Level Measurement

3. Measurement error

- A. Random error
- B. Systematic error

4. Reliability

5. Validity

1. Directness of Measurement

```
graph TD; A[1. Directness of Measurement] --> B(A. Direct measures); A --> C(B. Indirect measures);
```

**A. Direct
measures**

**B. Indirect
measures**

A. Direct measures

Direct measures involve determining the value of concrete factors such as weight, waist circumference, temperature, heart rate, BP, and respiration.

Technology is available to measure many bodily functions, biological indicators, and chemical characteristics

B. Indirect measures or indicators

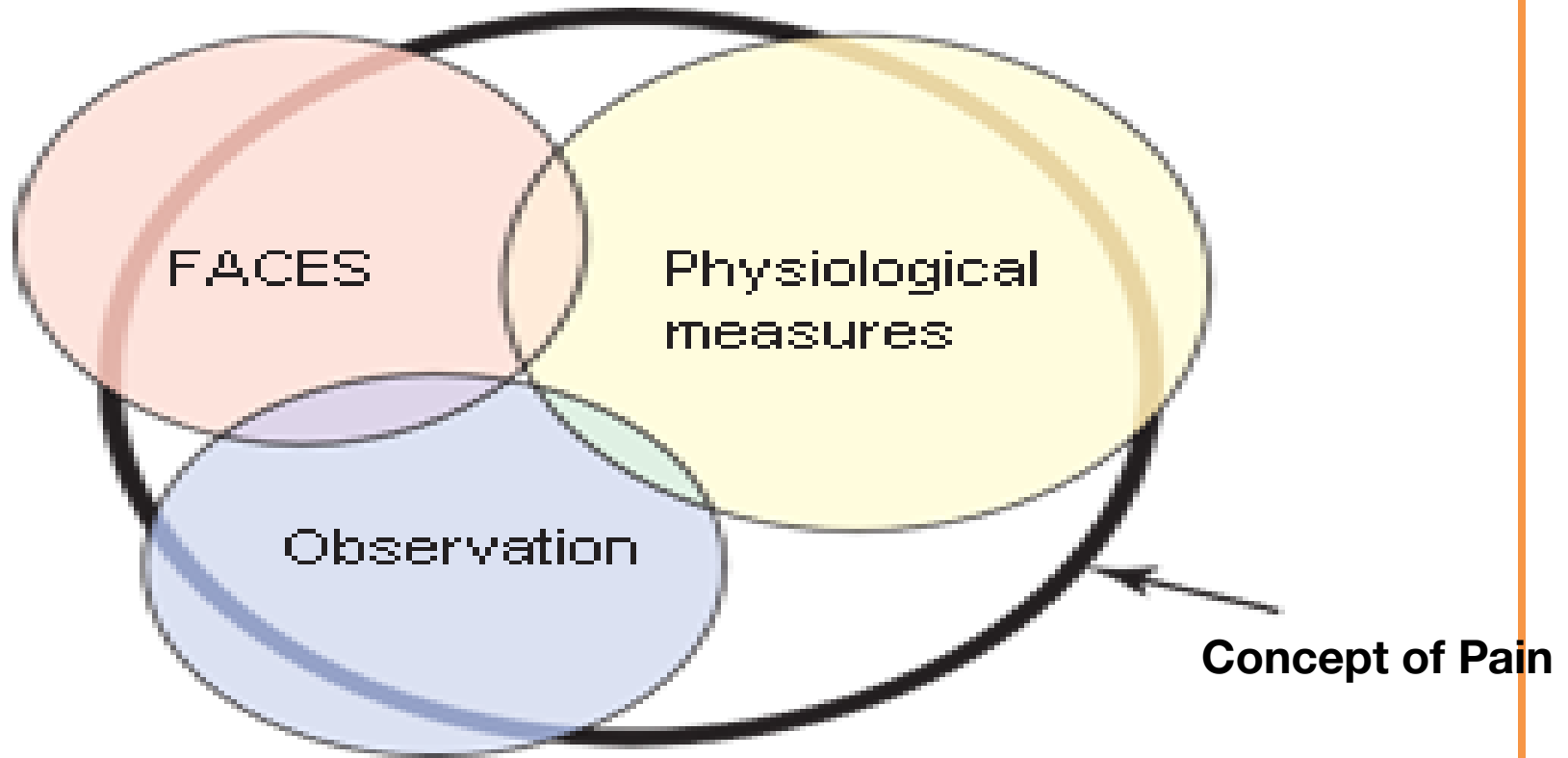
In many cases in nursing, the thing to be measured is not a **concrete object but an abstract idea**, characteristic, or concept such as **pain, stress, caring, coping, depression, anxiety, and adherence.**

- Researchers **cannot directly measure an abstract idea**, but they can capture some of its elements in their measurements, which are referred to as **indirect measures or indicators of the concepts.**

Therefore, multiple measurement methods or indicators are needed, and even then they cannot be expected to measure all elements of an abstract concept.

For example, multiple measurement methods might be used to describe pain in a study, which decreases the measurement error and increases the understanding of pain. The measurement methods of pain might include the FACES Pain Scale, observation (rubbing and/or guarding the area that hurts, facial grimacing, and crying), and physiological measures, such as pulse and blood

Example of indirect measure for the concept



Multiple measures of the concept of pain

2. Levels of Measurement

Various measurement methods produce data that are at different levels of measurement.

- 1. Nominal-level measurement**
- 2. Ordinal-level measurement**
- 3. Interval-level measurement**
- 4. Ratio-level measurement**

A. Nominal-level measurement

Nominal-level measurement is the lowest of the four measurement categories. It is used when data can be organized into categories of a defined property, but the categories cannot be rank-ordered.

E.g.

- Diagnosis (kidney stone, gastric ulcer, ovarian cyst)
- Gender (male, female)
- Marital status (single, married, divorced, widow)
- Race
- Ethnicity

B. Ordinal-level measurement

Ordinal-level measurement, data are assigned to categories that can be ranked. one category is judged to be higher or lower, or better or worse, than another category.

Many scales used in nursing research are ordinal levels of measurement. For example, it is possible to rank degrees of coping, levels of mobility, ability to provide self-care, or levels of dyspnea on an ordinal scale. For dyspnea with activities of daily living (ADLs), the scale could be:

- 0/ no shortness of breath with ADLs
- 1/ minimal shortness of breaths with ADLs
- 2/ moderate shortness of breath with ADLs
- 3/ extreme shortness of breath with ADLs

Not possible to claim that equal distances exist between the rankings

C. Interval-level measurement

Interval-level measurement uses interval scales, which have **equal numerical distances between intervals**.

E.g. The difference between the temperatures of 70°F and 80°F is 10°F and is the same as the difference between the temperatures of 30°F and 40°F.

It is not possible to provide the absolute amount of the attribute, because the interval scale lacks

D. Ratio-level measurement

Ratio-level measurement is the highest form of measurement and meets all the rules of other forms of measurement (mutually exclusive categories, exhaustive categories, ordered ranks, equally spaced intervals, and a continuum of values. In addition, ratio-level measures have absolute zero points.

e.g. Weight, length, and volume are commonly used as examples of ratio scales.

3. Measurement Error

- **True measure or score:** The ideal perfect measure.
- **Measurement error** is the difference between the true measure and what is actually measured
- **Two types of measurement errors:**
 1. Random measurement error
 2. systematic measurement error

A. Random measurement error

Random measurement error: the difference between the measured value and the true value is without pattern or direction (random).

In one measurement, the actual value obtained may be lower than the true value, whereas in the next measurement, the actual value obtained may be higher than the true value. A number of chance situations or factors can occur during the measurement process that can result in random error ([Waltz, Strickland, & Lenz, 2010](#)).

e.g, the person taking the measurements may not use the same procedure every time, a subject completing a paper and pencil scale may accidentally mark the wrong column, or the person entering the data into a computer may punch the wrong key. The purpose of measuring is to estimate the true value, usually by combining a number of values and calculating an average.

B. Systematic measurement error

In **systematic measurement error**: the variation in measurement values from the calculated average is primarily in the same direction.

e.g. a scale that shows weights that are 2 pounds over the true weights will give measures with systematic error. All the measured weights will be high, and as a result the mean will be higher than if an accurate weight scale were used.

4. Reliability (page 289)

Reliability is concerned with the **consistency of a measurement method**.

e.g. if you are using a paper and pencil scale to measure depression, it should indicate similar depression scores each time a subject completes it within a short time.

- Reliability testing is a measure of the amount of random error in the measurement technique.
- Because all measurement techniques contain some random error, reliability exists in degrees and usually is expressed as a reliability coefficient (r).
- A scale that does not produce similar scores for a subject with repeat testing is considered unreliable and results in increased

4. Reliability (page 289)

Reliability testing focuses on the following three aspects of reliability: stability, equivalence, and homogeneity

1. Stability: is concerned with the consistency of repeated measures of the same attribute with the use of the same scale or instrument. It is usually referred to as **test-retest reliability**. The authors report the coefficient of stability (test-retest reliability) at 2 weeks and 1 month (0.87 and 0.81, respectively).

Reliability

2. Equivalence/ interrater reliability: which involves the comparison of two versions of the same paper-and-pencil instrument or two observers measuring the same event.

Measured through Cohen's kappa

3. Homogeneity/ internal consistency reliability: reliability is defined as consistency across items of an instrument, with individual items being individual questions. To the extent that certain items “hang together” and measure the same thing, an instrument is said to possess high internal consistency reliability. Measured

through Cronbach's alpha (coefficient alpha)/e.g alpha

E. Validity (page 291)

Validity is a determination of how well the instrument reflects the abstract concept being examined. That's means the instrument measure what is intended to measure.

Three common types of validity are:

1. Content validity
2. Criterion-related validity
3. Construct validity

1. Content validity

Content validity is the extent to which an instrument or test measures an intended content area.

- Usually, this type of validity is used in the development of a questionnaire, interview schedule, interview guide, or instrument.
- Content validity is then determined by a panel of experts. Experts carefully evaluate all items or questions used by the instrument, as well as the instrument's overall appropriateness

2. Criterion -related validity

Criterion-related validity is a measure of how well an instrument measuring a particular concept compares with a criterion, providing more quantitative evidence on the accuracy of the instrument.

In some instances, a newer instrument is compared with an older, more reputable instrument.

3. Construct validity

Construct validity is the extent to which an instrument measures an intended hypothetical concept or construct.

The process of validating a construct is by no means easy. Construct validity is the most valuable, yet the most difficult, way to assess an instrument's validity

Data collection

Qualitative vs. Quantitative Data



Data collection is a major part of the research process. Methods and instruments for data collection must be chosen according to the nature of the problem, approach to the solution, and variables being studied. Whether the researcher is testing hypotheses or seeking answers to research questions, valid and reliable instruments are essential.

Data collection

Quantitative Data



Quantitative Data Collection Methods

Quantitative data collection methods include the use of instruments. An instrument is a device used to record or gather data on a particular concept.

- **Data-gathering instruments** used in research studies include questionnaires, rating scales, checklists, standardized tests, and biophysical measures.
- **Methods for collecting data** may be based on a form of self-report that asks individuals to complete a questionnaire, survey, or standardized test





1. Surveys and Surveys Questionnaires

are a popular method for collecting data to describe, compare, or explain knowledge, attitudes, and behavior. A survey is a series of questions posed to a group of subjects.

Questionnaires are structured self-administered surveys. The most common way of distributing questionnaires is through the mail, although many research situations allow for face-to-face administration. A questionnaire may be **closed-ended or open-ended.**

Characteristics of a good questionnaire

- Ask questions that are specific rather than general. E.g. How would you rate your level of satisfaction regarding being able to have input into your treatment preferences—excellent, very good, good, fair, or poor
- Use simple language
- Each question should represent one concept
- Phrase questions in as neutral a way as possible

TABLE 9.1 Example of Open-Ended and Close-Ended Questions

Open-Ended Question

Describe some reasons to wear a seat belt when riding in a car driven by someone else.

Closed-Ended Question

1. How often do you wear a seat belt when riding in a car driven by someone else?
- a. Never
 - b. Rarely
 - c. Sometimes
 - d. Most of the time
 - e. Always

Open-Ended Question

Explain your beliefs about drinking.

Closed-Ended Question

1. During the past 30 days, how many times did you drive a car when you had been drinking alcohol?
- a. 0 times
 - b. 1 time
 - c. 2–3 times
 - d. 4–5 times
 - e. 6 or more times

Questionnaire example

PATIENT SATISFACTION WITH NURSING CARE QUALITY QUESTIONNAIRE (Laschinger, McGillis Hall, Pedersen & Almost, 2005)

Please rate some things about the nursing care during your hospital stay in terms of whether they were Excellent, Very Good, Good, Fair or Poor. Please check only one rating for each statement.

	Excellent	Very Good	Good	Fair	Poor
INFORMATION YOU WERE GIVEN: How clear and complete the nurses' explanations were about tests, treatments, and what to expect.	☐	☐	☐	☐	☐
INSTRUCTIONS: How well nurses explained how to prepare for tests and operations.	☐	☐	☐	☐	☐
EASE OF GETTING INFORMATION: Willingness of nurses to answer your questions.	☐	☐	☐	☐	☐
INFORMATION GIVEN BY NURSES: How well nurses communicated with patients, families, and doctors.	☐	☐	☐	☐	☐
INFORMING FAMILY OR FRIENDS: How well the nurses kept them informed about your condition and needs.	☐	☐	☐	☐	☐
INVOLVING FAMILY OR FRIENDS IN YOUR CARE: How much they were allowed to help in your care.	☐	☐	☐	☐	☐



2. Scales

A scale is a set of numerical values assigned to responses, representing the degree to which subjects possess a particular attitude, value, or characteristic.

- a. Likert scale
- b. Semantic Differential Scales
- c. Visual Analogue Scales

a. Likert scale

Likert scales, also **called summative scales,** require subjects to respond to a series of statements to express a viewpoint.

Subjects read each statement and select an

appropriately ranked

response

TABLE 9.2 Example of Likert Responses

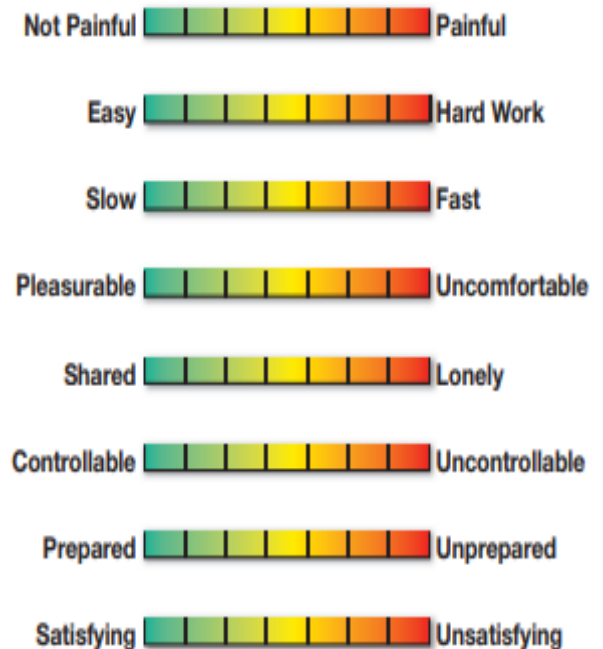
Strongly disagree	Strongly disagree	
Disagree	Disagree with reservations	
Agree	Agree with reservations	
Strongly agree	Strongly agree	
Almost never	Never	Always
Occasionally	Rarely	Frequently
Usually	Infrequently	Sometimes
Almost always	Sometimes	Rarely
	Frequently	Never
	Often	

An example of a Likert scale

Item	Answer				
	Totally Disagree (1), Disagree (2), Undecided (3), Agree (4) and Totally Agree (5)				
Denial & Anger					
DA1. I am not able to accept what is happening	(1)	(2)	(3)	(4)	(5)
DA2. I am in complete denial about what is happening	(1)	(2)	(3)	(4)	(5)
DA3. I feel like I am a coward for denying what is happening	(1)	(2)	(3)	(4)	(5)
DA4. I am more irritable as a result of this situation	(1)	(2)	(3)	(4)	(5)
DA5. I am easily irritated by anything	(1)	(2)	(3)	(4)	(5)
DA6. I hide my emotions from others	(1)	(2)	(3)	(4)	(5)
Bargaining					
BA1. I constantly repeat to myself that everything will be fine	(1)	(2)	(3)	(4)	(5)
BA2. Other people are helping me to overcome the situation	(1)	(2)	(3)	(4)	(5)
BA3. I make promises to myself in order to get through the situation	(1)	(2)	(3)	(4)	(5)

B. Semantic Differential Scales

- The semantic differential scale is composed of a set of scales using pairs of adjectives that reflect opposite feelings.

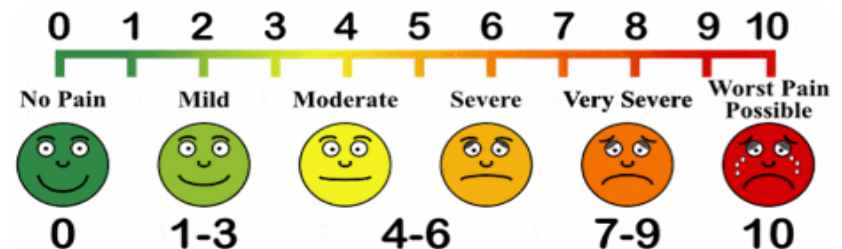


C. Visual Analogue Scales

The visual analogue scale (VAS) is useful for measuring subjective phenomena (e.g., pain, fatigue, shortness of breath, anxiety). The scale is unidimensional, quantifying intensity only.

Fig 9•3 Example of Visual Analogue Scale (VAS).

Directions: Please place an X at the point on the line that best reflects how much pain you experienced with the venipuncture.





3. Q Methodology

Q Methodology (Q sort) is a sorting technique used to study opinions, attitudes, or judgments. Participants receive a set of 40–100 cards, each with a statement, and are asked to rank them into 5–10 piles from low to high priority. The researcher specifies how many cards go in each pile, creating forced choices. Participants can rearrange cards until satisfied, usually starting with the extremes.

Q Methodology (Q-sort): Nurses' Attitudes Toward Artificial Intelligence in Healthcare

1. Overview

Q Methodology (Q-sort) is a research technique used to explore subjective viewpoints, attitudes, or beliefs. In this study, it is used to understand nurses' attitudes toward artificial intelligence (AI) in healthcare. Participants are provided with 50 statements and asked to rank them into 5 groups ranging from 'Strongly Disagree' to 'Strongly Agree' according to their personal opinions.

2. Forced Distribution

Group	Agreement Level	Number of Cards
1	Strongly Disagree	8
2	Disagree	12
3	Neutral	10
4	Agree	12
5	Strongly Agree	8

▲ 3. Q-set Statements (50)

1. 1. AI helps reduce medical errors.
2. 2. AI decreases decision-making time.
3. 3. Relying on AI may weaken nurses' clinical judgment.
4. 4. AI improves patient safety.
5. 5. Using AI increases trust in care quality.
6. 6. AI may increase the risk of technical failures.
7. 7. AI reduces human interaction with patients.
8. 8. AI provides rapid information in emergencies.
9. 9. Using AI requires continuous training.
10. 10. AI may limit nurses' professional autonomy.
11. 11. AI systems provide up-to-date scientific references.
12. 12. AI may make inaccurate decisions in some cases.
13. 13. AI reduces nurses' mental fatigue.
14. 14. AI is complex and difficult to understand.
15. 15. AI may cause loss of nursing jobs.
16. 16. AI helps prioritize patient care needs.
17. 17. AI use enhances documentation quality.
18. 18. AI may lead to overdependence on technology.
19. 19. AI improves healthcare system efficiency.
20. 20. AI may slow work when the internet is weak.
21. 21. AI helps detect complications early.
22. 22. AI systems require costly maintenance.
23. 23. AI provides strong support for novice nurses.

4. Instructions for Participants

1. Read all 50 statements carefully.
2. Sort them into five groups based on your level of agreement:
 - Group 1: Strongly Disagree
 - Group 2: Disagree
 - Group 3: Neutral
 - Group 4: Agree
 - Group 5: Strongly Agree
3. Follow the forced distribution (8, 12, 10, 12, 8).
4. You may rearrange statements until the distribution best reflects your viewpoint.



4. Instruments

Phenomena measured by psychosocial instruments (e. g., coping, stress, self-concept, self-esteem, motivation) are studied widely in nursing.



5. Biophysical Measures

Nursing research includes variables measured with physiological instruments.

Various monitoring devices have been developed to measure vital signs and other physiological data.

Data collection

Qualitative Data



1. Interview

An interview schedule, sometimes called an interview guide, is a list of topics or an open-ended questionnaire administered to subjects by a skilled interviewer.

The presence of the interviewer allows for probing of subjects' responses and decreases the possibility of vague answers.

Audio and video recordings may be used to complement interviews.



2. Participant Observation

This type of qualitative data collection is used in ethnographic research.

The technique involves direct observation through involvement with subjects in their natural setting, participating in their lifestyle activities.



3. Focus Group Interviews

Focus groups are acceptable subjects for in-depth interviews. The technique serves a variety of purposes, with the ultimate goal of observing the interactions among focus group members and detecting their attitudes, opinions, and solutions to specific topics posed by the facilitator.

Focus groups are designed to be nonthreatening, so participants can express and clarify their views in ways that are less likely to occur one-on-one.



Data collection

- Data collection is the process of acquiring subjects and collecting the data for a study.

Data collection tasks:

1. Recruiting Subjects
2. Maintaining Consistency
3. Maintaining Controls
4. Protecting Study Integrity

1. Recruiting Subjects: The research report needs to describe the study participant recruitment process. Study participants or subjects may be recruited only at the initiation of data collection or throughout the data collection period.

2. Maintaining Consistency: Consistency involves maintaining the data collection pattern for each collection event as it was developed in the research plan. A good plan will facilitate consistency and maintain the validity of the study.

3. Maintaining Controls: Researchers build controls into their study plan to minimize the influence of intervening forces on the findings. Report needs to reflect the controls implemented in a study and any problems that needed to be managed during the study.

4. Protecting Study Integrity Maintaining consistency and controls during subject selection and data collection protects the integrity or validity of the study.

Any question



