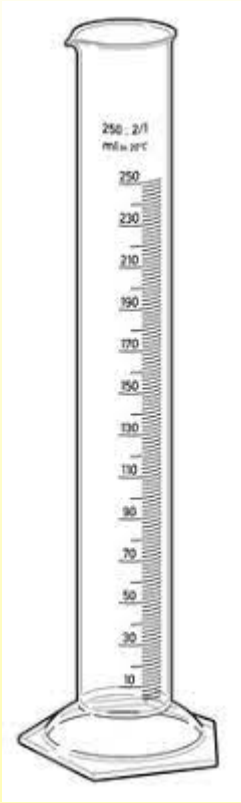
Scale and Questionnaire Development

Nutifafa Dey

PhD, Arizona State University

Measurement

- There are various approaches to measurement in the sciences



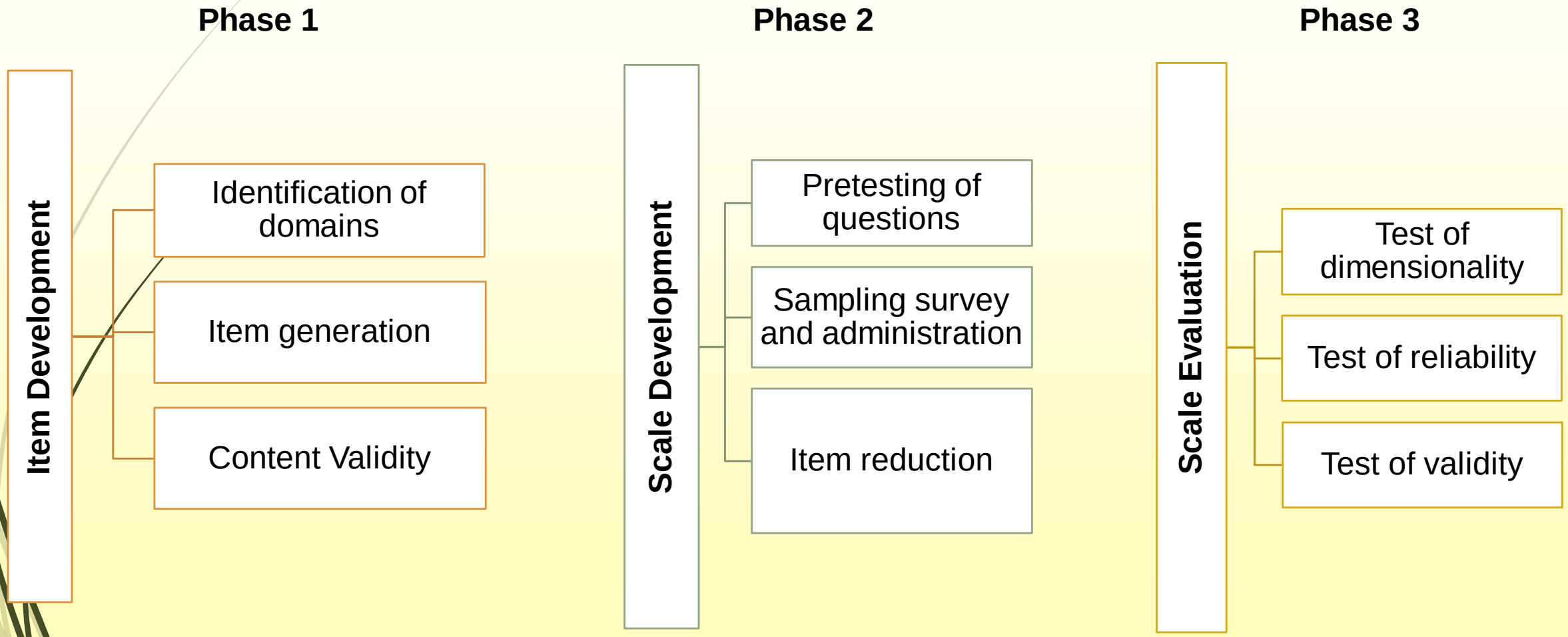
Measurement

- Social measures require measurement tools, such as scales and questionnaires
- Questionnaire (also called a test or a scale) is defined as a set of items designed to measure one or more underlying constructs (behavior, feelings, actions) also called latent variables (Fabrigar & Ebel-Lam, 2007)
- Examples of these social constructs are:
  - Stigma
  - Well-being
  - Quality of life
  - Illness perception
  - Life satisfaction
  - Depression
  - Religiosity

Scale/Questionnaire Development

- There are many steps to scale development
- Scale development is divided into three (3) phases and nine (9) processes
- It can be costly and time consuming, and statistical analysis is often required

Steps in Scale/Questionnaire Development



Boateng et al. 2018

Phase 1: Item Development
Identification of Domains

- A domain or construct refers to the concept, attribute, or unobserved behavior that is the target of the study
  - Define the construct you want to measure in detail
  - Specify the purpose and boundaries of the domain
    - Are you measuring a specific behavior or global construct?
  - Confirm that there are no existing instruments
  - Take into account existing theories, literature, and conceptual frameworks.

Phase 1: Item Development
Item Generation

Deductive and inductive methods are used to generate items

Deductive method

- Based on the description of the relevant domain and the identification of items.
- Done through literature review and assessment of existing scales and indicators of that domain

Inductive method

- Involves the generation of items from the responses of individuals
- Data is obtained through direct observations and exploratory research methodologies, such as focus groups and individual interviews

Phase 1: Item Development
Item Generation – things to consider

- Avoid double barreled items
  - E.g.
- Avoid the use of jargons
  - E.g.
- Avoid exceptionally lengthy items
  - E.g.
- Avoid redundancy
  - E.g.
- Avoid ambiguity
  - E.g.

Phase 1: Item Development
Item Generation – things to consider

- Consider the number of items do you need?
  - Usually larger than the final scale
- The more the items, the higher your reliability

Phase 1: Item Development Content Validity

- Content validity also known as “theoretical analysis” refers to the degree to which a measure assesses the domain of interest
- It is mainly assessed through evaluation by expert and target population judges
  - Evaluation by experts (5 to 7) : to determine whether each item adequately represents the domain in terms of content relevance, representatives and technical quality.
  - Evaluation by target population (interviews) : to determine whether each item adequately represents the domain in terms of representativeness of actual experience from the target community.

Phase 2: Scale Development Pretesting of Questions

- Pre-testing helps to ensure that items are meaningful to the target population before the survey is actually administered
- It minimizes misunderstanding and subsequent measurement error
- It assesses the extent to which the questions reflect the domain of interest and that answers produce valid measurements
- Administer draft questions to 5–15 interviewees in 2–3 rounds while allowing respondents to verbalize the mental process entailed in providing answers

Phase 2:
Survey and Administration

- Identify problems with the data collection and reduce measurement errors
- Requires a small sample size from the population
- Recommended sample size is 10 respondents per survey item and/or 200-300 observations
- Cross-sectional data obtained is used for exploratory factor analysis
- Make necessary revisions based on pilot testing results

Phase 2:
Item Reduction Theory

- The goal of this phase is to identify items that are not or least related to the construct of interest
- Reduction analysis is conducted to ensure that only parsimonious, functional, and internally consistent items are ultimately included
- Conduct an inter-item correlation to determine the relationship between the items

Phase 3:
Test of Dimensionality

- Determine whether the measurement of items, their factors, and function are the same across two independent samples or within the same sample at different time points
- Test if latent constructs are hypothesized
- Such tests can be conducted using independent cluster model (ICM)-confirmatory factor analysis, bifactor modeling, or measurement invariance

Phase 3:
Test for Reliability

- Assess the degree of consistency exhibited when a measurement is repeated under identical conditions
- Cronbach's alpha assesses the internal consistency of the scale items
- An alpha coefficient of 0.70 has often been regarded as an acceptable threshold for reliability
- Aim for a high level of internal consistency among items

Phase 3:
Test for Validity

- Evaluate the validity of the scale or questionnaire
- Assess content validity, construct validity, and criterion-related validity
  - Criterion-related validity: determine if scores predict future outcomes
  - Content validity: determine the extent to which scale covers all relevant domains of the construct
  - Construct validity: To examine if the same concept measured in different ways yields similar results
- Use established validation methods such as factor analysis or correlational analysis