

Measurement

Ray Block, Jr.

PS #585

Research Methods

Today's Blueprint

Last class

- Concepts
- Variables
- Hypotheses
- Theories
- Models

Today's Class

- Measurement
 - Conceptualization & Operationalization
 - Levels of Measurement
 - Validity & Reliability

Measurement

That which can be observed empirically can be measured scientifically ☺

Measurement

- What is it?
 - The assignment of values [not necessarily numbers] to empirical properties of concepts according to specified rules

Measurement

- What purpose does it serve?
 - Provides a set of procedures that tell us how to observe data
 - Provides a bridge between theoretical concepts and empirical variables

Conceptualization & Operationalization

Think conceptual and operational definitions ☺

Conceptualization

- Conceptualization:
 - The process of taking a raw idea and giving it a conceptual (or theoretical) definition

Conceptualization

- Conceptual definition
 - Defining ideas by using concepts
 - Concepts = The unobservable properties of ideas

Conceptualization

- Ex) Participation:
 - [General idea] = Those activities by private citizens aimed at influencing the government
 - [Conceptual definition] = Voter turnout in the 2000 presidential election

Operationalization

- Operationalization:
 - The process of taking a conceptual definition and making it more precise by linking it to an operational definitions

Operationalization

- Operational definition:
 - Links concepts to variables
 - Variables = The observable properties of concepts

Operationalization

- Ex) Participation:
 - [Conceptual definition] = Voter turnout in 2000 elections
 - [Operational definition] = The number of people who actually vote in the 2000 presidential election divided by the number people who are eligible to vote

Conceptualization vs. Operationalization

- In theory:
 - Conceptualization > Operationalization
 - Research *cannot* proceed w/out clearly defined concepts
- In Practice:
 - Conceptualization < Operationalization
 - Research *often* proceeds w/out clearly defined concepts

Levels of Measurement

Think N.O.I.R.; think precision ☺

Levels of Measurement

- N = Nominal:
 - Describes variables that are categorical in nature (have mutually exclusive categories)

Levels of Measurement

- Examples:
 - Nationality
 - Race/Ethnicity
 - Sex/gender
 - Religion
 - Marital status
 - State of residence
 - Zodiac sign
 - Class rank

Levels of Measurement

- O = Ordinal:
 - Describes variables that can be ordered or ranked (i.e. order of importance, more or less of something, big or little, strong or weak, etc.)

Levels of Measurement

- Examples:
 - Social class
 - Star rating system for movies or restaurants
 - Likert scales
 - Belt ranking (martial arts)
 - Class grades

Levels of Measurement

- I = Interval:
 - Describes variables that have more or less equal intervals, or meaningful distances between their ranks.

Levels of Measurement

- Examples:
 - IQ scores
 - Temperature

Levels of Measurement

- R = Ratio:
 - Describes variables that have equal intervals and a fixed zero (or reference) point
 - Zero of something \neq The absence of something

Levels of Measurement

- Examples:
 - Income
 - Education
 - Weight
 - Height

Levels of Measurement

- Precision:
 - Use measures that are as precise as possible
 - Level of measurement affects analyses
 - Nominal < Ordinal < Interval/Ratio

Levels of Measurement

- Precision:
 - Use measures that are as precise as possible
 - Level of measurement affects analyses
 - Nominal < Ordinal < Interval/Ratio

Levels of Measurement

- Precision:
 - Use measures that are as precise as possible
 - Level of measurement affects analyses
 - Nominal < Ordinal < Interval/Ratio

Levels of Measurement

- Measures of Central Tendency:
 - Measures the center of a set of observations
 - Best statistic for summarizing data
 - Mode = Most frequently occurring observation
 - Median = Middlemost observation (after they have been ordered from “smallest” to “biggest”)
 - Mean = The average of the observations

Validity and Reliability

How good are your measures?

Validity

- What is it?
 - When a variable *accurately/ truthfully/ correctly* measures what it is supposed to measure

Validity

- What purpose does it serve?
 - Checks (theoretically) the correspondence between the the measures we create and the concepts we are interested in

Validity

Decision Rules for Judging Validity:

- Content Validity:
 - Does the measure represent the universe of indicators it is suppose to represent?

- Criterion Validity:
 - Is there a relationship between the measure and the empirical outcomes?
- Construct Validity:
 - Does the measure relate to the appropriate theoretical framework?

Validity

Types of Content Validity:

- Face Validity:
 - Does it look like it measures what it is supposed to measure?
- Sampling Validity:
 - Does it contain a representative sample of the standard set of indicators commonly used represent the property being measured?

Validity

Types of Criterion Validity:

- Concurrent Validity:
 - Does the measure differentiate between groups (theoretically) known to be different?
- Predictive Validity:
 - Can you predict future meaningful outcomes from the measure?

Validity

Types of Construct Validity:

- Convergent Validity:
 - Does the measure correlate with concept(s) that it ought to correlate with?
- Divergent Validity:
 - Is the measure different from concept(s) that is should be different from?

Reliability

- What is is?
 - When a variable yields *consistently* similar results upon repeated tests
 - Can occur over time, across different situations, or between different instruments

Reliability

- What purpose does it serve?
 - Checks (statistically) the performance of a measure

Reliability

The Logic of Reliability:

$$\text{Observed Score} = \text{True Score} - \text{Score Error}$$

Where:

- *Observed Score* = What is actually recorded
- *True Score* = Theoretically, the perfect reflection of the value(s) of the variable
- *Score Error* = Everything that prevents the perfect realization of the true score
 - Method Error: Due to testing situation
 - Trait Error: Due to characteristics of individual participants

Reliability

How is reliability measured?

- Conceptually:

Reliability

In Practice:

- Reliability = Replicability (Repeatability)
 - Consistency across different tests of the same measure

- Consistency = Correlation between tests
- The stronger the correlation, the more reliable the measure (range -1 to +1)

Reliability

Types of Reliability:

- Test-Retest:
 - Consistent at two points in time using same sample

Reliability

- Alternative-Form:
 - Consistent at two points in time using different samples

Reliability

- Split-Half:
 - Consistent at same point in time using different samples (two halves created after measurement)

Reliability

- Inter-Rater, Inter-Coder, or Inter-Observer:
 - Consistency among researchers' (at least two) estimates of the same phenomenon

Reliability

- Internal Consistency:
 - Consistency of individual items within a scale or index

Validity vs. Reliability

- Reliability = Necessary condition for validity
- Validity ≠ Necessary condition for reliability
- Therefore, Reliability < Validity

Validity vs. Reliability

Threats to Validity

- Systematic Error:
 - Constant across the administration of a measurement instrument
 - Participant bias
 - Researcher bias
 - Faulty design

Threats to Reliability

- Random Error:
 - Varies across the administration of a measurement instrument
 - Setting
 - Mood
 - Data entry

Validity vs. Reliability

- Random Error = Unavoidable ☹
- Systematic Error = Controllable ☺
- Therefore, when minimizing measurement error:
 - Random error < Systematic error