Outline
- What are measurement scales?
- Limitations of ratings
- Types of rating scales
  - Graphic
  - Summated – Likert
  - Summated – Checklist
  - Other
    - Guttman
    - Thurstone
    - Semantic differential

What are Measurement Scales?

Measurement scales
- Use a scale to determine “how much” of a quantity.
- Global vs. Specific scales
  - e.g., job satisfaction
    - “how happy overall w/ your job?” (rate 1-5)
    - vs. “how happy are you with…” (rate each 1-5 and sum):
      - Autonomy
      - Work load
      - Salary
      - Coworker relations
- Unidimensional vs. Multidimensional scales
Psychologists try to do the same thing with the responses to items. Many different types of rating scales are used to capture or assess “how much” of a construct the person has. Attitudes (how strong an attitude against drunk driving), achievement motivation, personality (our friend the IPIP).

What are Measurement Scales?

- Nominal
- Ordinal
- Interval
- Ratio

...or from the book chapter which won the Pulitzer Prize for Best Book Chapter on Multiple-Item Measures?

Assumptions vs. reality of the data

Memory Effects in Rating Scales

- Must understand the strengths and limitations of the person using the scale.
- 7± 2 limit to working memory
- Frame-of-Reference effects (how much depends on previous experience)
- Halo Effects
- Other biases
  - generosity error
  - contrast error
  - acquiescent response style
Graphic Rating Scales
- Widely used – assumes people use the same frame of reference
- Person indicates how much by putting a mark at the appropriate point on a line that runs from one extreme to the other; 1 item used for scale
- The person scales “how much” for his or herself

Another Graphic Rating Scale
- Nursing Development

Summated Rating Scales
- Most common and most consistent with psychological constructs
- Likert scales are the most common example (5 or 7 point)
- Use multiple items to get multiple responses concerning a personality attribute or an attitude
- Sum the responses to form a measurement scale.
Summated Ratings

Attitude about war with Iraq

Dimensions of the attitude

Domain sampling: Should ask an item for each subcomponent of the attitude

Types of Rating Scales

Likert items for summated rating scales

- I worry that more non-combatants will be killed during this war.
- I worried that further terrorism would have occurred if Saddam was not stopped.

The resulting responses are added up to form a representation of the entire attitude

- For these two items, the resulting attitude would range from 0-10, instead of from 0-5 for each item
- Responses to multiple items have more construct validity and higher reliability than responses to single items
Checklists for summated rating scales

**Depression Adjective Checklist**
Have you felt the way described by the following adjectives in the last week?

1. Happy
2. Active
3. Blue
4. Downcast
5. Dispirited
6. Composed
7. Distressed
8. Lonely
9. Free
10. Broken

The number of endorsed responses is summed and the total score represents the person's score (because 1=yes 0=no).

Checklists work well, but generally do not provide as much information as a Likert scale of the same items.

Example: I have been happy over the last week.

Strongly Disagree
Neutral
Strongly agree

Other rating scales

- Guttman cumulative scales -
  - Should agree up to a point, then disagree with all remaining options.
  - Usually a small set of unidimensional items with increasing complexity.
- Examples:
  - Geometry test
  - Attitudes about gender equality
  - Spanking children (see next slide)
Guttman Scale Example

Attitudes about spanking children
(Monette et al., 1994)
- Slapping a child's hand is an appropriate way to teach the meaning of “no!”
- Spanking is sometimes necessary
- Sometimes discipline requires using a belt or paddle.
- Some children need a good beating to help keep them in line.

Other rating scales

- Thurstone's differential scales
  - Put the items on a number line relating to the strength of the attitude
  - e.g., attitude toward labor unions
  - Assume you will endorse those items that are at your 'level' of the attitude
  - Time consuming to construct
- Semantic differential scales
  - Measuring attitudes
  - Bipolar adjective scales (fair/unfair, light/heavy)
  - What if the IPIP were measured this way?