

Self Report Measures

I. Measuring Self-Report Variables

A. The Survey Research Method

- Many participants who are (should be) randomly selected.

- Provide Self reports of Behaviors, Feelings (current emotional states), Attitudes, and/or Personality Constructs

1. Methodological Validity

- Internal Validity
 - Low (this is correlational data)
- External Validity
 - Generalizability to Population = High (potentially)
 - Generalizability to other Settings
 - Mundane realism = Low
 - Experimental realism = Low
 - Ecological Validity = should be high

2. Strengths -

- More generalizable to the population as whole (external validity)
- though, difficult and expensive to collect truly random samples
- often use stratified samples (sample several different smaller populations).

3. Problems with Survey Research

- a. Information is less detailed than Observational and Case Study methods
- b. It is correlational research = Low Internal Validity (cause/effect relationships can not be established)
- c. Response style(set)
 - Response Acquiescence - saying yes
 - Response Deviation - saying no
 - Social Desirability- not wanting to look bad or be uncooperative. (Crown Marlow Scale)
 - Self Deceptive Positivity
 - Impression Management

d. Volunteer Problems:

- people who want to participate (agree to participate) in research may be different from people who refuse.
 - more intelligent
 - better educated
 - more cooperative
 - better adjusted
 - desire more social approval

B. Psychometric Scaling & Self-Report Methods:

1. Psychometric Scaling: Definition

- developed in early portion of this century.
- developed to measure mental phenomenon that have no clearly identifiable external cause, but the experience can be reported by the individual.
- (e.g. feelings, beliefs, Intelligence and other mental processes).



2. **Summative Scales:** Measures of attitudes and/or personality dimensions based on ratings of a series of statements selected to represent a specific topic.

- Typically these scales use **Numerical Rating Scales** on a 5 or 7 pt scales

Agreement:

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	Disagree	Neither Agree nor disagree	Agree	Strongly Agree

Endorsement:

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Very Undescriptive of me	Undescriptive of me	Somewhat Undescriptive of me	Neither Descriptive nor Undescriptive	Somewhat Descriptive of me	Descriptive of me	Very Descriptive of me

The response numbers from all the items comprising a scale are summed and averaged.

Example: Bloom County “No Freaks”

Measure of Freakiness

1. I like limber Eskimos.		
2. It is ok for a person to have one or many tattoos.		
3. I would associate with people who gargle Windex		
4. I would consider purchasing a Barry Manilow inflatable doll.		
5. It would not bother me to sleep with large tomatoes		
6. I feel that snake wrestling in Jell-O pudding is an acceptable social activity.		

- Ideally, half the items are worded in the affirmative (agreeing indicates more positive attitudes toward freaks) and half the items are worded in the negative (disagreeing indicates more positive attitudes toward freaks).
- Reverse Worded items must be reverse scored.

Balanced Example: Reverse scored items =

8 – score.

1s become 7s, 2s become 6s, 3s become 4s, etc.

C. **Measurement Reliability**: The consistency of a measure.

1. Consistency Across Time

Test-Retest:

- Give test at two points in time.
- High positive correlations indicates good Test-Retest Reliability (usually about .50 over 3 months is average).
- Not all psychological phenomena are expected to be stable.
 - Traits = e.g., personality variables
 - States = e.g., transient moods.
- Problem of Practice Effects / Testing Effects
 - Increase in performance due to item familiarity (abilities, knowledge, and performance tests).
 - Solution = Parallel Forms

2. Consistency Across Items

Internal Consistency: (Cronbach's Alpha)

- correlate all items in a scale with each other (item intercorrelations).
- All items should be responded to consistently (highly correlated).
- Adequate Consistency = (alpha) $\alpha \geq .70$
- Don't confuse Cronbach α with Type I error α .

D. Measurement Validity

- Does the measure assess what it should
- A measure must be reliable in order for it to be valid (reliability is a necessary, but not sufficient, criteria for validity)

1. Content Validity (Face Validity)

- Does a measure contain the information it is supposed to
- a - Representativeness = are all the relevant domains/dimensions/constructs represented.
 - Have experts and lay persons develop list: eyeball it
- b - Relevance = are irrelevant domains / dimensions / constructs excluded.
 - experts eyeball it

2. Construct Validity:

- Does the measure actually assess the theoretical construct that it is intended to.
- a. Convergent Validity = Is the measure correlated (associated) with other measures that it is expected to based on theory.
 - e.g.
 - Dog Liking should be associated with more general attitudes toward animals
 - Attitude toward Cheezy Poofs should be associated with other snack food attitudes.
 - Attitudes with Freakiness should be negatively associated with Disgust Sensitivity, Dogmatism, and Conservatism

b. Discriminant Validity: Is the measure uncorrelated with other measures that theory says it should not be related to.

If the measure was strongly related to an established and theoretically unrelated construct it would call into question the meaning of the new measure.

e.g.,

- Dog Liking should not be associated with Social Desirability, Political Attitudes (Conservatism vs. Liberalness), or Dogmatism.
- Attitude toward Cheezy Poofs should not be associated with social desirability or disgust sensitivity.
- Freakiness should not be associated with extroversion, agreeableness, or neuroticism.