I. Operationalization

- The processes or steps taken to measure an abstract concept in terms of its observable, measurable characteristics.

II. Types of Operational Definitions

A. Experimental Operational Definitions:
   Processes used to manipulate a variable

B. Measured Operational Definitions:
   Processes used to carefully measure or observe a variable

III. Methods of Operationalization

A. Observations
   s: often more accurate than reports
   s: many behaviors difficult to observe
   w: gives no meaning or interpretation

B. Self-Reports
   s: good for assessing attitudes and beliefs
   s: good for assessing “unobservable” behaviors
   w: social desirability

C. Direct Classification:
   Definitions describing how to decide the presence or absence of a variable
C. Observer Ratings
s: may help with problems of social desirability
w: observer bias

D. Triangulation of Measurement
Measuring the same concept using multiple methods

A Side-Bar...
• Scales are sets of questions (indicators) used to measure a given concept or variable.
• Sometimes you can operationalize with a “single indicator”
  ~ Are you a (1) Female or (2) Male?
• Sometimes you need to use “multiple indicators” or a scale to operationalize all facets of a concept.

Community Involvement Scale
How often do you get together with people who know what’s going on in San Francisco?
Never (1) Seldom (2) Sometimes (3) Frequently (4)
How often do you have ideas for improving things in San Francisco?
How often do you work to bring about change in San Francisco?
How often do you keep up with the local news in San Francisco?

IV. Scale Construction and Levels of Measurement
A. Nominal Scales
1. Scales that classify variables into different categories.
2. Requirements of Nominal Scales
   Categories must be mutually exclusive
   Categories must be exhaustive

What is your primary way of getting university news?
________ Newspapers
________ Television
________ Friends

What is your racial or ethnic category?
________ African American
________ Caucasian
________ Native American
B. Ordinal Scales

Scales that classify variables into categories, and the categories can be ordered. (Ranks)

_____ Blues
_____ Jazz
_____ Classical
_____ Easy-Listening
_____ Country

Ordinal scales cannot report the distance between categories -- only the order.

C. Interval Scales

1. Scales that...
   a. Classify variables into categories
   b. Have categories that can be ordered
   c. Have standardized distances between points on the scale
   d. Have an arbitrary zero point.

2. Likert Scales
   a. Agree/Disagree Responses to Statements

   1........2........3........4........5
   Strongly Disagree  Neither Agree  Strongly Agree

b. Likert-type Scales

   1........2........3........4........5
   Not at all  Somewhat  Very Much

   1........2........3........4........5
   Never  Sometimes  Often

   -2........-1........0........1........2
   Never  Sometimes  Often

c. Considerations
   i. Neutral Points
   ii. Polarity Rotation:

   Alternating the direction of the wording

   Helps to avoid Response Sets
3. Semantic Differential Scales

a. Scales with antonyms at the opposite ends of the scale.

| Good: _______ | _______ | _______ | _______ | _______ | Bad: _______ | _______ | _______ | _______ | _______ |
| Unfair: _______ | _______ | _______ | _______ | _______ | Fair: _______ | _______ | _______ | _______ | _______ |
| Active: _______ | _______ | _______ | _______ | _______ | Passive: _______ | _______ | _______ | _______ | _______ |
| Bad: _______ | _______ | _______ | _______ | _______ | Good: _______ | _______ | _______ | _______ | _______ |

b. Considerations

i. True Antonyms?

ii. Neutral Points

iii. Polarity Rotation: Alternating the direction of the wording.

4. Thurstone Scales

Statements to which respondents agree or disagree. Points for each response are determined by pre-ratings.

D. Ratio Scales

Scales that...

a. Classify variables into categories

b. Have categories that can be ordered

c. Have standardized distances between points on the scale

d. Have a true zero point.

Reliability and Validity
I. Measurement Validity

The extent to which a measure actually measures what it claims to measure.

A. Face Validity: The extent to which a measure appears to be a reasonable measure of some concept on the face of it.

B. Expert Jury Validity: A group of experts examine the face validity of some measure.

C. Criterion-related Validity: The degree to which some measurement relates to some other (external) measure or criterion.
   1. Concurrent validity: How well a new measure correlates with a previously validated measure of the same concept.
   2. Predictive validity: How well a measure predicts the occurrence of a concept.

D. Construct Validity: The degree to which a measure relates to other variables in a way that is predicted by theory.

II. Reliability

The stability and consistency of a measure.

A. Test-Retest:
   - The same people are given the measure at least two points in time.

B. Alternate Forms:
   - Alternate forms of the same measure from a common pool of items.

Questions
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Form A      = Form B
1 14 6 4 9 10 8              2 13 5 3 11 7 12
C. Split-half method.
- After a test has been given, dividing up the items within the test into two groups.
- Scores on the two groups of items should be similar

<table>
<thead>
<tr>
<th>Questions</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sam’s score on half of the items</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sam’s score on the other half of the items</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

D. Intercoder Reliability. The extent to which at least two observers agree in their ratings of the same event or object.

E. Internal Consistency:
- Compares answers to each question to answers on every other question.

<table>
<thead>
<tr>
<th>Exam Questions</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exam Questions</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>
Exam Questions

F. Statistical Methods:
1. Types of Statistics
   - Scott’s pi, Cohen’s kappa: Intercoder
   - K-R 20: Internal Consistency for correct/incorrect responses (exam)
   - Cronbach’s alpha: Internal Consistency for items that are NOT “correct/incorrect” (attitudes)

2. Reliability Coefficients: Numbers that indicate the level of reliability of some measure.
   - Ranges from 0 to 1
   - .9 or higher is very strong
   - Below .6 should not be analyzed

III. Relationships Between Reliability and Validity

Scale 1
Scale 2
Scale 1
Scale 2
A. In order for a measure to be valid, it must also be reliable.

B. A measure that is reliable is not necessarily valid.

C. Reliability is necessary but not sufficient for validity.