What is Scaling?

- Young's (1984) definition
- Scaling is the branch of measurement that involves the construction of an instrument that associates qualitative constructs with quantitative metric units.

Scaling and the Rule

- Scaling is the method for determining the rule
- A response scale (e.g., true-false, 1-5 agree scale, etc...) is not scaling.
  - A response scale is just a way to collect data from people using an instrument.
  - The just attaches a response to an object or statement
  - No consistent rule is used
- Scaling involves procedures that you do independent of the respondent so that you can come up with a numerical value for the object.

Major Scaling Methods

- Thurstone or Equal Appearing Intervals
- Likert or Summative Scaling
- Guttman or Cumulative Scaling
- Mokken or Probabilistic Cumulative Scale
- Multidimensional scaling

Thurstone Scaling

- Actually created 3 different ways of producing a unidimensional scale
  - Equal-Appearing intervals (easiest)
  - Successive intervals
  - Paired comparisons

Example adapted from
(https://www.socianalysis.com/thurstone.html)
**Equal Appearing Intervals**

- **Steps**
  1. State focus of scaling attempt and definition of the focal variable (e.g., attitudes toward individuals with AIDS).
  2. Generate potential scale items used to elicit attitudes:
     - Large set of initial items
     - Similar wording

**Item examples**

- People get AIDS by engaging in immoral behavior
- AIDS is the wrath of God
- AIDS is an epidemic that affects us all
- People with AIDS are bad
- People with AIDS are real people
- I treat everyone the same, regardless of whether or not they have AIDS
- AIDS costs the public too much
- AIDS is a death sentence
- AIDS is preventable
- People who contract AIDS deserve it
- AIDS doesn’t have a preference, anyone can get it.
- AIDS is the worst thing that could happen to you.
- AIDS is helps control the population.
- People with AIDS do not deserve our help
- AIDS will never happen to me
- People with AIDS deserve what they got
- AIDS distracts from other diseases that deserve our attention more
- People with AIDS deserve our understanding, but not necessarily special treatment
- People with AIDS should be quarantined to protect the rest of society
- Everyone affected with AIDS deserves it due to their lifestyle
- Someone with AIDS could be just like me
- People infected with AIDS did not have safe sex
- People with AIDS should be treated just like everybody else.

**Equal Appearing Intervals**

- **Steps (cont)**
  1. Participants rate the items in terms of the favorability of the attitude toward individuals with AIDS.

**Item stats**

<table>
<thead>
<tr>
<th>Statement</th>
<th>Median</th>
<th>Interquartile Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

**Final step**

- **Select final scale items**
  - Select statements that are at equal intervals across the range of medians (e.g., one for each rating level, 1-13)
  - Try to select the statement that has the least amount of variability across judges (smallest Interquartile Range)
For instance:

- People with AIDS are like my parents (6)
- Because AIDS is preventable, we should focus our resources on prevention instead of curing (5)
- People with AIDS deserve what they got. (1)
- AIDS affects us all (10)
- People with AIDS should be treated just like everybody else. (11)
- AIDS will never happen to me. (3)
- It's easy to get AIDS (5)
- AIDS doesn't have a preference, anyone can get it (9)
- AIDS is a disease that anyone can get if they are not careful (9)
- If you have AIDS, you can still lead a normal life (8)
- AIDS is good because it helps control the population. (2)
- I can't get AIDS if I'm in a monogamous relationship. (4)

Thurstone's Paired Comparisons

- Very similar process
- Difference is in how the initial ratings are collected
- Method of paired comparisons requires each judge to make a judgement about each pair of statements (Uuugh!)
  - 57 statements require 1,596 unique comparisons

Likert Scales

- Select the items
- Look for items with good rater agreement
- Compute the intercorrelations between all pairs of items, based on the ratings of the judges
  - Discard items that have a low item-total correlation

Details...

- The value in parentheses after each statement is its scale value.
- Random order of the statements with respect to scale values.
- Notice that there is no item with scale value of 7 and that we have two with values of 5 and 9
  (one of these pairs will average out to 6.5).
- You now have a scale -- a yardstick you can use for measuring attitudes towards people with AIDS.
- A person's total scale score is the average of the scale scores of all the items that person agreed with.

Administer the scale

- Each respondent is asked to rate each item on some response scale.
  - strongly disagree
  - disagree
  - undecided
  - agree
  - strongly agree
- Can also use a forced-choice response scale with an even number of responses (no middle neutral)
- Total score for the respondent on the scale is the sum of their ratings for all of the items (this is why this is sometimes called a "summated" scale).
Guttman Scales
- Also known as cumulative scaling or scalogram analysis.
- Goal: a set of items or statements such that a respondent who agrees with any specific question in the list will also agree with all previous statements.

Guttman
- Same as before, generate the items.
- Judges rate the statements or items in terms of how favorable they are to the concept of immigration (Yes=favorable; No=unfavorable).
- The key to Guttman scaling is in the analysis.

Guttman Scaling
- Construct a matrix that shows the responses of all the judges on all of the items.
- Sort this matrix so that respondents who agree with more statements are listed at the top and those agreeing with fewer are at the bottom.
- The resulting scale should be nearly cumulative.

Guttman Matrix

<table>
<thead>
<tr>
<th>Item</th>
<th>Item 1</th>
<th>Item 2</th>
<th>Item 3</th>
<th>Item 4</th>
<th>Item 5</th>
<th>Item 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 2</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 3</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 4</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Item 5</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Item 6</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

Exceptions are circled.

Guttman Scaling
- Can do this for a small number of items.
- If there are many items, then a scalogram analysis is used to do the same thing.
- Determines the subsets of items from our pool that best approximate the cumulative property.
- Many analyses used to determine how well the items approximate a true cumulative scale.
Guttman Scale

- **INSTRUCTIONS:** Place a check next to each statement you agree with.
  - _____ I would permit a child of mine to marry an immigrant.
  - _____ I believe that this country should allow more immigrants.
  - _____ I would be comfortable if a new immigrant moved next door.
  - _____ I would be comfortable if a new immigrant moved into my community.
  - _____ It would be fine with me if new immigrants moved onto my block.
  - _____ I would be comfortable if my child dated a new immigrant.

Each item has a scale value associated with it (obtained from the scalogram analysis). A respondent's scale score is the sum the scale values of every item they agree with.

Mokken Scaling

- Just like Guttman scales
- Individual's scale score is the rank of the highest item endorsed or the total number of items endorsed
- Key difference
  - Mokken scales are probabilistic
  - A respondent answering an item positively will have a large probability of answering a less difficult item in a positive way as well

Multidimensional Scaling

- Many interesting concepts are not unidimensional (e.g., attitude toward immigration)
- MDS maps the distances between points in a high dimensional space into a lower dimensional space without too much loss of information.
- Almost always 2-D

The “points” that are represented in multidimensional space can be just about anything.

These objects might be people, in which case MDS can identify clusters of people who are “close” versus “distant” in some real or psychological sense.

The objects might be physical objects (e.g., cities, cars, books) or psychological or cultural objects (e.g., personality traits, religions, political parties).

The only requirement is that some basis exist for rating or ranking the objects in terms of similarity.

As long as the “distance” between the objects can be assessed in some fashion, MDS can be used to find the lowest dimensional space that still adequately captures the distances between objects.

Once the number of dimensions is identified, must identify the meaning of the dimensions.
Multidimensional Scaling

- Basic data representation in MDS is a dissimilarity matrix that shows the distances between every possible pair of objects.
- The goal of MDS is to faithfully represent these distances with the lowest possible dimensional space.
- We seek a reference system that can capture the basic information in a data set—in this case, a data set characterized by distances between objects.

This table lists the distances between European cities. A multidimensional scaling of these data should be able to recover the two dimensions (North-South x East-West) that we know must underlie the spatial relations among the cities.

<table>
<thead>
<tr>
<th></th>
<th>Athens</th>
<th>Berlin</th>
<th>Dublin</th>
<th>London</th>
<th>Madrid</th>
<th>Paris</th>
<th>Rome</th>
<th>Warsaw</th>
</tr>
</thead>
<tbody>
<tr>
<td>Athens</td>
<td>0</td>
<td>1119</td>
<td>1777</td>
<td>1486</td>
<td>1475</td>
<td>1303</td>
<td>646</td>
<td>1013</td>
</tr>
<tr>
<td>Berlin</td>
<td>1119</td>
<td>0</td>
<td>817</td>
<td>577</td>
<td>1159</td>
<td>545</td>
<td>736</td>
<td>327</td>
</tr>
<tr>
<td>Dublin</td>
<td>1777</td>
<td>817</td>
<td>0</td>
<td>291</td>
<td>906</td>
<td>489</td>
<td>1182</td>
<td>1135</td>
</tr>
<tr>
<td>London</td>
<td>1486</td>
<td>577</td>
<td>291</td>
<td>0</td>
<td>783</td>
<td>213</td>
<td>897</td>
<td>904</td>
</tr>
<tr>
<td>Madrid</td>
<td>1475</td>
<td>1159</td>
<td>906</td>
<td>783</td>
<td>0</td>
<td>652</td>
<td>856</td>
<td>1483</td>
</tr>
<tr>
<td>Paris</td>
<td>1303</td>
<td>545</td>
<td>489</td>
<td>213</td>
<td>652</td>
<td>0</td>
<td>694</td>
<td>859</td>
</tr>
<tr>
<td>Rome</td>
<td>646</td>
<td>736</td>
<td>1182</td>
<td>897</td>
<td>856</td>
<td>694</td>
<td>0</td>
<td>839</td>
</tr>
<tr>
<td>Warsaw</td>
<td>1013</td>
<td>327</td>
<td>1135</td>
<td>904</td>
<td>1483</td>
<td>859</td>
<td>839</td>
<td>0</td>
</tr>
</tbody>
</table>

In the cities data, the meaning is quite clear.

The dimensions refer to the North-South x East-West surface area across which the cities are dispersed.

Expect MDS to faithfully recreate the map relations among the cities.

Common to evaluate the fit of a model using a "stress" statistic:

<table>
<thead>
<tr>
<th>Stress</th>
<th>Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>.20</td>
<td>Poor</td>
</tr>
<tr>
<td>.10</td>
<td>Fair</td>
</tr>
<tr>
<td>.05</td>
<td>Good</td>
</tr>
<tr>
<td>.02</td>
<td>Excellent</td>
</tr>
</tbody>
</table>

The stress for the one-dimensional model of the cities data is .31, clearly a poor fit.

The analysis proceeds by trying higher dimensional models, seeking the minimum number of dimensions necessary to get an acceptable fit.

A 2-D model fits very well (stress value = .009) indicating an exceptional fit. This is no great surprise for these data.
Multidimensional Scaling

- Rarely do social sciences have metric data on which to compare the objects. More typically, the “distances” are on psychological dimensions for which ordinal relations are the best scaling possible.
- Use similarity ratings instead. Similarity is a slippery slope.
- Can be very time consuming because they require the rating of each pair of objects.

Shepard & Chipman: Shapes of States

- Illinois
- Nevada
- Minn.
- Missouri
- Alabama
- Oregon
- Colo.
- Nebraska
- Maine
- La.
- Fla.
- S.C.
- W.Va.
- Idaho
- Oklahoma

- Vertical/Irregular
- Horizontal/Rectangular
- Smallish/Wiggly

Beer MDS

- Dim I is “Flavor”
- Dim II is “Image”
- BUDWEISER
- AMSTEL LIGHT
- HEINEKEN
- GUINESS
- MILLER LITE

- Inexpensive
- High Quality
- High Alcohol
- Good Taste
- Good Image