Measurement and Scaling: Noncomparative Scaling Techniques

Agenda

1) Overview
2) Noncomparative Scaling Techniques
3) Continuous Rating Scale
4) Itemized Rating Scale
   i. Likert Scale
   ii. Semantic Differential Scale
   iii. Stapel Scale
5) Noncomparative Itemized Rating Scale Decisions
   i. Number of Scale Categories
   ii. Balanced vs. Unbalanced Scales
   iii. Odd or Even Number of Categories
   iv. Forced vs. Non-forced Scales
   v. Nature and Degree of Verbal Description
   vi. Physical Form or Configuration
6) Multi-item Scales

Noncomparative Scaling Techniques

- Respondents evaluate only one object at a time, and for this reason noncomparative scales are often referred to as monadic scales.
- Noncomparative techniques consist of continuous and itemized rating scales.
Basic non-comparative scales

<table>
<thead>
<tr>
<th>Scale</th>
<th>Basic characteristics</th>
<th>Examples</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous rating scale</td>
<td>Place a mark on a continuous line</td>
<td>Reaction to TV commercials</td>
<td>Easy to construct</td>
<td>Scoring can be cumbersome unless computerised</td>
</tr>
</tbody>
</table>

Itemised rating scales

<table>
<thead>
<tr>
<th>Scale</th>
<th>Basic characteristics</th>
<th>Examples</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likert scale</td>
<td>Degree of agreement on a 1 (strongly disagree) to 5 (strongly agree) scale</td>
<td>Measurement of attitudes</td>
<td>Easy to construct, administer and understand</td>
<td>More time-consuming</td>
</tr>
<tr>
<td>Semantic differential</td>
<td>Seven-point scale with bipolar labels</td>
<td>Brand, product, and company images</td>
<td>Versatile</td>
<td>Controversy as to whether the data are interval</td>
</tr>
<tr>
<td>Stapel scale</td>
<td>Unipolar ten-point scale, –5 to +5, without a neutral point (zero)</td>
<td>Measurement of attitudes and images</td>
<td>Easy to construct, administer over phone</td>
<td>Confusing and difficult to apply</td>
</tr>
</tbody>
</table>

RATE: Rapid Analysis and Testing Environment

A relatively new research tool, the perception analyzer, provides continuous measurement of "gut reaction." A group of up to 400 respondents is presented with TV or radio spots or advertising copy. The measuring device consists of a dial that contains a 100-point range. Each participant is given a dial and instructed to continuously record his or her reaction to the material being tested.

As the respondents turn the dials, the information is fed to a computer, which tabulates second-by-second response profiles. As the results are recorded by the computer, they are superimposed on a video screen, enabling the researcher to view the respondents’ scores immediately. The responses are also stored in a permanent data file for use in further analysis. The response scores can be broken down by categories, such as age, income, sex, or product usage.
Continuous Rating Scale

Respondents rate the objects by placing a mark at the appropriate position on a line that runs from one extreme of the criterion variable to the other. The form of the continuous scale may vary considerably.

How would you rate Sears as a department store?

Version 1
Probably the worst ............................................. -I............................................. Probably the best

Version 2
Probably the worst ..................................... -I.................................................. Probably the best

Version 3
Very bad - Neither good nor bad - Very good

Itemized Rating Scales

• The respondents are provided with a scale that has a number or brief description associated with each category.
• The categories are ordered in terms of scale position, and the respondents are required to select the specified category that best describes the object being rated.
• The commonly used itemized rating scales are the Likert, semantic differential, and Stapel scales.
The *Likert scale* requires the respondents to indicate a degree of agreement or disagreement with each of a series of statements about the stimulus objects.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. FNAC sells high quality products.</td>
<td>1</td>
<td>2X</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2. FNAC has poor in-store service.</td>
<td>1</td>
<td>2X</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3. I like to shop at FNAC.</td>
<td>1</td>
<td>2X</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

- The analysis can be conducted on an item-by-item basis (profile analysis), or a total (summated) score can be calculated.
- When arriving at a total score, the categories assigned to the negative statements by the respondents should be scored by reversing the scale.
**Examples**

**NEW FRAGRANCE SURVEY**

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>I usually buy may fragrances in a department store</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>I usually buy may fragrances in a supermarket</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>I only buy one brand of fragrance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>I would be interested in a new fragrance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>I frequently try new fragrances</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**JOB SATISFACTION**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>I get a feeling of accomplishment from the work I am doing</td>
</tr>
</tbody>
</table>

---

**Semantic Differential Scale**

The **semantic differential** is a seven-point rating scale with end points associated with bipolar labels that have semantic meaning.

- **SEARS IS:**
  - Powerful
  - Reliable
  - Modern
- **SEARS IS NOT:**
  - Weak
  - Unreliable
  - Old-fashioned

- The negative adjective or phrase sometimes appears at the left side of the scale and sometimes at the right.
- This controls the tendency of some respondents, particularly those with very positive or very negative attitudes, to mark the right- or left-hand sides without reading the labels.
- Individual items on a semantic differential scale may be scored on either a -3 to +3 or a 1 to 7 scale.
- The resulting data are commonly analyzed through Profile analysis.
Instructions
This part of the study measures what certain banks mean to you. You judge them on a series of descriptive scales bounded at each end by one of two bipolar adjectives. Please mark the blank that best indicates how accurately each adjective describes what the bank means to you.

Form
Dresdner Bank is:

- Powerful: __________:X: __________: Weak
- Unreliable: __________:X: __________: Reliable
- Modern: __________:X: __________: Old-fashioned
- Hi-tech: __________:X: __________: Low-tech
- Careful: __________:X: __________: Careless


1) Rugged __________: Delicate
2) Excitable __________: Calm
3) Uncomfortable __________: Comfortable
4) Dominating __________: Submissive
5) Thrifty __________: Indulgent
6) Pleasant __________: Unpleasant
7) Contemporary __________: Obsolete
8) Organized __________: Unorganized
9) Rational __________: Emotional
10) Youthful __________: Mature
11) Formal __________: Informal
12) Orthodox __________: Liberal
13) Complex __________: Simple
14) Colorless __________: Colorful
15) Modest __________: Vain
The Stapel scale is a unipolar rating scale with ten categories numbered from -5 to +5, without a neutral point (zero). This scale is usually presented vertically.

SEARS

+5  +5
+4  +4
+3  +3
+2  +2X
+1  +1

HIGH QUALITY  POOR SERVICE

-1  -1
-2  -2
-3  -3
-4X  -4
-5  -5

The data obtained by using a Stapel scale can be analyzed in the same way as semantic differential data.
## Summary of Itemized Scale Decisions

<table>
<thead>
<tr>
<th>Decision</th>
<th>Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Number of categories</td>
<td>Although there is no single, optimal number, traditional guidelines suggest that there should be between five and nine categories.</td>
</tr>
<tr>
<td>2) Balanced vs. unbalanced</td>
<td>In general, the scale should be balanced to obtain objective data.</td>
</tr>
<tr>
<td>3) Odd/even no. of categories</td>
<td>If a neutral or indifferent scale response is possible from at least some of the respondents, an odd number of categories should be used.</td>
</tr>
<tr>
<td>4) Forced vs. non-forced</td>
<td>In situations where the respondents are expected to have no opinion, the accuracy of the data may be improved by a non-forced scale.</td>
</tr>
<tr>
<td>5) Verbal description</td>
<td>An argument can be made for labeling all or many scale categories. The category descriptions should be located as close to the response categories as possible.</td>
</tr>
<tr>
<td>6) Physical form</td>
<td>A number of options should be tried and the best selected.</td>
</tr>
</tbody>
</table>

### Balanced and unbalanced scales

#### Balanced scale

*Clinique moisturiser for men is:*

- Extremely good
- Very good
- Good
- Bad
- Very Bad
- Extremely bad

#### Unbalanced scale

*Clinique moisturiser for men is:*

- Extremely good
- Very good
- Good
- Somewhat good
- Bad
- Very bad
Rating Scale Configurations

A variety of scale configurations may be employed to measure the gentleness of Cheer detergent. Some examples include:

Cheer detergent is:

1) Very harsh  ---  ---  ---  ---  ---  ---  Very gentle
2) Very harsh  1  2  3  4  5  6  7  Very gentle
3) Very harsh
   .
   . Neither harsh nor gentle
   .
   . Very gentle
4)

<table>
<thead>
<tr>
<th>Very harsh</th>
<th>Harsh</th>
<th>Somewhat Harsh</th>
<th>Neither harsh nor gentle</th>
<th>Somewhat gentle</th>
<th>Gentle</th>
<th>Very gentle</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>0</td>
<td>+1</td>
<td>+2</td>
<td>+3</td>
</tr>
</tbody>
</table>

Rating scale configurations

A variety of scale configurations may be employed to measure the gentleness of Clinique Face Scrub for Men.

Some examples include:

Clinique Face Scrub for Men is:

1. Very harsh  ---  ---  ---  ---  ---  ---  Very gentle
2. Very harsh  1  2  3  4  5  6  7  Very gentle
3. Very harsh
   .
   . Neither harsh nor gentle
   .
   . Very gentle
4.

<table>
<thead>
<tr>
<th>Very harsh</th>
<th>Harsh</th>
<th>Somewhat Harsh</th>
<th>Neither harsh nor gentle</th>
<th>Somewhat gentle</th>
<th>Gentle</th>
<th>Very gentle</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>0</td>
<td>+1</td>
<td>+2</td>
<td>+3</td>
</tr>
</tbody>
</table>
Some unique rating scale configurations

**Thermometer scale**

**Instructions**
Please indicate how much you like McDonald's 'Big Macs' by colouring in the thermometer with your red pen. Start at the bottom and colour up to the temperature that shows how much you prefer McDonalds 'Big Macs'.

### Form

<table>
<thead>
<tr>
<th>Like very much</th>
<th>Dislike very much</th>
</tr>
</thead>
<tbody>
<tr>
<td>100°</td>
<td>0°</td>
</tr>
<tr>
<td>75°</td>
<td></td>
</tr>
<tr>
<td>50°</td>
<td></td>
</tr>
<tr>
<td>25°</td>
<td></td>
</tr>
</tbody>
</table>

Some unique rating scale configurations

**Smiling face scale**

**Instructions**
Please tell me how much you like Barbie Doll by pointing to the face that best shows how much you like it. If you did not like the Barbie Doll at all, you would point to Face 1. If you liked it very much, you would point to Face 5. Now tell me, how much did you like the Barbie Doll?

### Form

1.  
2.  
3.  
4.  
5.
Development of a Multi-item Scale

Figure 9.4

1. Develop Theory
2. Generate Initial Pool of Items: Theory, Secondary Data, and Qualitative Research
3. Select a Reduced Set of Items Based on Qualitative Judgement
4. Collect Data from a Large Pretest Sample
5. Statistical Analysis
6. Develop Purified Scale
7. Collect More Data from a Different Sample
8. Evaluate Scale Reliability, Validity, and Generalizability
9. Final Scale

Scale Evaluation

- Scale Evaluation
  - Reliability
  - Validity
  - Generalizability

Reliability:
- Test/Retest
- Alternative Forms
- Internal Consistency

Validity:
- Content
- Criterion
- Construct

Generalizability:
- Convergent
- Discriminant
- Nomological