

ELEVENTH EDITION

EXPLORING  
**Marketing** Research



Barry Babin | William Zikmund

# Chapter 10

## Measurement and Attitude Scaling



# LEARNING OUTCOMES

*After studying this chapter, you should*

1. Determine what things need to be measured to address a research question
2. Distinguish levels of scale measurement
3. Create an index or composite measure
4. Assess scale reliability and validity
5. Understand why the concept of attitude is so important in marketing research
6. Design a scale that measures an attitudinal concept
7. Implement a multi-attribute model

# Introduction

- Anyone who has ever followed a recipe knows the importance of good measurement
  - Just as in the culinary (cooking) arts, researchers can measure business and marketing in more than one way
  - Researchers often may have to use imperfect measurement devices
  - Measure a concept poorly and the “recipe” is a likely disaster
  - Only in this case, the “recipe” is usually an important business decision poorly made instead of a ruined dish

# What Needs to Be Measured?

- Business and marketing concepts can often be measured in more than one way
  - Before the measurement process can be defined, researchers have to decide exactly what it is that needs to be produced
- The decision statement, corresponding research questions, and research hypotheses can be used to decide what concepts need to be measured

# What Needs to Be Measured? (cont'd.)

- Measurement is the process of describing some property of a phenomenon of interest usually by assigning numbers in a reliable and valid way
  - When numbers are used, the researcher must have a rule for assigning a number to an observation in a way that provides an accurate description
  - All measurement systems present the potential for error

# Concepts

- Generalized ideas that represent something of meaning
  - Concepts such as age, sex, and number of children are relatively concrete properties and present few problems in either definition or measurement
  - Concepts such as brand loyalty, personality, and so on, are more abstract and are more difficult to both define and measure

# Operational Definitions

- Researchers measure concepts through a process known as operationalization
  - A process that involves identifying scales that correspond to variance in the concept
- Scales provide:
  - A range of values that correspond to different values in the concept being measured
  - Correspondence rules that indicate that a certain value on a scale corresponds to some true value of a concept, hopefully in a truthful way

# Variables

- Researchers use variance in concepts to make diagnoses
- Variables capture different concept values
- Scales capture variance in concepts and as such, the scales provide the researcher's variables
- Once a research project is underway, there is little difference between a concept and a variable

# Constructs

- A term used for concepts that are measured with multiple variables
  - Sometimes a single variable cannot capture a concept alone
  - Using multiple variables to measure one concept can provide a more complete account of some concept than could any single variable
- Operational definitions translate conceptual definitions into measurement scales

# Levels of Scale Measurement

- Important because it determines the mathematical comparisons that are allowed
- The four levels of scale measurement
  - Nominal
  - Ordinal
  - Interval
  - Ratio

# Nominal Scales

- Represent the simplest type of scale and assign a value to an object for identification or classification purposes
  - The value can be but does not have to be a number since no quantities are being represented
  - A qualitative scale
  - Useful even though they can be considered elementary
  - Marketing researchers use nominal scales quite often

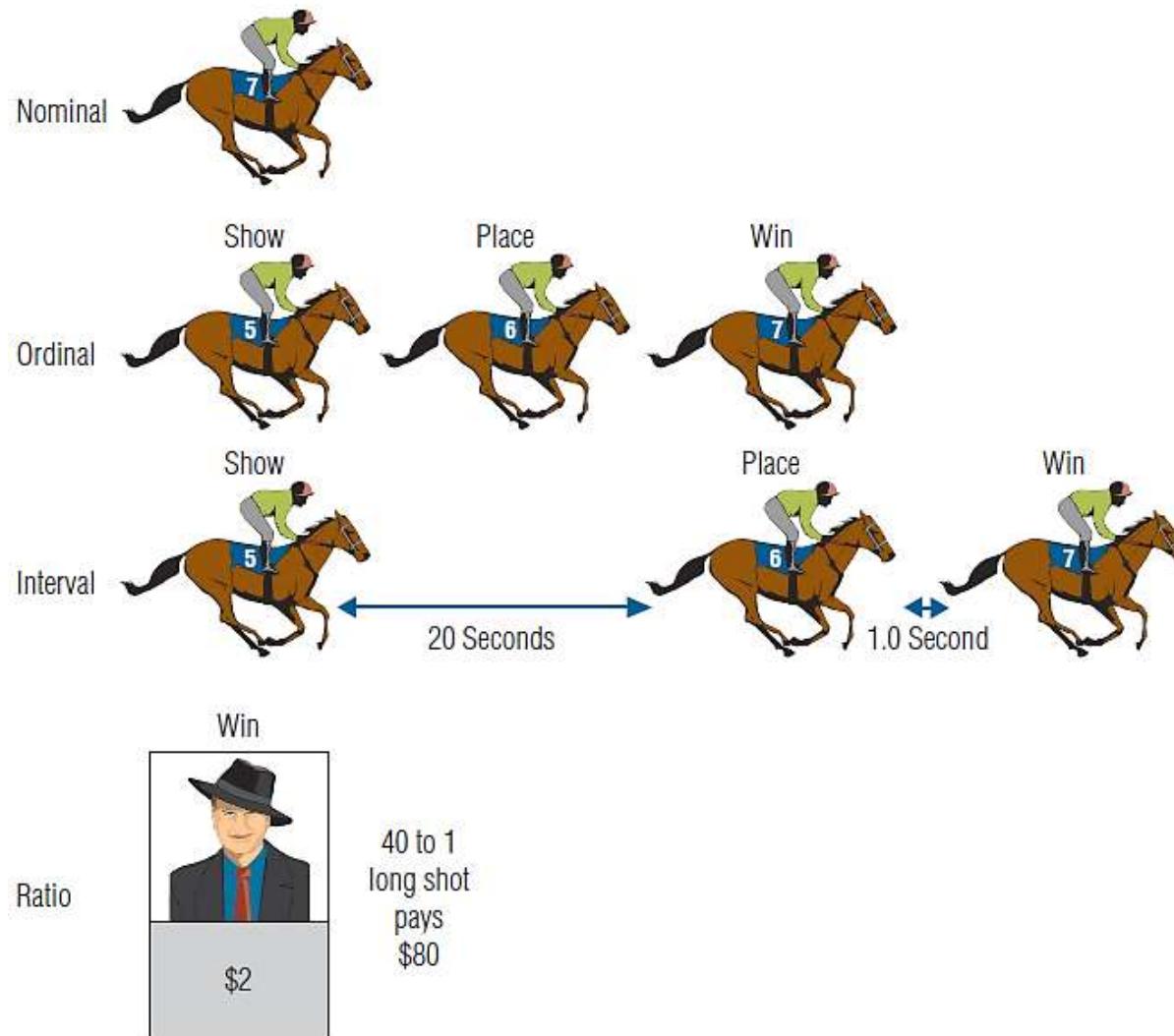
## Nominal Scales (cont'd.)

- Arbitrary in the sense that each label can be assigned to any of the categories without introducing error
- Examples
  - Uniform numbers
  - Airport terminals
  - School bus numbers

# Ordinal Scales

- Have nominal properties
- Allow things to be arranged based on how much of some concept they possess
- Include ranking scales
- Are somewhat arbitrary
- Example: “win,” “place,” and “show” in a horse race tells which horse was first, second, and third, but does not tell by how much a horse won

## EXHIBIT 10.2 Nominal, Ordinal, Interval, and Ratio Scales Provide Different Information



# Interval Scales

- Have both nominal and ordinal properties
- Capture information about differences in quantities of a concept
  - Classic example: Fahrenheit temperature scale:
    - ❖ 80 degrees is hotter than 40 degrees
    - ❖ One cannot conclude that the  $40^{\circ}$  is twice as cold as  $80^{\circ}$
    - ❖ Illustration: convert the temperatures to the Celsius scale  
 $80^{\circ} \text{ F} = 26.7^{\circ} \text{ C}$  and  $40^{\circ} \text{ F} = 4.4^{\circ} \text{ C}$ .
- Do not exactly represent some phenomenon

# Ratio Scales

- Represents the highest form of measurement
  - Have all the properties of interval scales with the additional attribute of representing absolute quantities
- Represent absolute meaning
- Provide iconic measurement

## Ratio Scales (cont'd.)

- Zero, therefore, has meaning in that it represents an absence of some concept
  - An absolute zero is a defining characteristic in determining between ratio and interval scales
  - For example, money is a way to measure economic value

## EXHIBIT 10.3 Facts About the Four Levels of Scales

Level	Examples	Numerical Operations	
Nominal	Yes – No Female – Male Buy – Did Not Buy Postal Code: _____	Counting	<ul style="list-style-type: none"><li>• Frequencies</li><li>• Mode</li></ul>
Ordinal	Rankings Choose from the Following: <ul style="list-style-type: none"><li>• Dissatisfied</li><li>• Satisfied</li><li>• Very Satisfied</li><li>• Delighted</li></ul> Indicate Your Level of Education: <ul style="list-style-type: none"><li>• HS Diploma</li><li>• Some College</li><li>• Bachelor's Degree</li><li>• Graduate Degree</li></ul>	Counting and Ordering	<ul style="list-style-type: none"><li>• Frequencies</li><li>• Mode</li><li>• Median</li><li>• Range</li></ul>

## EXHIBIT 10.3 Facts About the Four Levels of Scales (cont'd.)

Level	Examples	Numerical Operations	
Interval	100-Point Job Performance Ratings Assigned by Supervisors: 0% = Worst Performers 100% = Best Performers Temperature-Type Attitude Scales: Low Temperature = Bad Attitude High Temperature = Good Attitude	Common Arithmetic Operations	<ul style="list-style-type: none"> <li>• Mean</li> <li>• Median</li> <li>• Variance</li> <li>• Standard Deviation</li> </ul>
Ratio	Amount Purchased Salesperson Sales Volume Likelihood of performing some act: <ul style="list-style-type: none"> <li>• 0% = No Likelihood to</li> <li>• 100% = Certainty</li> </ul> Number of stores visited Time spent viewing a particular web page Number of web pages viewed	All Arithmetic Operations	<ul style="list-style-type: none"> <li>• Mean</li> <li>• Median</li> <li>• Variance</li> <li>• Standard Deviation</li> </ul>

Source: © Cengage Learning 2013.

# Mathematical and Statistical Analysis of Scales

- A researcher must know the meaning behind the numbers before useful conclusions can be drawn
  - Example: averaging the numbers used to identify school busses is meaningless

# Discrete Measures

- Take on only one of a finite number of values
  - Most often used to represent a classificatory variable
  - Do not represent intensity of measures
  - Common discrete scales include any yes-no response, matching, etc.
- Can possess ordinal properties and nominal properties
- The central tendency of a nominal or ordinal discrete measure is best captured by the mode

# Continuous Measures

- Assign values anywhere along some scale range in a place that corresponds to the intensity of some concept
- Ratio measures are continuous measures
- Marketing researchers generally treat interval scales with five or more categories as continuous

# Continuous Measures (cont'd.)

- Researchers should keep in mind the distinction between ratio and interval measures
  - Errors in judgment can be made when interval measures are treated as ratio
  - An attitude of 0 means nothing as attitude only has meaning in a relative sense
- The means and standard deviation may be calculated from continuous data
- Using the actual quantities from arithmetic operations is permissible with ratio scales

# Reliable and Valid Measures

- An attribute is a single characteristic or fundamental feature of an object, person, situation, or issue
- Attribute assessment is common in marketing research
- The measures of attributes are often combined to represent some less concrete concept

# Indexes and Composites

- Multi-item instruments for measuring a construct are called index measures, or composite measures
  - An index measure assigns a value based on how much of the concept being measured is associated with an observation
  - An index is often formed by putting several variables together
  - Composite scales also assign a value based on a mathematical derivation of multiple variables

# Computing Scale Values

- A summated scale is created by summing the response to each item making up the composite measure
- A researcher may sometimes choose to average the scores rather than summing them because the composite measure is expressed on the same scale as is the items that make it up
- Reverse coding means that the value assigned for a response is treated oppositely from the other items

## EXHIBIT 10.4 Computing a Composite Scale

Item	Strongly Disagree (SD) → Strongly Agree (SA)				
This site appears to be more trustworthy than other sites I have visited.	SD	<input checked="" type="radio"/> D	N	A	SA
My overall trust in this site is very high.	SD	D	<input checked="" type="radio"/> N	A	SA
My overall impression of the believability of the information on this site is very high.	SD	<input checked="" type="radio"/> D	N	A	SA
My overall confidence in the recommendations on this site is very high.	SD	<input checked="" type="radio"/> D	N	A	SA
The company represented in this site delivers on its promises.	SD	D	N	<input checked="" type="radio"/> A	SA
Computation: Scale Values: SD=1, D=2, N=3, A=4, SA=5					
Thus, the Trust score for this consumer is	$2 + 3 + 2 + 2 + 4 = 13$				

Source: © Cengage Learning 2013.

# Reliability

- An indicator of a measure's internal consistency
- A measure is reliable when different attempts at measuring something converge on the same result
- When the measuring process provides reproducible results, the measuring instrument is reliable

# Internal Consistency

- Represents a measure's homogeneity
- Internal consistency of a multiple-item measure can be measured by correlating scores on subsets of items making up a scale
- The split-half method of checking reliability
  - Take half the items from the scale and checking them against the results from the other half
  - The two scale halves should:
    - ❖ Correlate highly
    - ❖ Produce similar scores

# Coefficient Alpha (α)

- Represents internal consistency by computing the average of all possible split-half reliabilities for a multiple item scale
  - The coefficient demonstrates whether or not the different items converge
  - Ranges in value from 0 (no consistency) to 1 (complete consistency)
    - ❖ 0.80 - 0.96: very good reliability
    - ❖ 0.70 - 0.80: good reliability
    - ❖ 0.60 - 0.70: fair reliability
    - ❖ Below 0.60: poor reliability

# Test-Retest Reliability

- The test-retest method of determining reliability involves administering the same scale or measure to the same respondents at two separate times to test for stability
  - If the measure is stable over time, the test, administered under the same conditions each time, should obtain similar results
  - Represents a measure's repeatability

# Test-Retest Reliability (cont'd.)

- Reliability is a necessary but insufficient condition for validity
  - A reliable scale may not be valid
  - A reliable but invalid instrument will yield consistently inaccurate results

# Validity

- Good measures should be both precise (i.e., reliable) and accurate (i.e., valid).
- Validity is the accuracy of a measure or the extent to which a score truthfully represents a concept
  - Achieving validity is not a simple matter
  - Addresses the problem of whether a measure indeed measures what it is supposed to measure
  - When a measure lacks validity, any conclusions based on that measure are also likely to be faulty

# Establishing Validity

- The three approaches to establishing validity
  - Face (content) validity refers to the subjective agreement among professionals that a scale logically reflects the concept being measured
  - Criterion validity addresses the question: “Does my measure correlate with measures of similar concepts or known quantities?”
    - ❖ May be classified as either concurrent validity or predictive validity
  - Construct validity

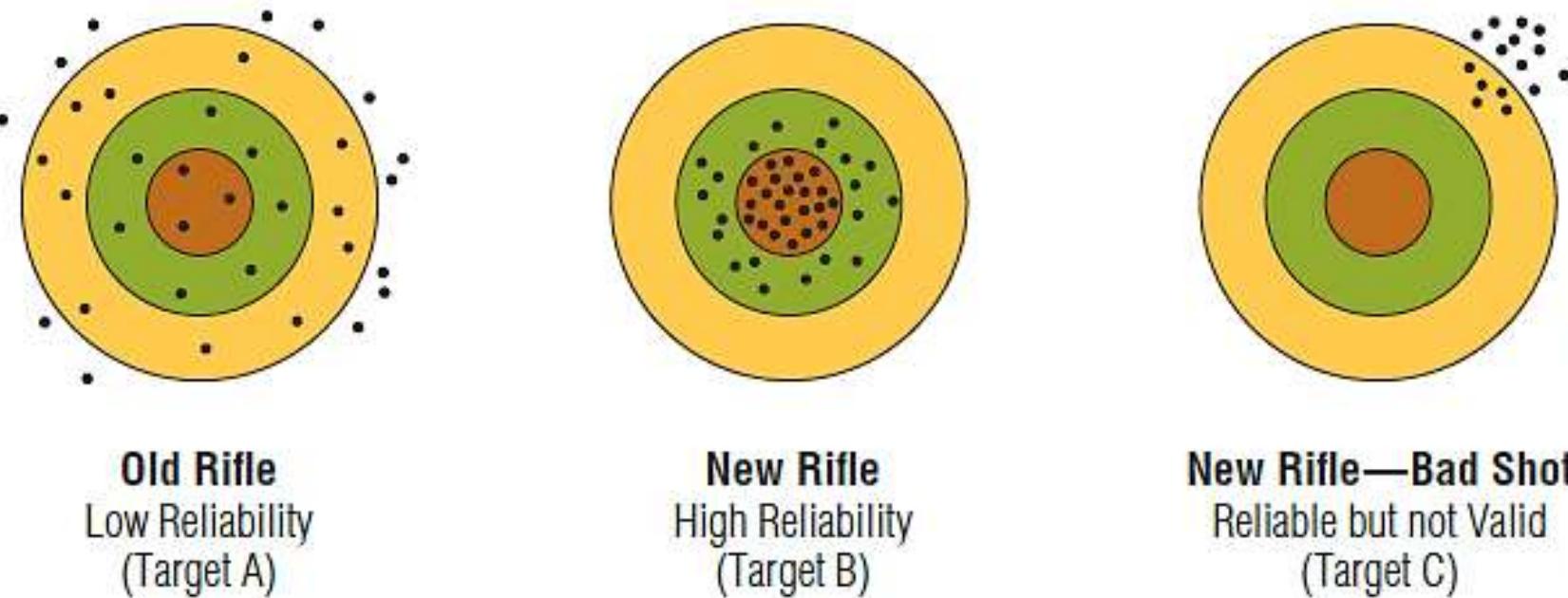
# Construct Validity

- Exists when a measure reliably measures and truthfully represents a unique concept and consists of several components:
  - Face validity
  - Criterion validity
  - Convergent validity
  - Discriminant validity
  - Fit validity

# Reliability versus Validity

- The differences between the two are illustrated with rifle targets (see Exhibit 10.5)
  - A: The shots from the older gun are scattered ► low reliability
  - B: The shots from the newer gun are closely clustered and on target ► high reliability and validity
  - C: The shots from a newer gun are closely clustered but off target ► high reliability but low validity

## EXHIBIT 10.5 Computing a Composite Scale



SOURCE: © Cengage Learning 2013.

# What Is an Attitude?

- An enduring disposition to respond consistently in a given manner to various aspects of the world, including persons, events, and objects
  - There are three components of attitude:
    - ❖ Affective component – expresses how much affinity someone has toward the relevant matter
    - ❖ Cognitive component – represents an individual's awareness of the relevant matter
    - ❖ Behavioral component – represents the action that corresponds to a certain type of attitude

# Attitude Measures and Scaling

- A remarkable variety of techniques has been devised to measure attitudes, stemming in part from lack of consensus about the exact definition of the concept
- The affective, cognitive, and behavioral components of an attitude may be measured by different means, i.e., direct to indirect, physiological to verbal, etc.

# Physiological Measures

- Research may assess the affective (emotional) components of attitudes through physiological measures
  - Examples: galvanic skin response (GSR), blood pressure, magnetic resonance imaging (MRI), and pupil dilation
- These measures provide:
  - A means of assessing affect without verbally questioning the respondent
  - A gross measure of likes or dislikes; they are not extremely sensitive to the different gradients of an attitude

# Self-Report Scales

- Involve gaining a respondent's structured response to some specific query or stimulus
- Researchers can ask respondents to rank, rate, sort, or choose one of multiple responses
  - A ranking task requires the respondent to rank order a small number of stores, brands, feelings, or objects based on overall preference or of some characteristic of the stimulus
    - ❖ Rankings provide ordinal measurement

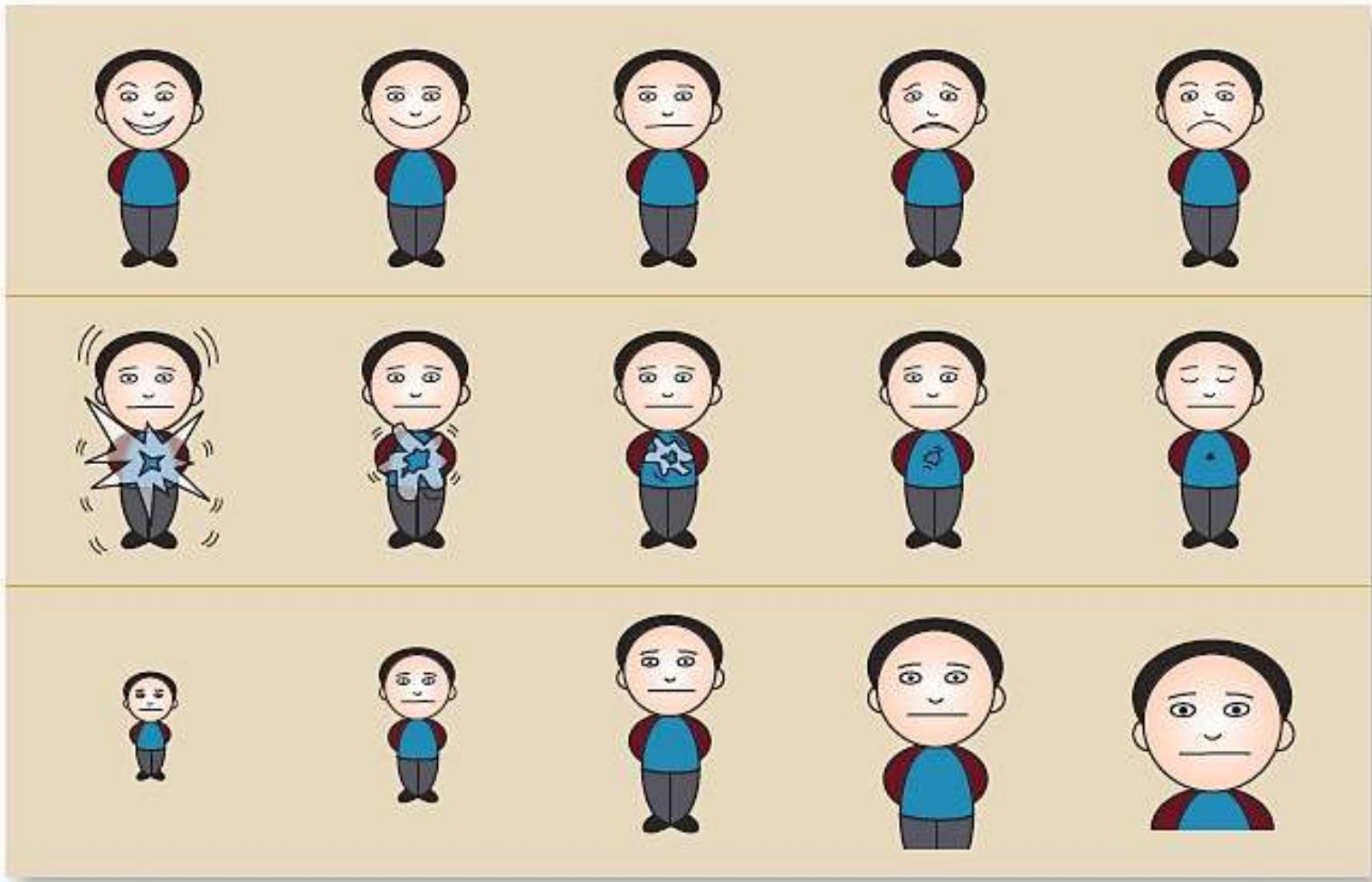
## Self-Report Scales (cont'd.)

- Rating requires the respondent to estimate the magnitude or the extent to which some characteristic exists or some choice is preferred
- A sorting task presents the respondent with several product concepts to classify the concepts
- Another type of attitude measurement requires a choice between two or more alternatives

# Measuring the Affect Component

- Exhibit 10.6 displays an approach to ask respondents to rate their affect toward an object
  - Respondents are asked to choose a “manikin” from each or three rows to show: 1) their general affective state (like or dislike); 2) the degree to which they get emotionally aroused; and 3) how much the thing being rated makes them feel small (unimportant) or important

## EXHIBIT 10.6 Novel Approach to Rating Affect



# Category Scales

- Rating scales that consist of several response categories, often providing respondents with alternatives to indicate positions on a continuum
  - Ordered categories based on a particular descriptive or evaluative dimension can provide further information
- Measure attitude with greater sensitivity than a two-point response scale
- Question wording is an extremely important factor in the usefulness of these scales

## EXHIBIT 10.7 Commonly Applied Category Scale Descriptions

Quality				
Excellent	Good	Fair	Poor	
Very good	Fairly good	Neither good nor bad	Not very good	Not good at all
Well above average	Above average	Average	Below average	Well below average
Importance				
Very important	Fairly important	Neutral	Not so important	Not at all important
Interest				
Very interested		Somewhat interested		Not very interested
Satisfaction				
Completely satisfied	Somewhat satisfied	Neither satisfied nor dissatisfied	Somewhat dissatisfied	Completely dissatisfied
Very satisfied	Quite satisfied	Somewhat satisfied	Not at all satisfied	

## EXHIBIT 10.7 Commonly Applied Category Scale Descriptions (cont'd.)

Frequency				
All of the time	Very often	Often	Sometimes	Hardly ever
Very often	Often	Sometimes	Rarely	Never
All of the time	Most of the time	Some of the time	Just now and then	
Truth				
Very true	Somewhat true	Not very true	Not at all true	
Definitely yes	Probably yes	Probably no	Definitely no	
Uniqueness				
Very different	Somewhat different	Slightly different	Not at all different	
Extremely unique	Very unique	Somewhat unique	Slightly unique	Not at all unique

# The Likert Scale

- Respondents indicate their attitudes by checking how strongly they agree or disagree with carefully constructed statements, ranging from very positive to very negative attitudes
  - Individuals generally choose from approximately five, e.g., “strongly agree,” “agree,” “uncertain,” “disagree,” and “strongly disagree”
  - Researchers assign scores to each possible response (e.g., 1, 2, 3, 4, and 5, respectively)
- Attitude scores are at best interval and not ratio

# Selecting Items for a Likert Scale

- A researcher may use multiple items to represent a single attitudinal concept and generate a large number of statements before putting together the final questionnaire
- A pretest may be conducted using these items, allowing for an item analysis to be performed
  - Allows for selecting items that evoke a wide response

# Semantic Differential

- A scale type that has respondents describe their attitude using a series of bipolar rating scales
  - Bipolar adjectives—such as “good” and “bad,” “modern” and “old-fashioned,” or “clean” and “dirty”—anchor both ends (or poles) of the scale
- Semantic differentials and meaning
  - The validity of the semantic differential depends on finding scale anchors that are semantic opposites

## EXHIBIT 10.8 An Example of a Semantic Differential Scale

How does the Cheerios ad make you feel? Use the scales below to indicate your feelings.

Sad	<input type="radio"/>	Happy						
Pleased	<input type="radio"/>	Displeased						
Depressed	<input type="radio"/>	Cheerful						
Good	<input type="radio"/>	Bad						

# Scoring Semantic Differentials

- For scoring purposes, a numerical score is assigned to each position on the rating scale (i.e., 1, 2, 3, 4, 5, 6, 7 or -3, -2, -1, 0, +1, +2, +3).
  - Researchers generally assume that the semantic differential provides interval data
  - Critics argue that the data have only ordinal properties because the numerical scores are arbitrary
- Practically, marketing researchers treat them as metric (at least interval) because the amount of error introduced by assuming the intervals between choices are equal is fairly small

# Constant-Sum Scale

- With a constant-sum scale, respondents are asked to divide a fixed number of points among several attributes to indicate their relative importance
  - If respondents follow the instructions correctly, the results will approximate interval measures
  - The scale is flawed because the last response is completely determined by the way the respondent has scored the other choices

# Graphic Rating Scales

- Present respondents with a graphic continuum
  - Respondents are allowed to choose any point on the continuum to indicate their attitude
- Ways to determine a respondent's score
  - Measure the length (in millimeters) from one end of the continuum to the point marked by the respondent
  - Divide the line into predetermined scoring categories (lengths) and record respondents' marks accordingly
- Advantage
  - Allows the researcher to choose any interval desired for purposes of scoring

## EXHIBIT 10.9 Graphic Rating Scale

Please evaluate each attribute in terms of how important it is to you by placing an X at the position on the horizontal line that most reflects your feelings.

Seating comfort	Not important _____	Very important
In-flight meals	Not important _____	Very important
Airfare	Not important _____	Very important

Use the scales below to indicate how well each adjective describes the wine you just sampled.



Source: © Cengage Learning 2013.

## EXHIBIT 10.10 Using Face Scales to Assess Attitudes

Respondent 1

How pleased are you with the breakfast items at McDonalds?



Respondent 2

How pleased are you with the breakfast items at McDonalds?



# Ranking

- Consumers often rank order their preferences
  - An ordinal scale may be developed by asking respondents to rank order (from most preferred to least preferred) a set of objects or attributes
- The ranking scale suffers from inflexibility in that if we know how someone ranked five out of six alternatives, we know the answer to the sixth

# Paired Comparisons

- In paired comparisons, the respondents are presented with two objects at a time and asked to pick the one they prefer
  - Ranking objects with respect to one attribute is not difficult if only a few products are compared, but as the number of items increases, the number of comparisons increases geometrically  $[(n)(n - 1)/2]$
- If the number of comparisons is too large
  - Respondents may fatigue and no longer carefully discriminate among them

# Direct Assessment of Consumer Attitudes

- Attitudes, as hypothetical constructs, cannot be observed directly
- One's attitude can be inferred by the way he or she responds to multiple attitude indicators
- The decision whether to use ranking, sorting, rating, or a choice technique is determined by:
  - The problem definition
  - The type of statistical analysis desired

# How Many Scale Categories or Response Positions?

- A matter of sensitivity
- More categories are better than fewer
  - Be careful that the number category responses does not become taxing to respondents

# Balanced or Unbalanced Rating Scale?

- The fixed-alternative format may be balanced or unbalanced
  - A balanced scale has an equal number of positive and negative categories and a neutral point
  - Unbalanced rating scales have more response categories at one end than the other
  - The choice generally depends on the concept or the researcher's knowledge about attitudes toward the stimulus to be measured

# Forced-Choice Scales?

- In many situations, a respondent has not formed an attitude towards a concept
  - If a forced-choice rating scale compels the respondent to answer, the response is merely a function of the question
  - If answers are not forced, the midpoint of the scale may be used by the respondent to indicate unawareness as well as indifference
  - If many respondents are expected to be unaware of the attitudinal object, this problem may be eliminated by using a “no opinion” category

# Single or Multiple Items?

- Depends on:
  - The complexity of the issue to be investigated
  - The number of dimensions the issue contains
  - The level of abstraction of the phenomenon
- The researcher's conceptual definition will be helpful in making this choice

# Attitudes and Intentions

- Attitudes are considered a function of a person's beliefs about some activity weighted by their evaluations of those characteristics
- Research in this area is sometimes referred to as a multi-attribute model or reasoned action approach

# Multi-Attribute Attitude Score

- Attitudes are modeled with a multi-attribute approach
  - Take belief scores assessed with some type of rating scale
  - Multiply each belief score by an evaluation also supplied using some type of rating scale
  - Sum each resulting product
- Advantages
  - Results can provide management with feedback on the relative attitude scores and identify characteristics that are most in need of being improved

# Behavioral Intention

- Intentions represent the behavioral expectations of an individual toward an attitudinal object
- The component of interest to marketers is buying intention, a tendency to seek additional information, or plans to visit a showroom
  - Category scales for measuring the behavioral component of an attitude ask about a respondent's likelihood of purchase or intention to perform some future action, e.g., "How likely is it that you will purchase...?"