

MEASUREMENT

UNIT 3

BY:- HIMANSHU GOEL

What is to be measured?

- There is more than one way to measure a particular concept.
 - For example, salesperson's success can be measured in dollar or unit sales volume or share of accounts lost.
- Process of precisely assigning scores or numbers to attributes of people or objects.
 - Precision is the goal.

Operational definitions

- A **concept** is a generalised idea about a class of objects, attributes, occurrences, or processes.
- An operational definition gives meaning to a concept by specifying the activities or operations necessary to measure it.
- A conceptual definition is a verbal explanation of the meaning of a concept.
 - What it is and what it is not?

Operational definitions

.....
TABLE 8.1 Subjective happiness: an operational definition⁸

Concept	Conceptual Definition	Operational Definition
Subjective happiness	A global subjective assessment of whether one is a happy or unhappy person. Not equivalent to levels of affect and satisfaction with life. One may have an unfortunate life and yet still be happy.	Self report. Four items (questions): 1. In general I consider myself (1 Unhappy) to (7 Happy). 2. Compared to most of my peers I consider myself (1 Unhappy) to (7 Happy). 3. Some people are generally happy. They enjoy life regardless of what is going on, getting the most out of everything. To what extent does this characterisation describe you? (1 Not at all) (7 To a large extent). 4. Some people are generally not very happy. Although they are not depressed, they never seem as happy as they might be. To what extent does this characterisation describe you? (1 To a large extent) (7 Not at all). Happiness score is the average of the responses across each of the four scales. Ranges from 1 (Very Unhappy) to 7 (Very Happy).

Measurement and Scaling cont.

- **Measurement**
 - assigning numbers to certain characteristics of objects according to certain pre-specified rules
- **Scaling**
 - creating a continuum upon which objects are located

Types of scales

- A **scale** is any series of items that are arranged progressively according to value or magnitude, into which an item can be placed according to its quantification.
- Four types of scale:
 - nominal
 - ordinal
 - interval
 - ratio.

Types of scales

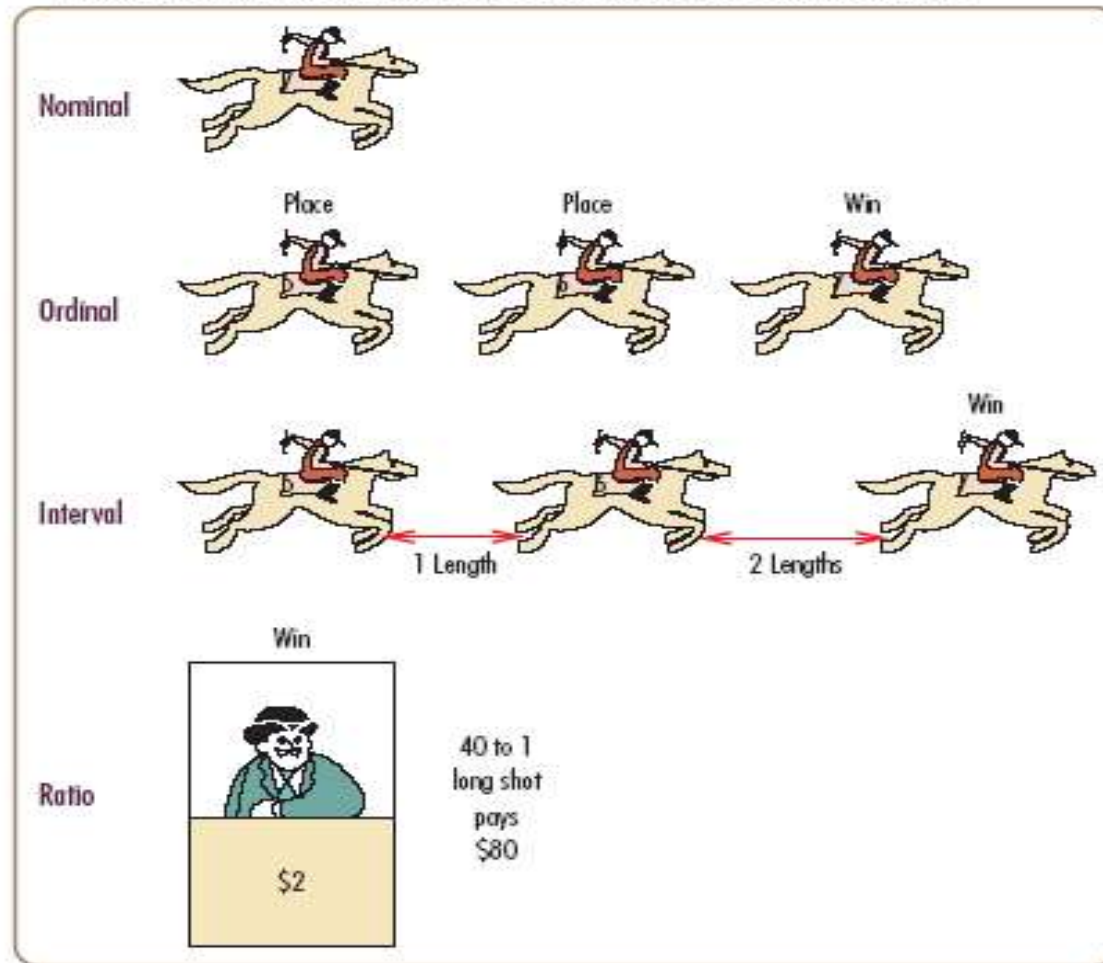
- **Nominal**: numbers or letters assigned to objects serve as labels for identification or classification.
 - Ronaldinho wears #10 for Brazil.
 - The number 7 on a horse's colours.
- **Ordinal**: arranges objects or alternatives according to their magnitude in an ordered relationship.
 - A horse finishing first in a race.

Types of scales

- **Interval**: arranges objects according to their magnitudes and distinguishes this ordered arrangement in units of equal intervals.
 - Winning horse is two lengths ahead of the second–placed horse.
- **Ratio**: absolute rather than relative quantities and an absolute zero where there is an absence of a given attribute.
 - Horse #7 is twice as heavy as horse #5.
 - 40 to 1 long shot pays \$80.

Types of scales

EXHIBIT 8.2 Nominal, ordinal, interval, and ratio scales provide different information



Scale Characteristics

- **Assignment**
 - Identify each object in a set
- **Order**
 - Relative magnitude of each scale point
- **Distance**
 - Difference between each descriptor
- **Origin**
 - Having a true zero point

Properties of Scale Measurements

	Description	Order	Distance	Origin
Nominal	X			
Ordinal	X	X		
Interval	X	X	X	
Ratio	X	X	X	X

Nominal

- Numbers identify and classify objects
- Mutually exclusive categories

Male.....1

Female.....2

- Attitude scales: dichotomous
- Used for counting purposes
- Statistics: frequency/percentage distribution and the mode

Ordinal

- Ranking system
- Indicates relative position, not magnitude
- Attitude scales: comparative, rank order, paired comparison
- Indicate your most preferred subject (as 1) and so on
 - Marketing Research 4
 - International Marketing 3
 - Introduction to Marketing 1
 - Statistics for Marketing 5
 - Consumer Behaviour 2
- Statistics: Frequencies, mode, median, percentiles

Interval

- Numerically equal distances on the scale represent equal values in the characteristic being measured
- Attitude scales: Likert, Stapel, semantic-differential
- Differences can be compared

To what extent do you involve your children in family decision making?

Not at all
involved

Extremely

1 2 3 4 5 6 7 8 9 10

- Possesses order and distance, not magnitude
- Statistics: mean and standard deviation, hypothesis testing

Ratio

- Same properties as interval scaled data with the advantage that it possesses a natural or absolute origin (zero point)
- Identify, rank and compare intervals
- Statistics: all those as interval plus coefficient of variation

Mathematical and statistical analysis of scales

- Type of scale used will determine the form of statistical analysis.

TABLE 8.2 Descriptive statistics for types of scales

Type of Scale	Numerical Operation	Descriptive Statistics
Nominal	Counting	Frequency in each category Percentage in each category Mode
Ordinal	Rank ordering	Median Range Percentile ranking
Interval	Arithmetic operations that preserve order and relative magnitudes	Mean Standard deviation Variance
Ratio	Arithmetic operations on actual quantities	Geometric mean Coefficient of variation

Note: All statistics appropriate for lower-order scales (nominal being the lowest) are appropriate for higher-order scales (ratio being the highest).

Index measures

- An **attribute** is a single characteristic or fundamental feature of an object, person, situation or issue.
- **Index** or **composite** measures are multi-item instruments for measuring a single concept with several attributes.
 - For example, index of social class based on residence, occupation, and education.
- One method for increasing accuracy.

Three criteria for good measurement

- **Reliability**
 - The degree to which measures are free from random error and therefore yield consistent results.
- **Validity**
 - The ability of a scale to measure what was intended to be measured.
- **Sensitivity**
 - The ability to accurately measure variability in stimuli or responses.

Reliability

- Applies to a measure when similar results are obtained over time and across situations.
 - For example, Tailor measuring with a tape measure obtains a true value of length repeatedly.
- Two dimensions: **repeatability** and **internal consistency**
- **Test–retest method** used to determine repeatability by administering the same scale at two separate points in time to test for stability.

Reliability

- **Split-half method** used to determine internal consistency by checking the results of one-half of a set of scaled items against the results from the other half.
- **Equivalent-form method** measures the correlation between alternative instruments, designed to be as equivalent as possible, administered to same group of subjects.

Validity

- To measure what we intend to measure.
 - For example, a hard-working student who understands statistics does poorly in a test that measures ability to do arithmetic and memorise formulas instead.
- Three approaches to establishing validity:
 - Face or content validity
 - Criterion validity
 - Construct validity.

Establishing validity

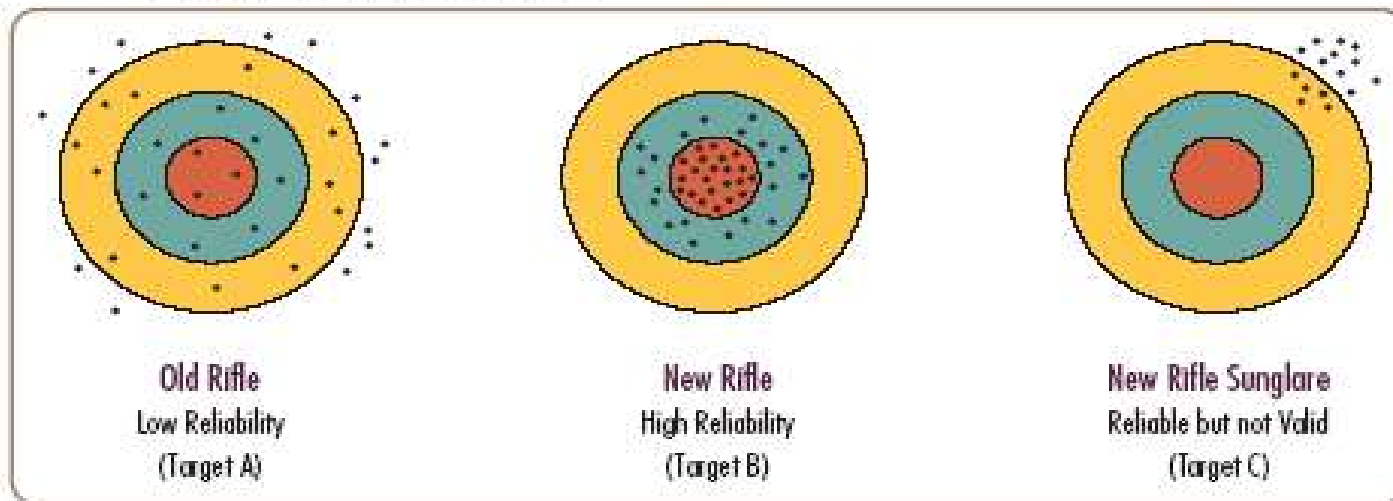
- **Face or content validity** refers to a professional agreement that a scale's content logically appears to accurately reflect what was intended to be measured.
- **Criterion validity** refers to the ability of a measure to correlate with other standard measures of the same construct or established criterion.

Establishing validity

- **Construct validity** refers to the ability of a measure to provide empirical evidence consistent with a theory–based concept.

Reliability versus validity

EXHIBIT 8.3 Reliability and validity on target



Attitudes defined

- **Attitude** is an enduring disposition to consistently respond in a given manner to various aspects of the world.
- Attitude is a hypothetical construct that is not directly observable but is measurable through indirect indicators.
- Measuring attitudes is important because changing consumers' attitudes is a major marketing goal.

Definitions

- Construct
 - to describe or explain some aspect of nature (tends to be theoretical) e.g., attitude towards...
- Dimension
 - Components of the construct e.g., cognitive, affective, conative
- Item
 - The statement in a questionnaire that attempts to measure the construct e.g., I will quit smoking this year.

Attitude Measurement

- Majority of questions in marketing research are designed to measure attitudes
- Mental states used by individuals to structure the way they perceive and respond to their environment. Attitudes include:
 - **cognitive** or **knowledge** component
 - **affective** or **liking** component
 - **intention** or **action** component

The attitude–measuring process

- Obtaining verbal statements from respondents generally requires that respondents perform a task such as ranking, rating, sorting, or making choices.
 - **Ranking**: rank order a small number of objects on the basis of preference.
 - **Rating**: estimate the magnitude of a characteristic.
 - **Sorting**: arrange objects into piles or classifications.

Attitude rating scales

- Rating scales designed to enable respondents to report the intensity of their attitudes.
 - Simple attitude scales
 - Category scales
 - Likert scale.
 - Allow respondents to rate how strongly they agree or disagree with a statement.
 - Semantic differential
 - 7–point rating scales that use bipolar adjectives to anchor both ends of each scale.

The Likert scale

TABLE 8.4 Likert scale items for measuring attitudes toward patients' interaction with a physician's service staff¹⁸

- | | |
|---|------------------------------------------------------------------------------------------------------------|
| 1 | <i>My doctor's office staff takes a warm and personal interest in me.</i> |
| 2 | <i>My doctor's office staff is friendly and courteous.</i> |
| 3 | <i>My doctor's office staff is more interested in serving the doctor's needs than in serving my needs.</i> |
| 4 | <i>My doctor's office staff always acts in a professional manner.</i> |

Example: Identify the components of attitude for the following attitude statements
Malaysia as a travel destination

- The weather in Malaysia is too humid.
- There are lots of sights to see in Malaysia.
- Malaysian people are warm and friendly.
- I will definitely visit Malaysia in the next year.
- I love Malaysian cuisine.
- I enjoy shopping in Malaysia.
- Malaysia is culturally different to Australia.

Attitude Rating Scale: Single-item Scales

- Only have one item to measure a construct
- Itemised-category scales - limited number of response categories

How satisfied are you with your last purchase of Nike runners?

- Very satisfied
- Quite satisfied
- Somewhat satisfied
- Not at all satisfied

Attitude Rating Scale: Single-item Scales cont.

- **Comparative**

- Compares one object, concept or person against another
- Between 3 – 9 response alternatives are generally best
- Odd number of categories is preferable
- Compared with Nike, Adidas produces sports wear that is:

Much

neither trendier

much

More

nor unfashionable

more

Trendier

Attitude Rating Scale: Single-item Scales cont.

- **Rank-order**
 - respondent compares one item with another (or many others) and ranks them
- Place a 1 beside your most preferred soft drink; 2...
 - Diet Pepsi
 - Pepsi
 - AC Cola
 - Coke
- Ordinal
- Used to measure preferences for brands and attributes

Advantages and Disadvantages of Rank-order Scales

Advantages

- Respondents discriminate brands/objects
- Takes less time and eliminates intransitive responses
- Instructions are easy to understand

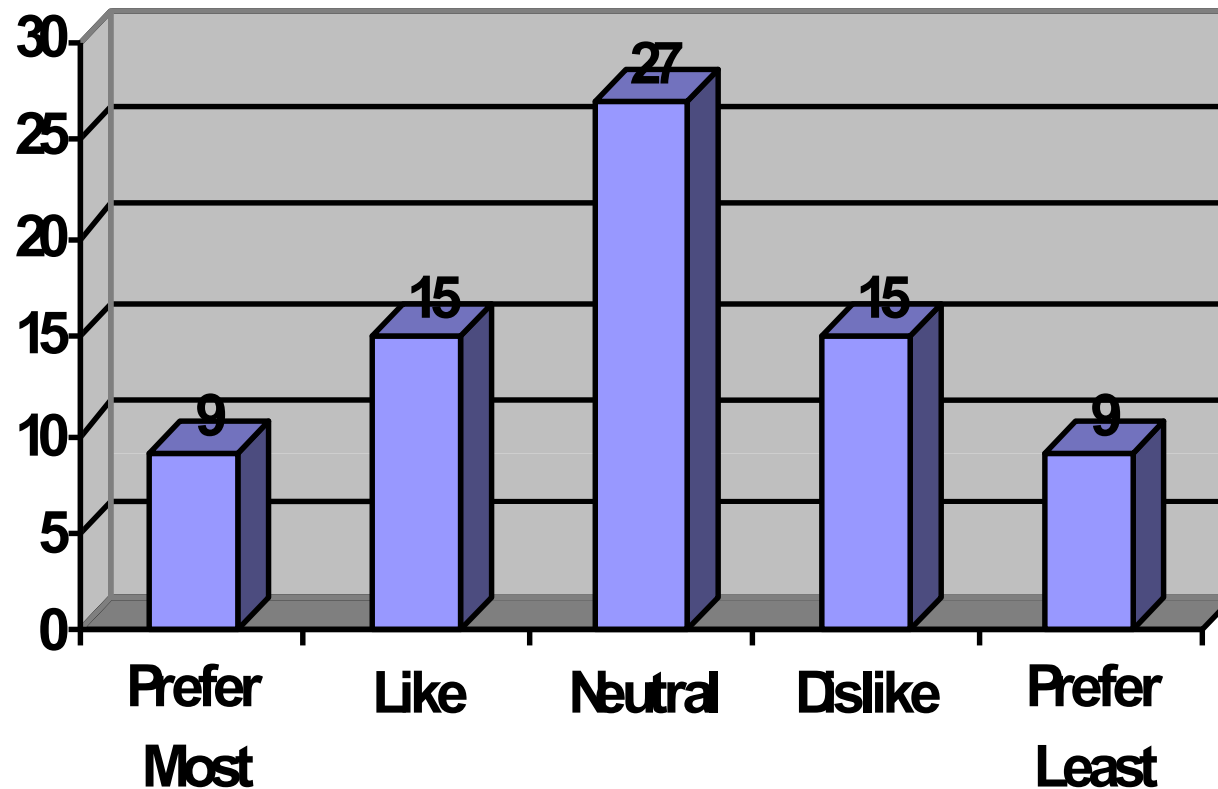
Disadvantages

- Produces ordinal data thus limits the statistics that can be performed

Attitude Rating Scale: Single-item Scales cont.

- **Q-sort**
 - Respondents sort comparative characteristics into groups
 - Between 60-90 objects can be ranked
 - Pre-specified number in each class
 - Each object is then assigned a rank order
 - Advantage: discriminates among a relatively large number of objects quickly

Sort 75 Magazines according to:



Attitude Rating Scale: Single-item Scales cont.

- **Constant-sum scaling**
 - A respondent is allocated a constant sum of units such as points among alternatives presented
 - Average is taken from all respondents
 - Ordinal in nature

Example: Constant Sum Scaling

Please divide 100 points among the following characteristics of buying a car indicating the relative importance of each characteristic to you

Characteristics	Number of points
Style	50
Ride	10
Economical	35
Warranty	5
Friendly Dealer	<u>0</u>
	100

Advantages and Disadvantages of Constant Sum Scaling

Advantages

- Finer discrimination among alternatives

Disadvantages

- Respondents may allocate more or less points available
- Cognitive effort required

Attitude Rating Scale: Single-item Scales cont.

- **Pictorial scales**

- Various categories of the scale are depicted pictorially

- How satisfied were you with the helpfulness of staff during your stay?

Λ 1 2 3 4 5 6 7 8 9 10 9

Attitude rating scales

Other scales include:

- Numerical scales
- Stapel scale
- Constant–sum scale
- Graphic rating scales
- Thurstone interval scale.

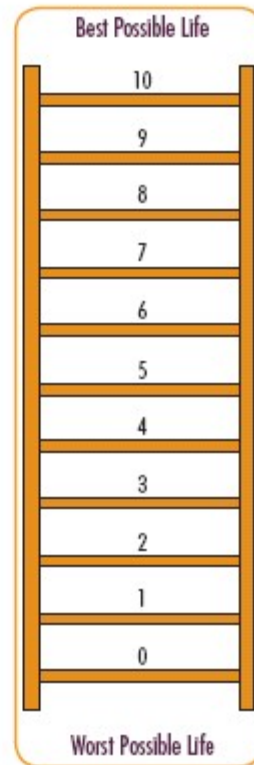
Graphic rating scales

EXHIBIT 8.6 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

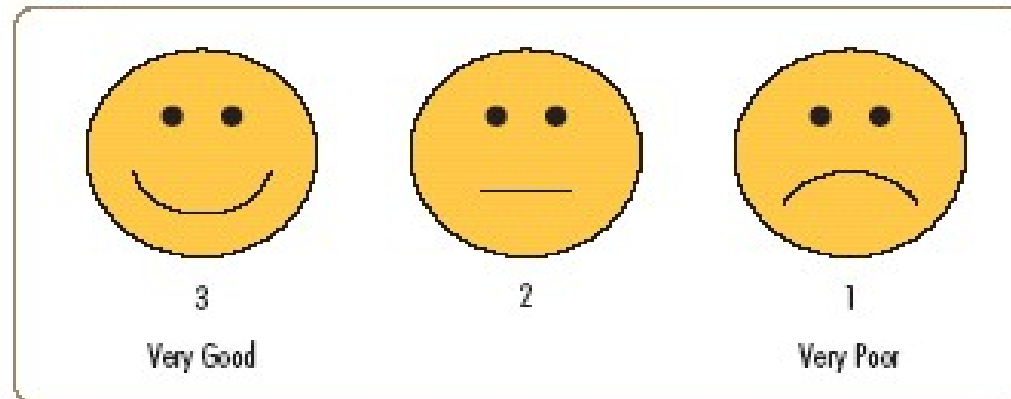
EXHIBIT 8.7 A ladder scale



Graphic rating scales

EXHIBIT 8.8 Graphic rating scale with picture response categories that stress visual communication

Happy Face Scale



Measuring behavioural intention

- The behavioural component of an attitude, typically representing a buying intention.
 - Category scales
 - Behavioural differential.

Selecting a measurement scale

- No single best answer.
- The choice of scale will depend on the nature of the attitudinal object to be measured.
- Will also depend on the manager's problem definition and backward and forward linkages to choices already made.
 - Example: to suit survey method chosen.

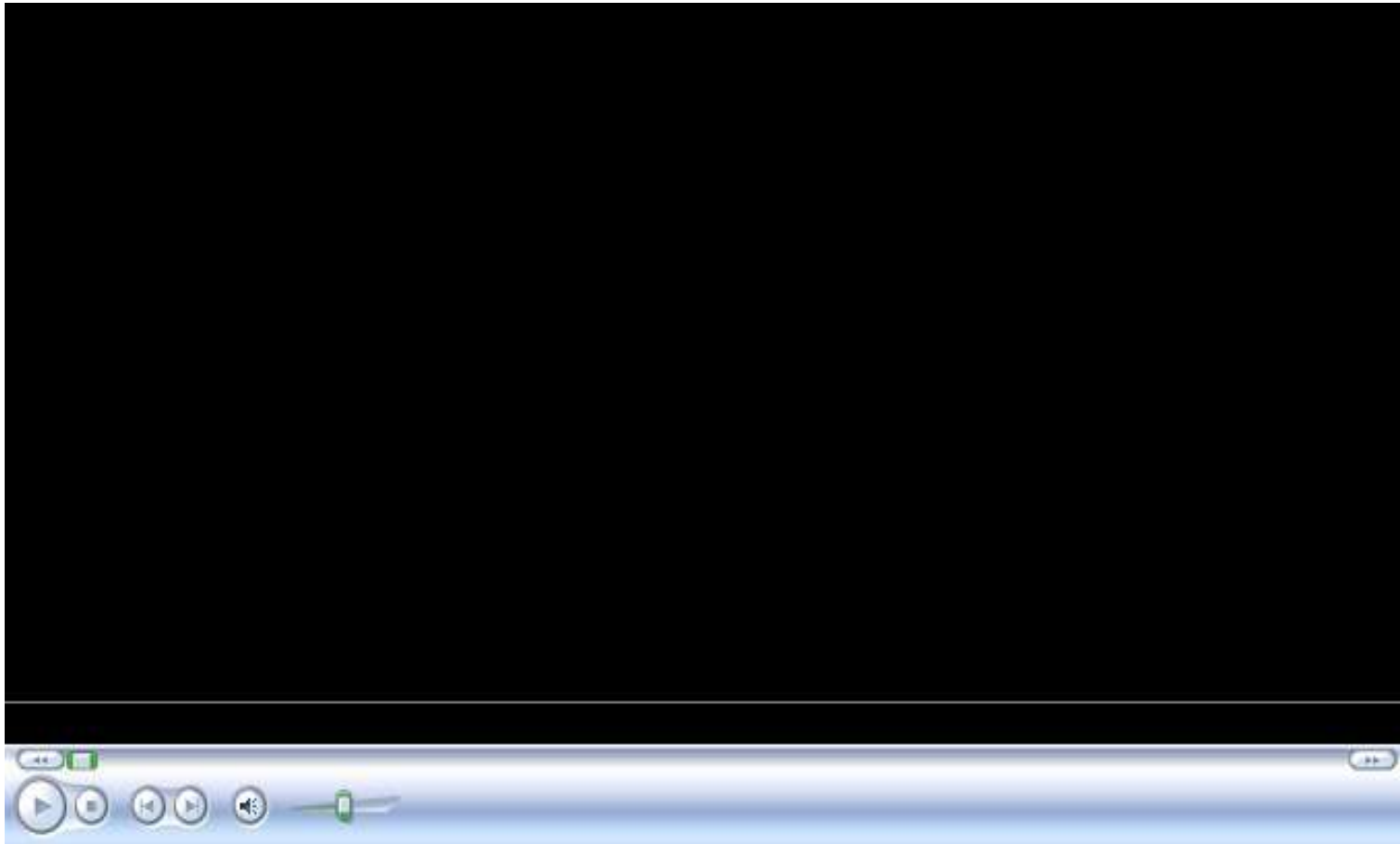
Selecting a measurement scale

- Some useful questions include:
 - Ranking, sorting, rating, or choice technique?
 - Monadic or comparative scale?
 - What type of category labels, if any?
 - How many scale categories or response positions?
 - Balanced or unbalanced rating scale?
 - A scale that forces a choice among predetermined options?
 - Single measure or index measure?

Learning Objectives

- Describe the concepts of measurement and scaling
- Identify different measurement scales
- Explain the criteria for judging good measurement
- Describe the different types of scales used for measuring attitudes
- Choose an appropriate scale for a specific survey
- Discuss some issues that may arise when developing scales to be used in cross-national research

Explain how you develop response scales



**Michael Sparks, Director
Intuitive Solutions**

Decisions in Single-item Scales

- Number of scale categories
- Types of poles used in scale
- Strength of the anchors
- Labelling of the categories
- Balance of the scale

Number of Scale Categories

5-point scale

The website is innovative

- Strongly disagree
- Disagree
- Neither agree nor disagree
- Agree
- Strongly Agree

7-point scale

The website is innovative

- Strongly disagree
- Disagree
- Disagree somewhat
- Neither agree nor disagree
- Agree somewhat
- Agree
- Strongly agree

Balanced vs. Unbalanced Scales

Balanced

How important is price, as far as you are concerned?

- Very important
- Of some importance
- Of little importance
- Of absolutely no importance

Unbalanced scales

How important is price, as far as you are concerned?

- Extremely important
- Very important
- Of some importance
- Of little importance

Forced vs. Non-forced Scales

Forced scales

In general how interested are you in trying a new brand of toothpaste?

- Very interested
- Somewhat interested
- Not too interested
- Not at all interested

Non-forced scales

In general how interested are you in trying a new brand of toothpaste?

- Very interested
- Somewhat interested
- Unsure**
- Not too interested
- Not at all interested

Multiple-item Scales

- A scale where several items are used to measure a construct
 - **Likert scale**
 - requires respondent to indicate the degree of agreement or disagreement with a variety of statements related to the attitude object

Example: Likert Scale

	Strongly disagree	Disagree	Neither agree nor	Agree	Strongly Agree
<i>The website displays is visually pleasing design</i>	1	2	3	4	5
<i>Learning to operate the website is easy for me</i>	1	2	3	4	5
<i>The website is innovative</i>	1	2	3	4	5
<i>The website loads quickly</i>	1	2	3	4	5
<i>I trust the website to keep my personal information safe</i>	1	2	3	4	5
<i>Shopping at this website is usually a satisfying experience</i>	1	2	3	4	5
<i>I will continue to purchase from this Website.</i>	1	2	3	4	5

A modified Likert scale can include 6, 7 or 10 items

Multiple-item Scales cont.

- **Thurstone scales**
 - Generate large number of statements
 - Then classified by judges according to degree of favourableness or unfavourableness
 - Time consuming and expensive
- **Semantic-differential scale**
 - Five or seven point rating scale with end points associated with bipolar labels that have semantic meaning
 - Used in comparing brand, product and company images

Multiple-item Scales cont.

- **Stapel scales**
 - Uni-polar rating scale with 10 categories without a neutral point
 - No need to pre-test adjectives

The store provides a fast checkout service	-5	-4	-3	-2	-1	+1	+2	+3	+4	+5
The staff at the store are friendly	-5	-4	-3	-2	-1	+1	+2	+3	+4	+5

Continuous Rating Scales

- 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
- Interval data
- Easy to construct but scoring is cumbersome and unreliable

Example: Continuous Rating Scales

To what extent do you involve your children in family decision making? (Place an X on the line)



Not at all

Extremely involved

To what extent do you involve your children in family decision making?

Not at all

Extremely involved

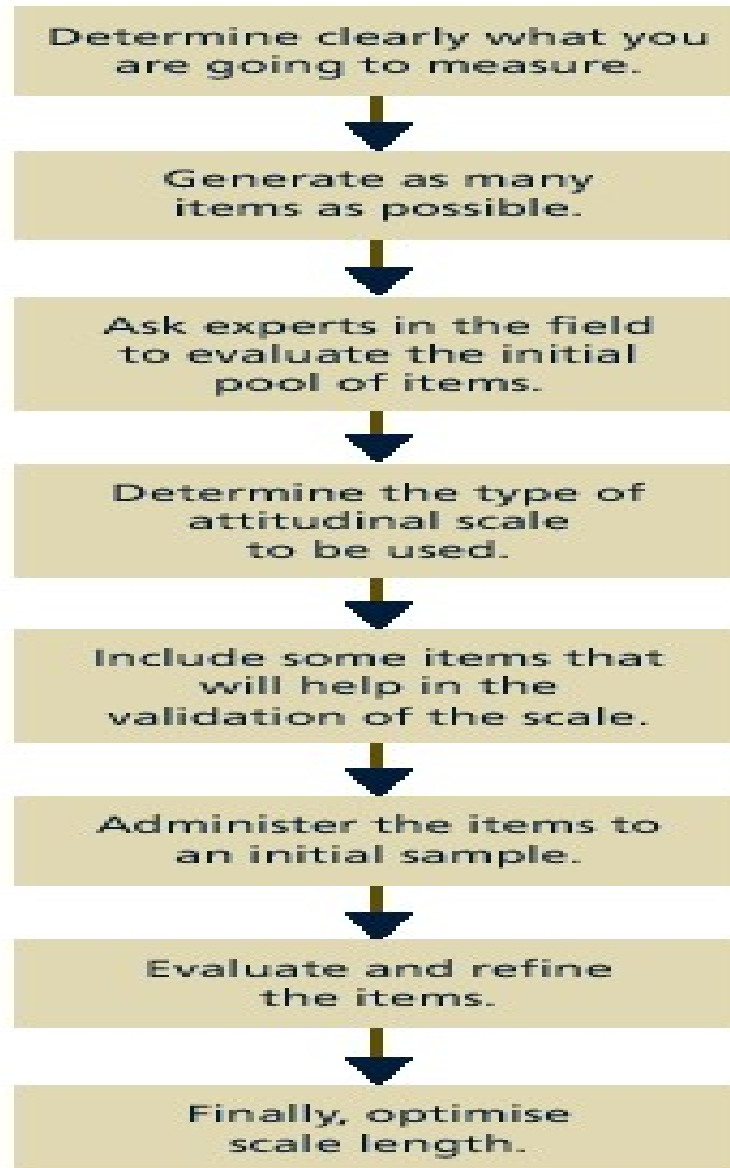
1 2 3 4 5 6 7 8 9 10

Example: Rapid Analysis and Testing Environment

- **RATE** is a continuous measurement of “gut reaction”. Roy Morgan’s The Worm can measure response to concept testing, advertising, even election debates.
- Between 50 to 100 respondents are provided with a Worming dial and asked to record their reaction to the material being tested.
- The information is fed into a computer which tabulates second-by-second response profiles. The responses are also stored in a data file for further analysis.
- Respondent demographics are obtained via a questionnaire and used to cross-tab the key information.

Source: www.roymorgan.com

Developing Multiple-item Scales



Choosing a Scale

- **Complicated by:**
 - Many different techniques each with own strengths and weaknesses
 - Virtually any technique can be adapted for a given situation
- **Influenced by:**
 - Information required
 - Budget constraints
 - Compatibility of scale

Scales in Cross-national Research

- **Responses can be affected by:**
 - Low literacy and educational levels
 - Cultural differences
 - Points on a scale vary between countries
 - Semantic differential closest to pan cultural scale