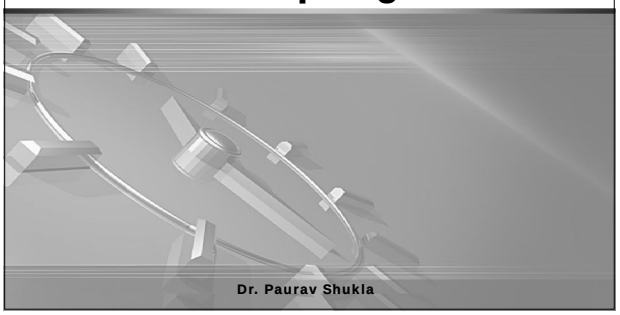
Measurement, Scaling, Sampling



Dr. Paurav Shukla

Session Objectives

- Measurement techniques
- Scaling techniques
- Sampling techniques

2

Important issues

- Validity
- Reliability

3

Primary Scales of Measurement

Scale					
Nominal	Numbers Assigned to Runners	 7	 8	 3	Finish
Ordinal	Rank Order of Winners	 Third place	 Second place	 First place	Finish
Interval	Performance Rating on a 0 to 10 Scale	8.2	9.1	9.6	
Ratio	Time to Finish, in Seconds	15.2	14.1	13.4	

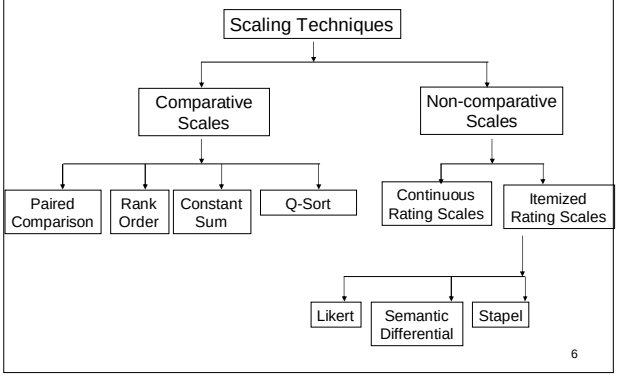
4

Primary Scales of Measurement

Scale	Basic Characteristics	Common Examples	Marketing Examples	Permissible Statistics	
				Descriptive	Inferential
Nominal	Numbers identify & classify objects	Social Security nos., numbering of football players	Brand nos., store types	Percentages, mode	Chi-square, binomial test
Ordinal	Nos. indicate the relative positions of objects but not the magnitude of differences between them	Quality rankings, rankings of teams in a tournament	Preference rankings, market position, social class	Percentile, median	Rank-order correlation, Friedman ANOVA
Interval	Differences between objects can be compared, zero point is arbitrary	Temperature (Fahrenheit, Celsius)	Attitudes, opinions, index nos.	Range, mean, standard deviation	Product-moment correlation, t tests, regression
Ratio	Zero point is fixed, ratios of scale values can be compared	Length, weight	Age, sales, income, costs	Geometric mean, harmonic mean	Coefficient of variation

5

A Classification of Scaling Techniques



6

Basic Non-comparative Scales

Scale	Basic Characteristics	Examples	Advantages	Disadvantages
Continuous Rating Scale	Place a mark on a continuous line	Reaction to TV commercials	Easy to construct	Scoring can be cumbersome unless computerized
Itemized Rating Scales				
Likert Scale	Degrees of agreement on a 1 (strongly disagree) to 5 (strongly agree) scale	Measurement of attitudes	Easy to construct, administer, and understand	More time-consuming
Semantic Differential	Seven-point scale with bipolar labels	Brand, product, and company images	Versatile	Controversy as to whether the data are interval
Stapel Scale	Unipolar ten-point scale, -5 to +5, without a neutral point (zero)	Measurement of attitudes and images	Easy to construct, administer over telephone	Confusing and difficult to apply

7

Semantic Differential Scale

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

8

Balanced Scale		Unbalanced Scale	
Nivia for Men is		Nivia for Men is	
Extremely good	_____	Extremely good	_____
Very good	_____	Very good	_____
Good	_____	Good	_____
Bad	_____	Somewhat good	_____
Very bad	_____	Bad	_____
Extremely bad	_____	Very bad	_____

9

Rating Scale Configurations

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

____	____	____	____	____	____	____
Very harsh	Harsh	Somewhat harsh	Neither harsh nor gentle	Somewhat gentle	Gentle	Very gentle
-3	-2	-1	0	+1	+2	+3

5)

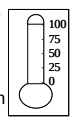
Very harsh	Neither harsh nor gentle	Very gentle

10

Some Unique Rating Scale Configurations

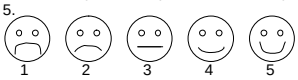
Thermometer Scale
Instructions: Please indicate how much you like McDonald's hamburgers by coloring in the thermometer. Start at the bottom and color up to the temperature level that best indicates how strong your preference is.
Form:

Like very much



Dislike very much

Smiling Face Scale
Instructions: Please point to the face that shows how much you like the Barbie Doll. If you do 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.
Form:



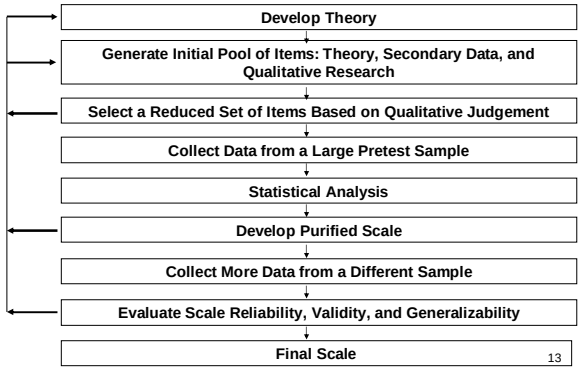
11

Summary of Itemized Scale Decisions

1) Number of categories	Although there is no single, optimal number, traditional guidelines suggest that there should be between five and nine categories
2) Balanced vs. unbalanced	In general, the scale should be balanced to obtain objective data
3) Odd/ even no. of categories	If a neutral or indifferent scale response is possible from at least some of the respondents, an odd number of categories should be used
4) Forced vs. non-forced	In situations where the respondents are expected to have no opinion, the accuracy of the data may be improved by a non-forced scale
5) Verbal description	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
6) Physical form	A number of options should be tried and the best selected

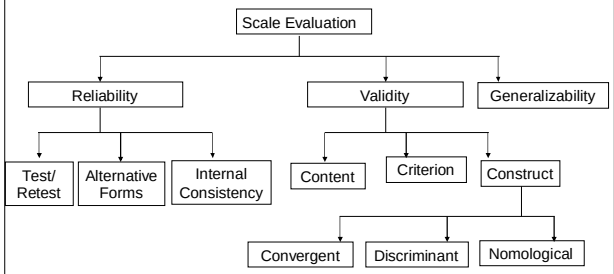
12

Development of a Multi-item Scale



13

Scale Evaluation



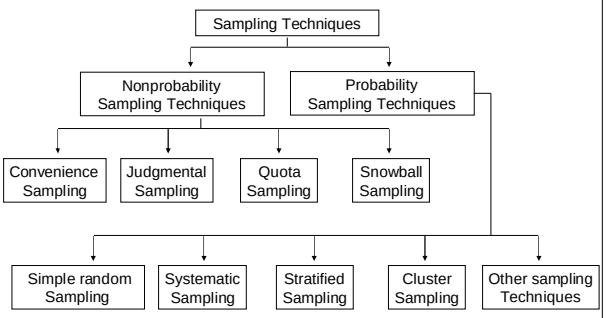
14

Potential Sources of Error on Measurement

1) Other relatively stable characteristics of the individual that influence the test score, such as intelligence, social desirability, and education.
2) Short-term or transient personal factors, such as health, emotions, fatigue.
3) Situational factors, such as the presence of other people, noise, and distractions.
4) Sampling of items included in the scale: addition, deletion, or changes in the scale items.
5) Lack of clarity of the scale, including the instructions or the items themselves.
6) Mechanical factors, such as poor printing, overcrowding items in the questionnaire, and poor design.
7) Administration of the scale, such as differences among interviewers.
8) Analysis factors, such as differences in scoring and statistical analysis.

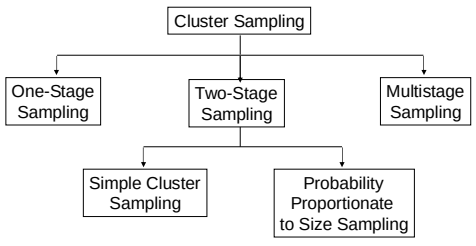
15

Classification of Sampling Techniques



16

Types of Cluster Sampling



17

Technique	Strengths	Weaknesses
Nonprobability Sampling		
Convenience sampling	Least expensive, least time-consuming, most convenient	Selection bias, sample not representative, not recommended for descriptive or causal research
Judgmental sampling	Low cost, convenient, not time-consuming	Does not allow generalization, subjective
Quota sampling	Sample can be controlled for certain characteristics	Selection bias, no assurance of representativeness
Snowball sampling	Can estimate rare characteristics	Time-consuming
Probability sampling		
Simple random sampling (SRS)	Easily understood, results projectable	Difficult to construct sampling frame, expensive, lower precision, no assurance of representativeness.
Stratified sampling	Include all important subpopulations, precision	Difficult to select relevant stratification variables, not feasible to stratify on many variables, expensive
Cluster sampling	Easy to implement, cost effective	Imprecise, difficult to compute and interpret results
Systematic sampling	Can increase representativeness, Easier to implement than SRS, sampling frame not necessary	Can decrease representativeness

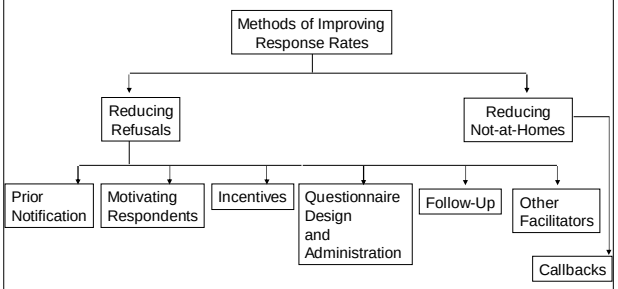
18

Choosing Nonprobability vs. Probability Sampling

Factors	Conditions Favouring the Use of Nonprobability sampling	Probability sampling
Nature of research	Exploratory	Conclusive
Relative magnitude of sampling and Non-sampling errors	Non-sampling errors are larger	Sampling errors are larger
Variability in the population	Homogeneous (low)	Heterogeneous (high)
Statistical considerations	Unfavourable	Favourable
Operational considerations	Favourable	Unfavourable

19

Improving Response Rates



20

Tennis's Systematic Sampling Returns a Smash

Tennis magazine conducted a mail survey of its subscribers to gain a better understanding of its market. Systematic sampling was employed to select a sample of 1,472 subscribers from the publication's domestic circulation list. If we assume that the subscriber list had 1,472,000 names, the sampling interval would be 1,000 (1,472,000/1,472). A number from 1 to 1,000 was drawn at random. Beginning with that number, every 1,000th subscriber was selected.

A brand-new dollar bill was included with the questionnaire as an incentive to respondents. An alert postcard was mailed one week before the survey. A second, follow-up, questionnaire was sent to the whole sample ten days after the initial questionnaire. There were 76 post office returns, so the net effective mailing was 1,396. Six weeks after the first mailing, 778 completed questionnaires were returned, yielding a response rate of 56%.

21