What and why of data analysis I

Prof. Ajai Pratap Singh

Dept. of Applied Psychology

VBS Purvanchal University Jaunpur
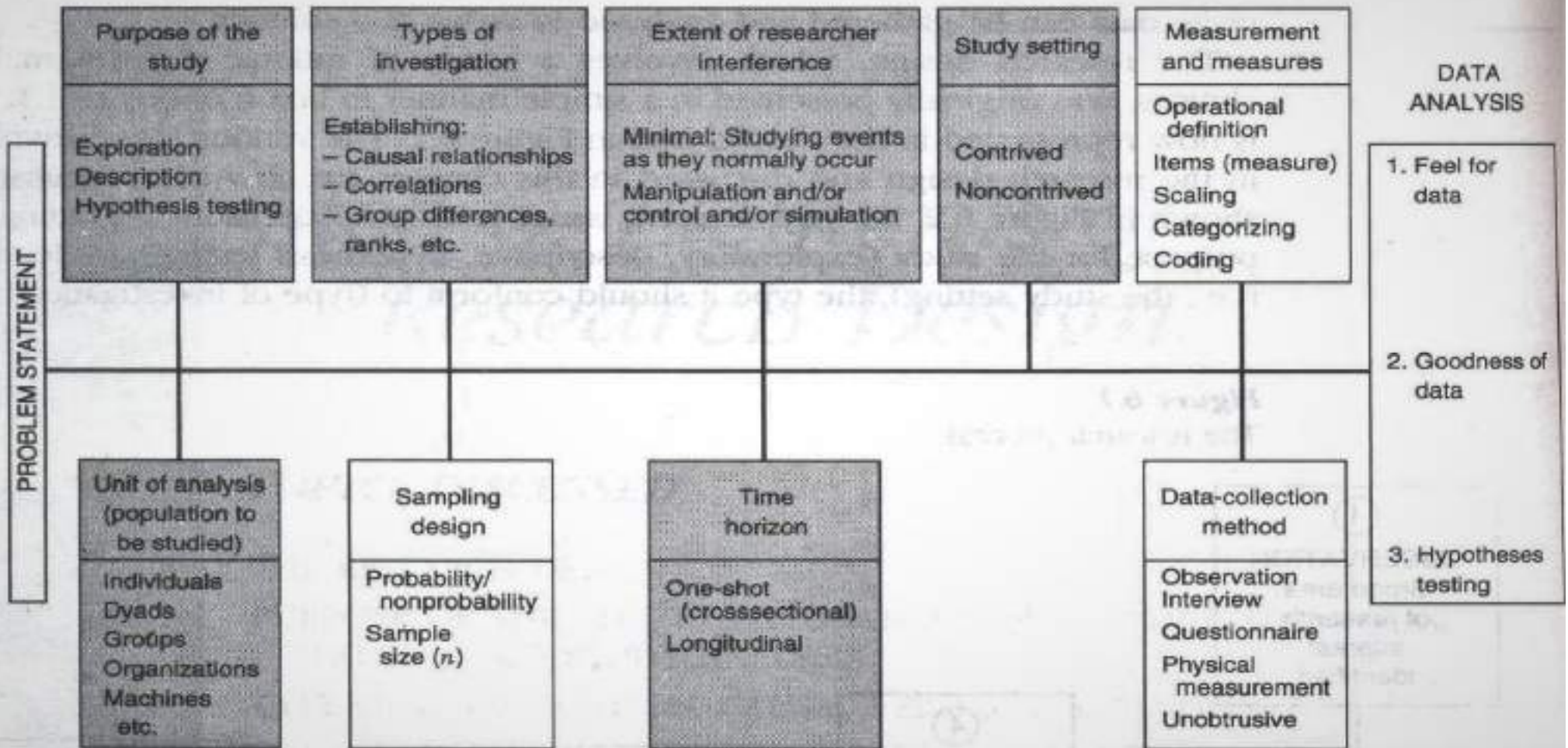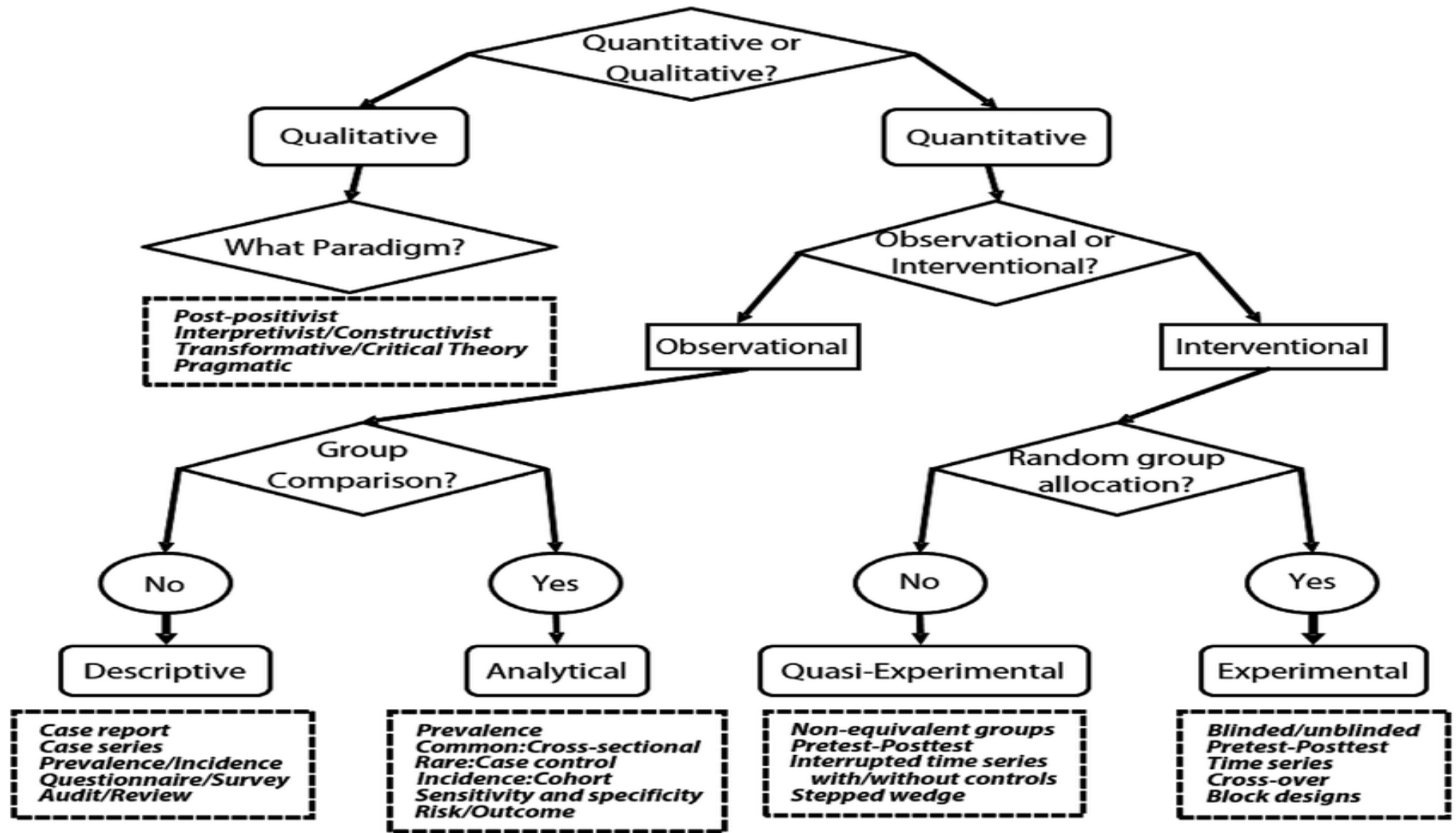
Uttar Pradesh
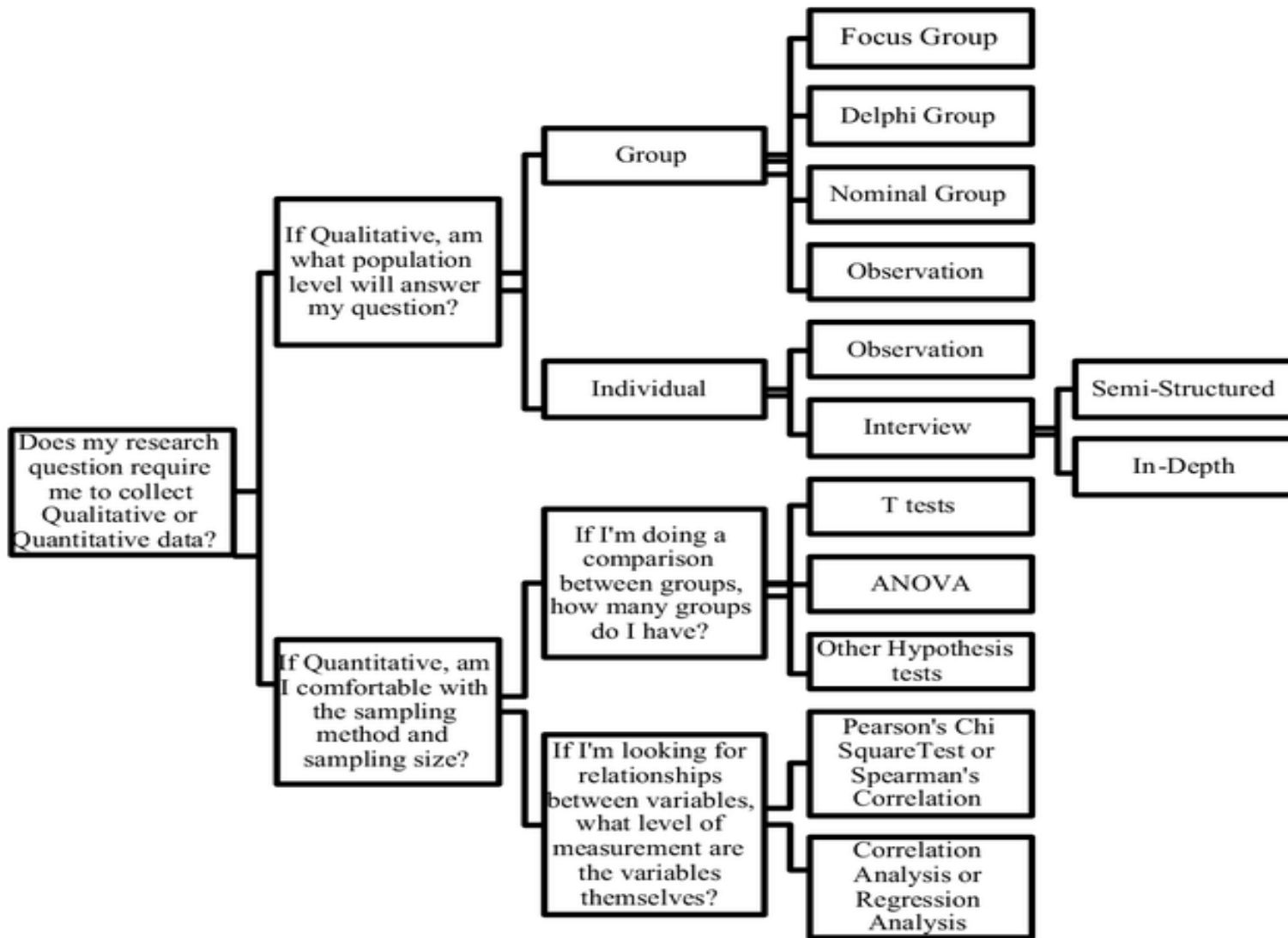
Email: ajaisingh27@gmail.com

The research design.

DETAILS OF STUDY				MEASUREMENT	DATA ANALYSIS

Purpose of the study	Types of investigation	Extent of researcher interference	Study setting	Measurement and measures	
Exploration Description Hypothesis testing	Establishing: — Causal relationships — Correlations — Group differences, ranks, etc.	Minimal: Studying events as they normally occur Manipulation and/or control and/or simulation	Contrived Noncontrived	Operational definition Items (measure) Scaling Categorizing Coding	1. Feel for data

PROBLEM STATEMENT

Unit of analysis (population to be studied)	Sampling design	Time horizon		Data-collection method	
Individuals Dyads Groups Organizations Machines etc.	Probability/ nonprobability Sample size (n)	One-shot (crosssectional) Longitudinal		Observation Interview Questionnaire Physical measurement Unobtrusive	3. Hypotheses testing

2. Goodness of data

Quantitative or Qualitative?

Qualitative

Quantitative

What Paradigm?

***Post-positivist
Interpretivist/Constructivist
Transformative/Critical Theory
Pragmatic***

Observational or Interventional?

Observational

Interventional

Group Comparison?

Random group allocation?

No

Yes

No

Yes

Descriptive

Analytical

Quasi-Experimental

Experimental

***Case report
Case series
Prevalence/Incidence
Questionnaire/Survey
Audit/Review***

***Prevalence
Common:Cross-sectional
Rare:Case control
Incidence:Cohort
Sensitivity and specificity
Risk/Outcome***

***Non-equivalent groups
Pretest-Posttest
Interrupted time series
    with/without controls
Stepped wedge***

***Blinded/unblinded
Pretest-Posttest
Time series
Cross-over
Block designs***

```
Does my research question require me to collect Qualitative or Quantitative data?

├── If Qualitative, am what population level will answer my question?
│   ├── Group
│   │   ├── Focus Group
│   │   ├── Delphi Group
│   │   ├── Nominal Group
│   │   └── Observation
│   └── Individual
│       ├── Observation
│       └── Interview
│           ├── Semi-Structured
│           └── In-Depth
│
└── If Quantitative, am I comfortable with the sampling method and sampling size?
    ├── If I'm doing a comparison between groups, how many groups do I have?
    │   ├── T tests
    │   ├── ANOVA
    │   └── Other Hypothesis tests
    └── If I'm looking for relationships between variables, what level of measurement are the variables themselves?
        ├── Pearson's Chi SquareTest or Spearman's Correlation
        └── Correlation Analysis or Regression Analysis
```

Statistical Techniques and Tests Classified According to Type, Number, and Measurement Scale of Variables[a]

			Criterion Variables					
			One			Two or More		
			Nominal	Ordinal	Interval	Nominal	Ordinal	Interval
Variates	One	Nominal	Chi-square test for independence Cochran Q test Fisher exact probability	Sign test Median test Mann-Whitney U test Kruskal-Wallis one-way analysis of variance	Analysis of variance			Multiple discriminant analysis
		Ordinal		Spearman's rank correlation Kendall's rank correlation	Analysis of variance with trend analysis			
		Interval	Analysis of variance		Regression analysis	Analysis of variance		Multiple regression analysis
	Two or More	Nominal		Friedman two-way analysis of variance	Analysis of variance (factorial design)			Analysis of variance
		Ordinal						
		Interval	Multiple discriminant analysis		Multiple regression analysis		Multiple discriminant analysis	Canonical correlation

[a] Adapted from R. L. Baker & R. E. Schultz (Eds.) *Instructional product research*. New York: Van Nostrand Co., 1972, p. 110.
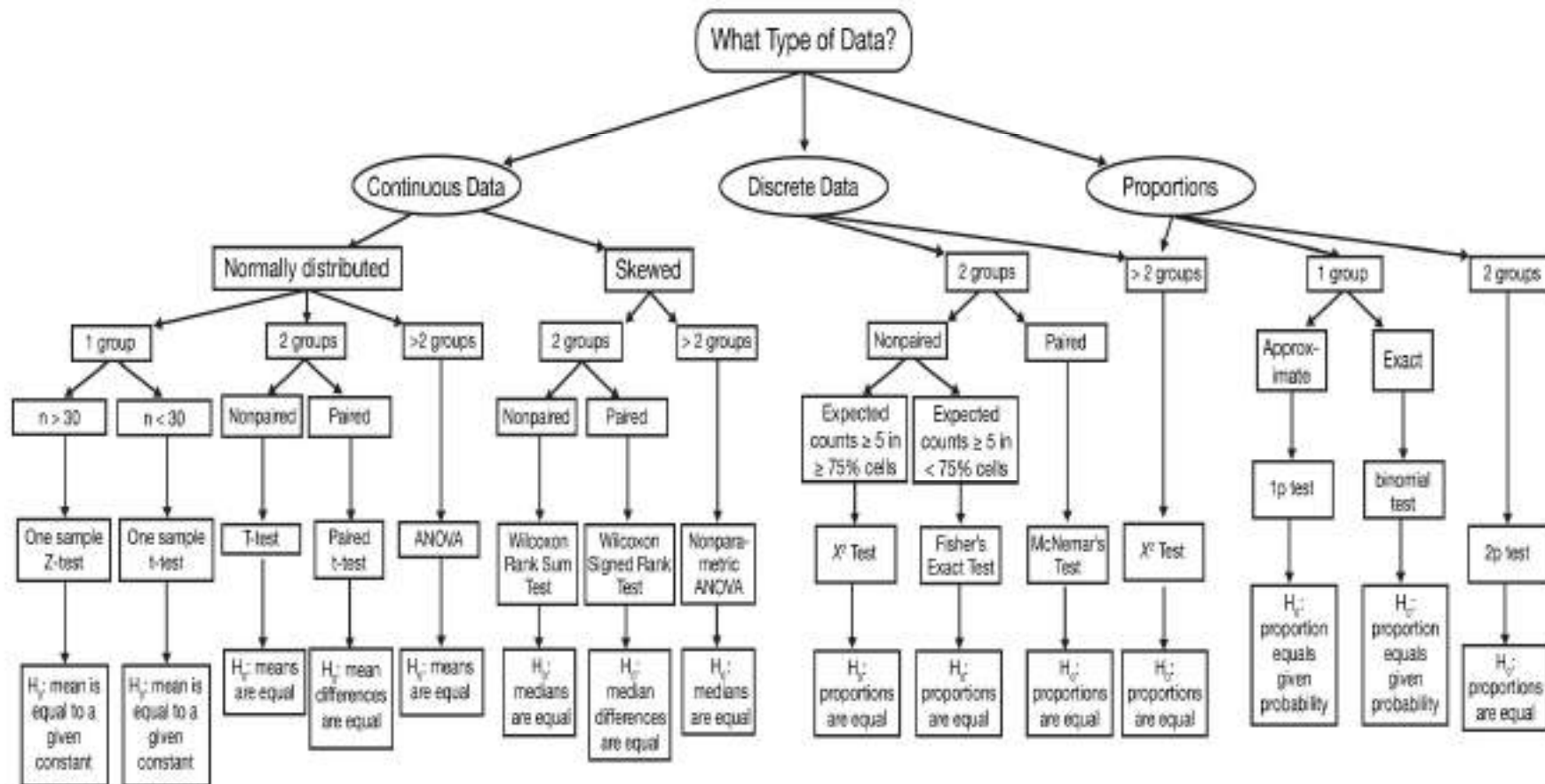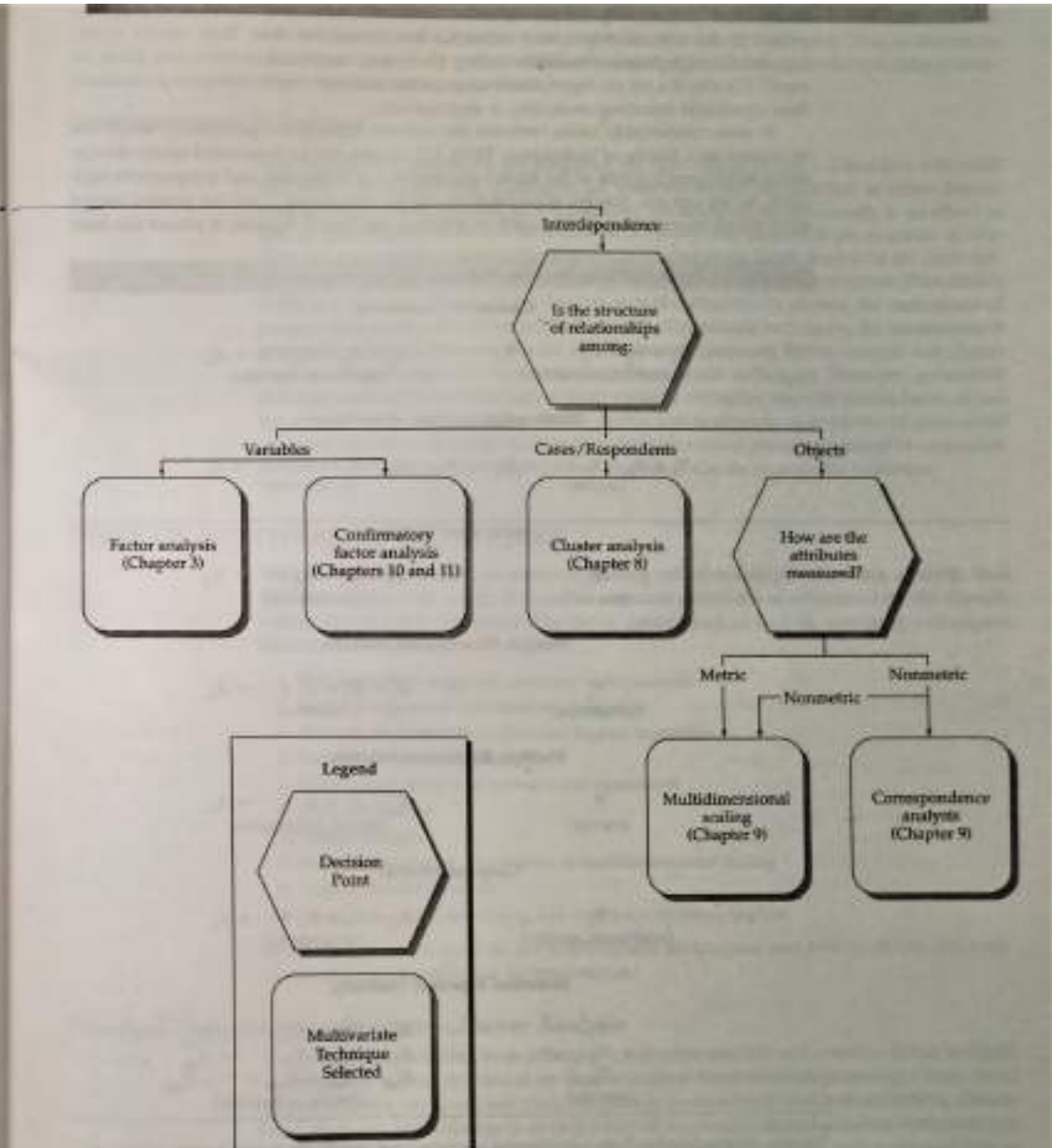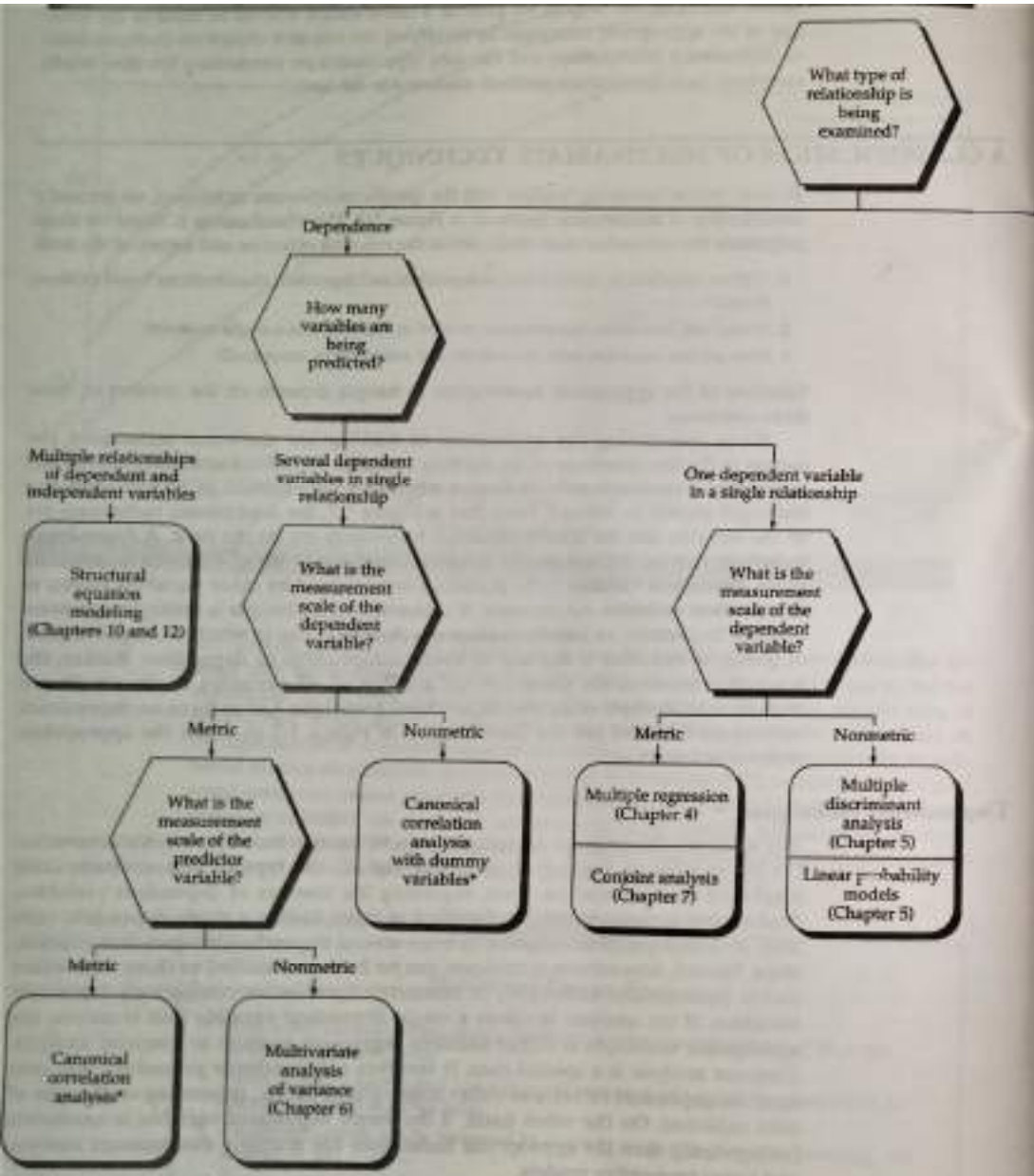
Flow chart: which test statistic should you use?

Table M 4
Use of Some Nonparametric Tests

Test	When Used	Function
Chi-square	With **nominal** data for one sample or two or more independent samples.	Tests for independence of variables.
Cochran Q	With more than two related samples measured on **nominal** scale.	Helps when data fall into two natural categories.
Fisher exact probability	With two independent samples measured on **nominal** scale.	More useful than χ^2 when expected frequencies are small.
Sign test	With two related samples measured on **ordinal** scale.	A good test for ranked data.
Median test	With one sample, to see if randomly drawn measurements are from a population with a specified median.	In a symmetric distribution, the mean and median will be the same.
Mann–Whitney U test	With two independent samples on **ordinal** data.	Analogue of the two independent sample t-tests.
Kruskal–Wallis one-way ANOVA	With more than two independent samples on an **ordinal** scale.	An alternative to one-way ANOVA where normality of distributions cannot be assumed.
Friedman two-way ANOVA	With more than two related samples on **ordinal** data.	A good alternative to two-way ANOVA where normality cannot be assumed.
Kolmogorov-Smirnov	With one sample or two independent samples measured on an **ordinal** scale.	Is a more powerful test than χ^2 or Mann-Whitney U.

What type of
relationship is
being
examined?

Dependence

Interdependence

How many
variables are
being
predicted?

Is the structure
of relationships
among:

Multiple relationships
of dependent and
independent variables

Several dependent
variables in single
relationship

One dependent variable
in a single relationship

Variables

Cases/Respondents

Objects

Structural
equation
modeling
(Chapters 10 and 12)

What is the
measurement
scale of the
dependent
variable?

What is the
measurement
scale of the
dependent
variable?

Factor analysis
(Chapter 3)

Confirmatory
factor analysis
(Chapters 10 and 11)

Cluster analysis
(Chapter 8)

How are the
attributes
measured?

Metric

Nonmetric

Metric

Nonmetric

Metric

Nonmetric

What is the
measurement
scale of the
predictor
variable?

Canonical
correlation
analysis
with dummy
variables*

Multiple regression
(Chapter 4)

Conjoint analysis
(Chapter 7)

Multiple
discriminant
analysis
(Chapter 5)

Linear probability
models
(Chapter 5)

Legend

Decision
Point

Multidimensional
scaling
(Chapter 9)

Correspondence
analysis
(Chapter 9)

Metric

Nonmetric

Canonical
correlation
analysis*

Multivariate
analysis
of variance
(Chapter 6)

Multivariate
Technique
Selected

* Additional materials on this subject are available on the web www.prenhall.com/hair

Sampling Design and sample size

- Roscoe(1975) proposes the following rule of thumb for determining sample size

1. Sample size larger than 30 and less than 500 are appropriate for most research.

2. Where sample are to broken into subsamples(male/ females, juniors/seniors etc) a minimum sample size of 30 for each category is necessary.

3. In multivariate research the sample size should be several times(preferably 10 times or more) as large as the number of variables in the study.

4. For simple experimental research with tight experimental controls, successful research is possible with samples as small as 10 to 20.

Thank you