

# Social Media Marketing Management

## 社會媒體行銷管理

# 探索性因素分析

## (Exploratory Factor Analysis)

1002SMMM11

TLMXJ1A

Tue 12,13,14 (19:20-22:10) D325

Min-Yuh Day

戴敏育

Assistant Professor

專任助理教授

Dept. of Information Management, Tamkang University

淡江大學 資訊管理學系

<http://mail.tku.edu.tw/myday/>

2013-05-28

# 課程大綱 (Syllabus)

週次	日期	內容 (Subject/Topics)
1	102/02/19	社會媒體行銷管理課程介紹 (Course Orientation of Social Media Marketing Management)
2	102/02/26	社群網路 (Social Media: Facebook, Youtube, Blog, Microblog)
3	102/03/05	社群網路行銷 (Social Media Marketing)
4	102/03/12	行銷管理 (Marketing Management)
5	102/03/19	社群網路服務與資訊系統理論 (Theories of Social Media Services and Information Systems)
6	102/03/26	行銷理論 (Marketing Theories)
7	102/04/02	教學行政觀摩日 (Off-campus study)
8	102/04/09	行銷管理論文研討 (Paper Reading on Marketing Management)
9	102/04/16	社群網路行為研究 (Behavior Research on Social Media)

# 課程大綱 (Syllabus)

週次	日期	內容 (Subject/Topics)
10	102/04/23	期中報告 (Midterm Presentation)
11	102/04/30	社群網路商業模式 [Invited Speaker: Dr. Rick Cheng-Yu Lu] (Business Models and Issues of Social Media)
12	102/05/07	社群網路策略 (Strategy of Social Media)
13	102/05/14	社群口碑與社群網路探勘 (Social Word-of-Mouth and Web Mining on Social Media)
14	102/05/21	社群網路論文研討 (Paper Reading on Social Media)
15	102/05/28	探索性因素分析 (Exploratory Factor Analysis)
16	102/06/04 (> 6/01)	確認性因素分析 (Confirmatory Factor Analysis)
17	102/06/11 (> 6/04)	期末報告1 (Term Project Presentation 1)
18	102/06/18 (> 6/11)	期末報告2 (Term Project Presentation 2)

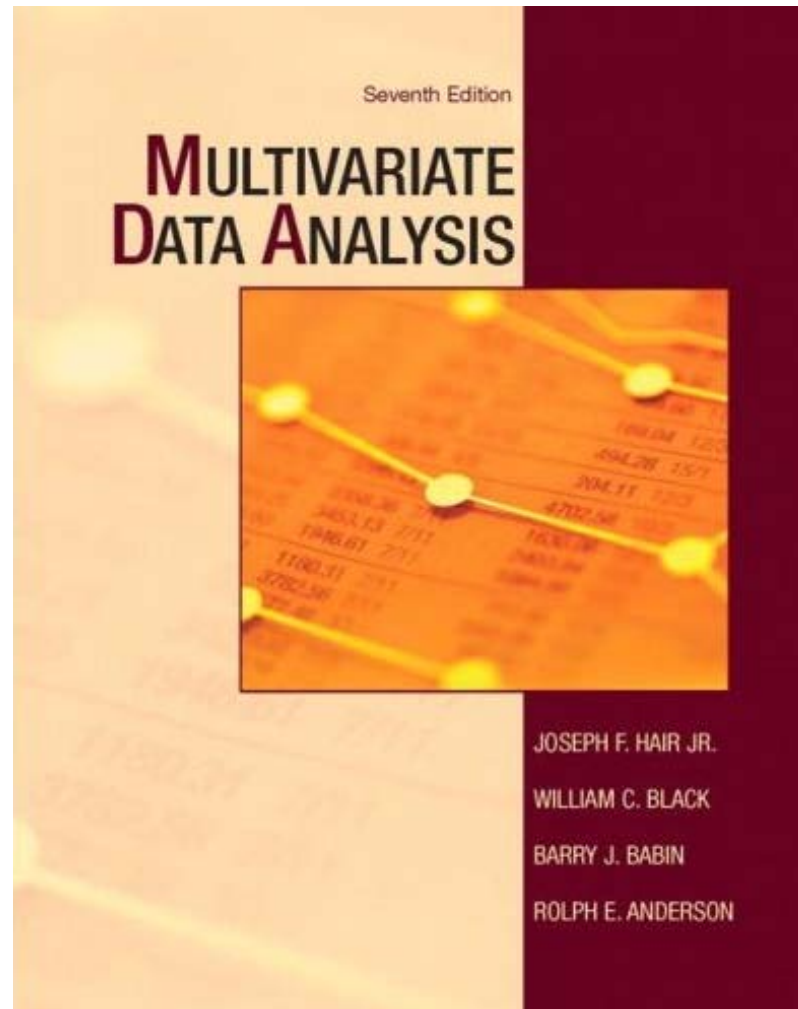
# Outline

- Seven stages of applying factor analysis
- Exploratory Factor Analysis (EFA) vs. Confirmatory Factor Analysis (CFA)
- Identify the differences between component analysis and common factor analysis models
- How to determine the number of factors to extract
- How to name a factor

# Types of Factor Analysis

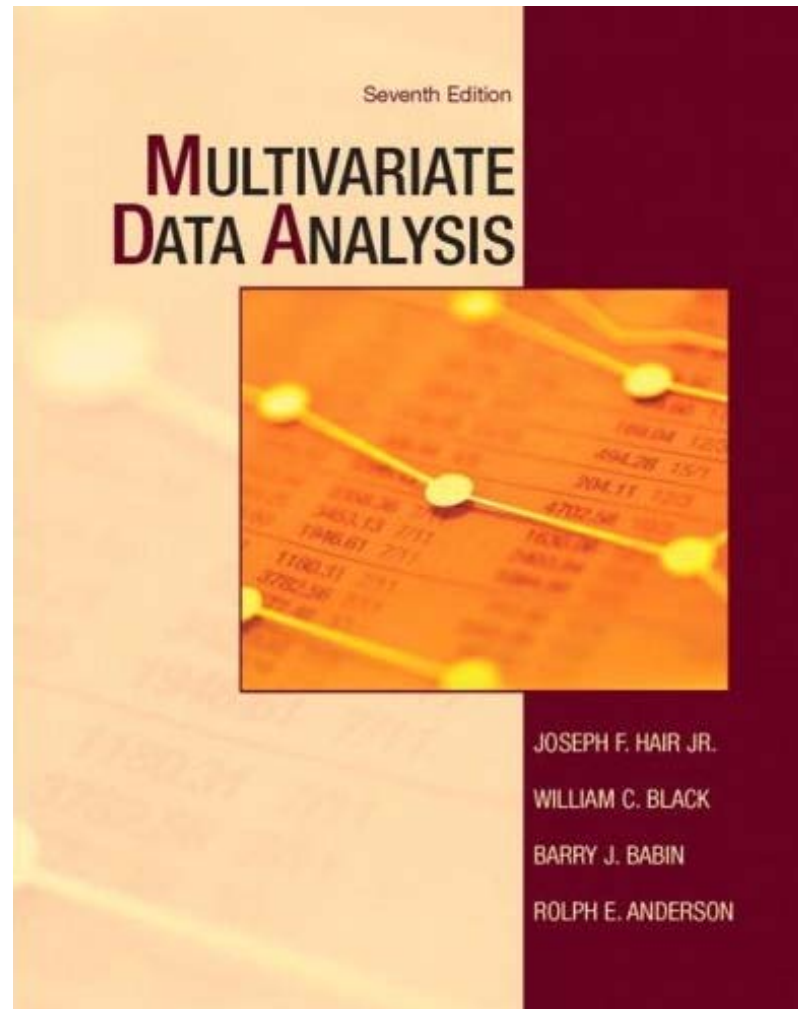
- **Exploratory Factor Analysis (EFA)**
  - is used to discover the factor structure of a construct and examine its reliability.  
It is **data driven**.
- **Confirmatory Factor Analysis (CFA)**
  - is used to confirm the fit of the hypothesized factor structure to the observed (sample) data.  
It is **theory driven**.

**Joseph F. Hair, William C. Black, Barry J. Babin, Rolph E. Anderson,  
Multivariate Data Analysis, 7th Edition,  
Prentice Hall, 2009**



# Chapter 3

## Exploratory Factor Analysis



# Exploratory Factor Analysis (EFA)

- Definition
  - Exploratory factor analysis (EFA) is an **interdependence** technique whose primary purpose is to define the underlying **structure** among the **variables** in the analysis.



# Exploratory Factor Analysis (EFA)

- Examines the **interrelationships** among a large number of **variables** and then attempts to explain them in terms of their **common** underlying dimensions.
- These **common** underlying dimensions are referred to as **factors**.
- A **summarization** and **data reduction** technique that does not have independent and dependent variables, but is an interdependence technique in which all variables are considered simultaneously.

# Correlation Matrix for Store Image Elements

	<b>V<sub>1</sub></b>	<b>V<sub>2</sub></b>	<b>V<sub>3</sub></b>	<b>V<sub>4</sub></b>	<b>V<sub>5</sub></b>	<b>V<sub>6</sub></b>	<b>V<sub>7</sub></b>	<b>V<sub>8</sub></b>	<b>V<sub>9</sub></b>
<b>V<sub>1</sub> Price Level</b>	1.00								
<b>V<sub>2</sub> Store Personnel</b>	.427	1.00							
<b>V<sub>3</sub> Return Policy</b>	.302	.771	1.00						
<b>V<sub>4</sub> Product Availability</b>	.470	.497	.427	1.00					
<b>V<sub>5</sub> Product Quality</b>	.765	.406	.307	.472	1.00				
<b>V<sub>6</sub> Assortment Depth</b>	.281	.445	.423	.713	.325	1.00			
<b>V<sub>7</sub> Assortment Width</b>	.354	.490	.471	.719	.378	.724	1.00		
<b>V<sub>8</sub> In-Store Service</b>	.242	.719	.733	.428	.240	.311	.435	1.00	
<b>V<sub>9</sub> Store Atmosphere</b>	.372	.737	.774	.479	.326	.429	.466	.710	1.00

# Correlation Matrix of Variables After Grouping Using Factor Analysis

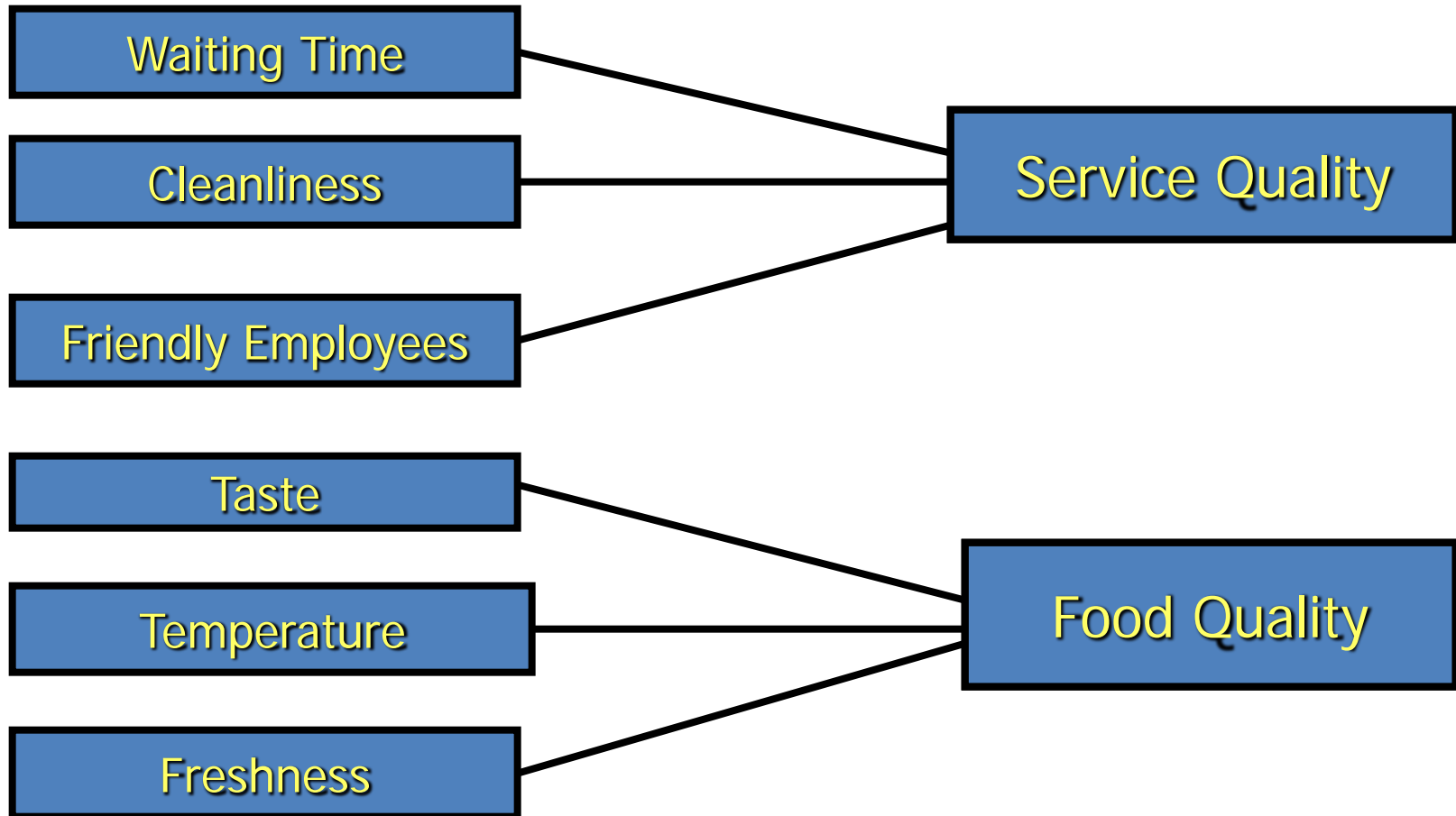
	V <sub>3</sub>	V <sub>8</sub>	V <sub>9</sub>	V <sub>2</sub>	V <sub>6</sub>	V <sub>7</sub>	V <sub>4</sub>	V <sub>1</sub>	V <sub>5</sub>
V <sub>3</sub> Return Policy	1.00								
V <sub>8</sub> In-store Service	.733	1.00							
V <sub>9</sub> Store Atmosphere	.774	.710	1.00						
V <sub>2</sub> Store Personnel	.741	.719	.787	1.00					
V <sub>6</sub> Assortment Depth	.423	.311	.429	.445	1.00				
V <sub>7</sub> Assortment Width	.471	.435	.468	.490	.724	1.00			
V <sub>4</sub> Product Availability	.427	.428	.479	.497	.713	.719	1.00		
V <sub>1</sub> Price Level	.302	.242	.372	.427	.281	.354	.470	1.00	
V <sub>5</sub> Product Quality	.307	.240	.326	.406	.325	.378	.472	.765	1.00

Shaded areas represent variables likely to be grouped together by factor analysis.

# Application of Factor Analysis to a Fast-Food Restaurant

## Variables

## Factors



# Factor Analysis Decision Process

- Stage 1: Objectives of Factor Analysis
- Stage 2: Designing a Factor Analysis
- Stage 3: Assumptions in Factor Analysis
- Stage 4: Deriving Factors and Assessing Overall Fit
- Stage 5: Interpreting the Factors
- Stage 6: Validation of Factor Analysis
- Stage 7: Additional uses of Factor Analysis Results

# Stage 1: Objectives of Factor Analysis

1. Is the objective exploratory or confirmatory?
2. Specify the unit of analysis.
3. Data summarization and/or reduction?
4. Using factor analysis with other techniques.

# Factor Analysis Outcomes

- Data summarization
  - derives underlying dimensions that, when interpreted and understood, describe the data in a much smaller number of concepts than the original individual variables.
- Data reduction
  - extends the process of data summarization by deriving an empirical value (factor score or summated scale) for each dimension (factor) and then substituting this value for the original values.

# Stage 2: Designing a Factor Analysis

- Three Basic Decisions:
  1. Calculation of input data – R vs. Q analysis.
  2. Design of study in terms of number of variables, measurement properties of variables, and the type of variables.
  3. Sample size necessary.



# Rules of Thumb 3–1

## Factor Analysis Design

- Factor analysis is performed most often only on metric variables, although specialized methods exist for the use of dummy variables. A small number of “dummy variables” can be included in a set of metric variables that are factor analyzed.
- If a study is being designed to reveal factor structure, strive to have **at least five variables** for each proposed factor.
- For sample size:
  - the sample must have more observations than variables.
  - the **minimum absolute sample size** should be **50 observations**.
- Maximize the number of observations per variable, with a minimum of **five** and hopefully at least **ten** observations per variable.

# Stage 3: Assumptions in Factor Analysis

- Three Basic Decisions
  1. Calculation of input data – R vs. Q analysis.
  2. Design of study in terms of number of variables, measurement properties of variables, and the type of variables.
  3. Sample size required.

# Assumptions

- Multicollinearity
  - Assessed using MSA (measure of sampling adequacy).
    - The MSA is measured by the Kaiser-Meyer-Olkin (KMO) statistic. As a measure of sampling adequacy, the KMO predicts if data are likely to factor well based on correlation and partial correlation. KMO can be used to identify which variables to drop from the factor analysis because they lack multicollinearity.
    - There is a KMO statistic for each individual variable, and their sum is the KMO overall statistic. KMO varies from 0 to 1.0. Overall KMO should be .50 or higher to proceed with factor analysis. If it is not, remove the variable with the lowest individual KMO statistic value one at a time until KMO overall rises above .50, and each individual variable **KMO is above .50.**
- Homogeneity of sample factor solutions

# Rules of Thumb 3–2

## Testing Assumptions of Factor Analysis

- There must be a strong conceptual foundation to support the assumption that a structure does exist before the factor analysis is performed.
- A statistically significant **Bartlett's test of sphericity (sig. < .05)** indicates that sufficient correlations exist among the variables to proceed.
- **Measure of Sampling Adequacy (MSA) values must exceed .50** for both the overall test and each individual variable.  
**Variables with values less than .50 should be omitted from the factor analysis one at a time**, with the smallest one being omitted each time.

# Stage 4: Deriving Factors and Assessing Overall Fit

- Selecting the factor extraction method – **common** vs. **component** analysis.
- Determining the number of factors to represent the data.

# Extraction Decisions

- Which method?
  - Principal Components Analysis
  - Common Factor Analysis
- How to rotate?
  - Orthogonal or Oblique rotation

# Extraction Method Determines the Types of Variance Carried into the Factor Matrix

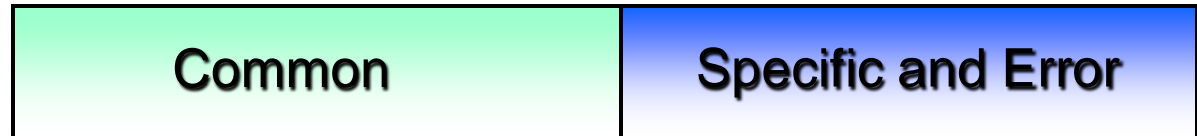
Diagonal Value

Variance

Unity (1)



Communality



Variance extracted



Variance not used

# Principal Components vs. Common?

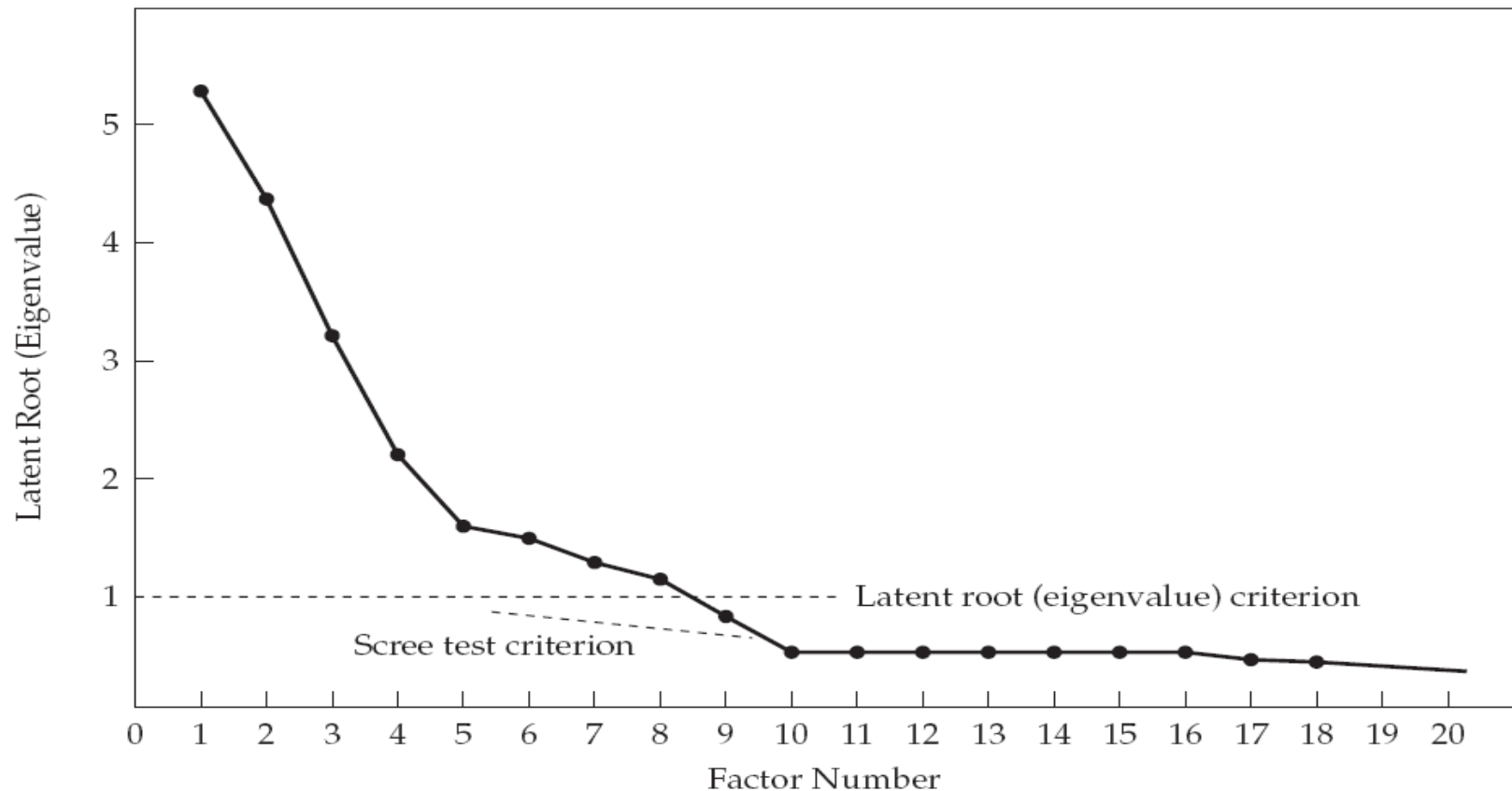
- Two Criteria
  - Objectives of the factor analysis.
  - Amount of prior knowledge about the variance in the variables.



# Number of Factors?

- A Priori Criterion
- Latent Root Criterion
- Percentage of Variance
- Scree Test Criterion

# Eigenvalue Plot for Scree Test Criterion



**FIGURE 3-6** Eigenvalue Plot for Scree Test Criterion

# Rules of Thumb 3–3

## Choosing Factor Models and Number of Factors

- Although both component and common factor analysis models yield similar results in common research settings (30 or more variables or communalities of **.60** for most variables):
  - the component analysis model is most appropriate when data reduction is paramount.
  - the common factor model is best in well-specified theoretical applications.
- Any decision on the number of factors to be retained should be based on several considerations:
  - use of several stopping criteria to determine the initial number of factors to retain.
  - **Factors With Eigenvalues greater than 1.0.**
  - A pre-determined number of factors based on research objectives and/or prior research.
  - Enough factors to meet a **specified percentage of variance explained, usually 60% or higher.**
  - Factors shown by the scree test to have substantial amounts of common variance (i.e., factors before inflection point).
  - More factors when there is heterogeneity among sample subgroups.
- Consideration of several alternative solutions (one more and one less factor than the initial solution) to ensure the best structure is identified.

# Processes of Factor Interpretation

- Estimate the Factor Matrix
- Factor Rotation
- Factor Interpretation
- Respecification of factor model, if needed, may involve . . .
  - Deletion of variables from analysis
  - Desire to use a different rotational approach
  - Need to extract a different number of factors
  - Desire to change method of extraction

# Rotation of Factors

- Factor rotation

- the reference axes of the factors are turned about the origin until some other position has been reached.

Since unrotated factor solutions extract factors based on how much variance they account for, with each subsequent factor accounting for less variance.

The ultimate effect of rotating the factor matrix is to redistribute the variance from earlier factors to later ones to achieve a simpler, theoretically more meaningful factor pattern.

# Two Rotational Approaches

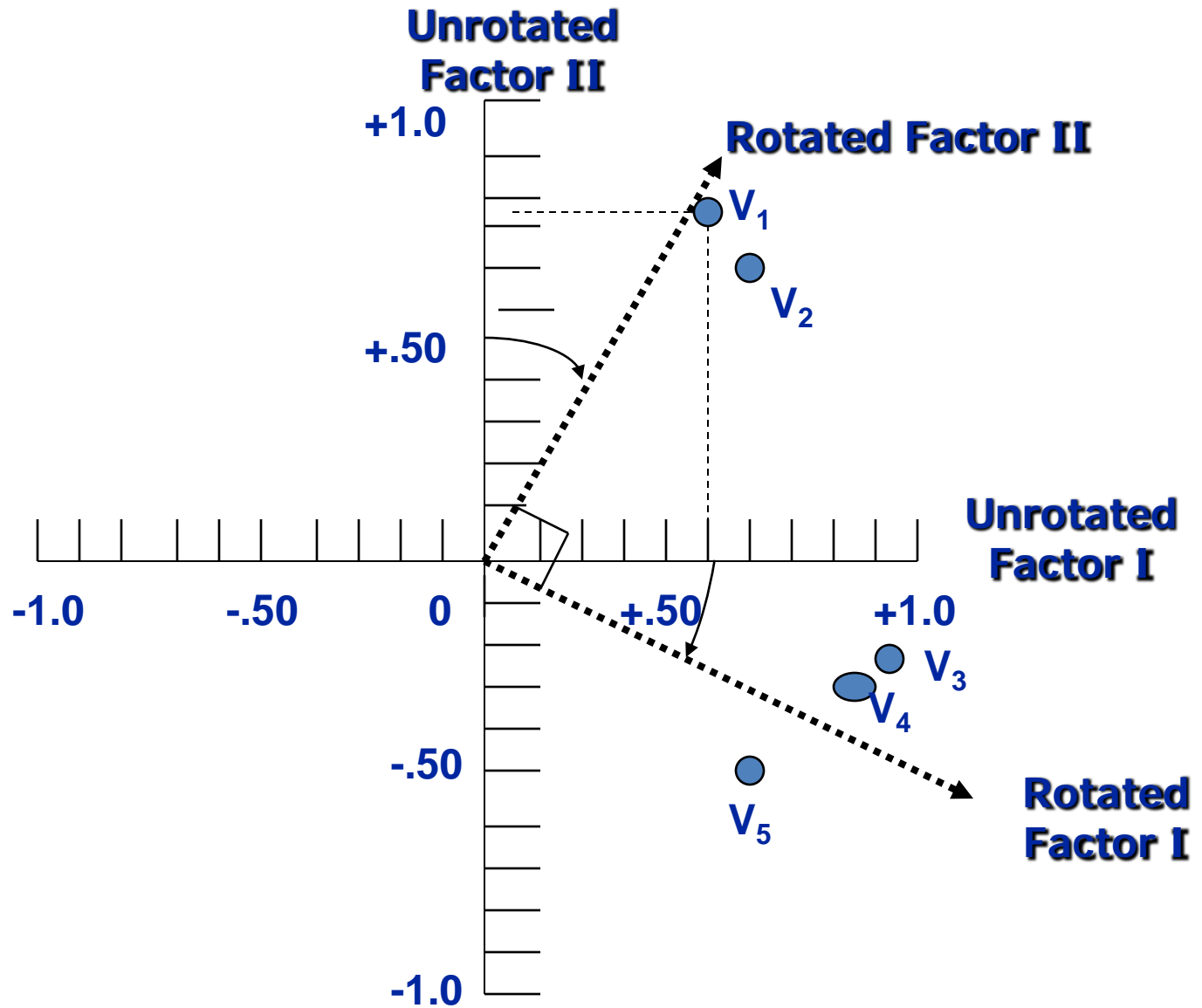
## 1. Orthogonal

- axes are maintained at 90 degrees.

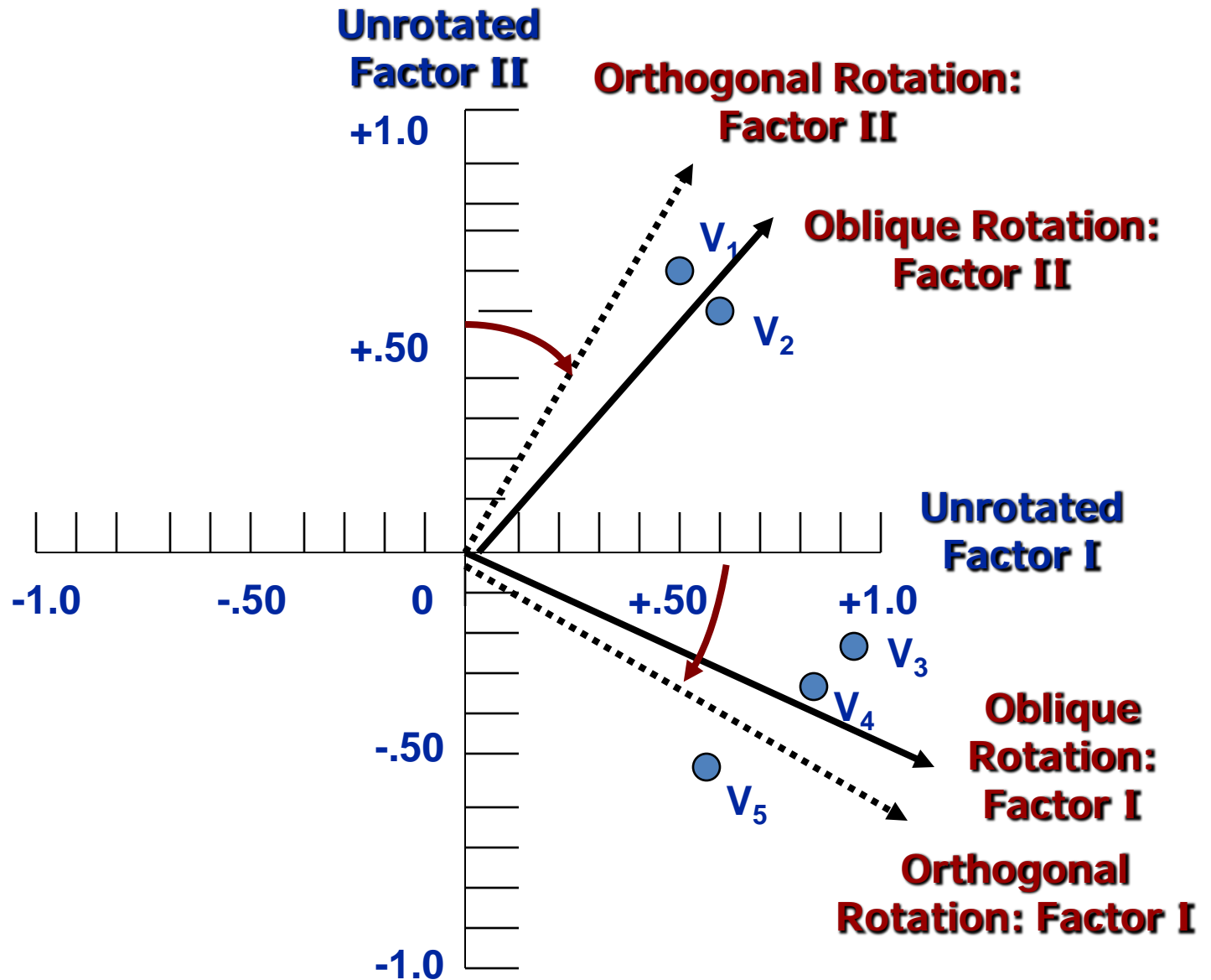
## 2. Oblique

- axes are not maintained at 90 degrees.

# Orthogonal Factor Rotation



# Oblique Factor Rotation





# Orthogonal Rotation Methods

- Quartimax (simplify rows)
- Varimax (simplify columns)
- Equimax (combination)

# Rules of Thumb 3–4

## Choosing Factor Rotation Methods

- Orthogonal rotation methods
  - are the **most widely used rotational methods**.
  - are The preferred method when the research goal is **data reduction** to either a smaller number of variables or a set of uncorrelated measures for subsequent use in other multivariate techniques.
- Oblique rotation methods
  - best suited to the goal of obtaining several **theoretically meaningful factors or constructs** because, realistically, very few constructs in the “real world” are uncorrelated

# Which Factor Loadings Are Significant?

- Customary Criteria = Practical Significance.
- Sample Size & Statistical Significance.
- Number of Factors ( $\uparrow = >$ ) and/or Variables ( $\uparrow = <$ ).

# Guidelines for Identifying Significant Factor Loadings Based on Sample Size

Factor Loading

Sample Size Needed  
for Significance\*

---

.30	350
.35	250
.40	200
.45	150
.50	120
.55	100
.60	85
.65	70
.70	60
.75	50

---

\*Significance is based on a .05 significance level ( $\alpha$ ), a power level of 80 percent, and standard errors assumed to be twice those of conventional correlation coefficients.

# Rules of Thumb 3–5

## Assessing Factor Loadings

- While factor loadings of +.30 to +.40 are **minimally acceptable**, values greater than +.50 are considered necessary for **practical significance**.
- To be considered significant:
  - A smaller loading is needed given either a larger sample size, or a larger number of variables being analyzed.
  - A larger loading is needed given a factor solution with a larger number of factors, especially in evaluating the loadings on later factors.
- Statistical tests of significance for factor loadings are generally very conservative and should be considered only as starting points needed for including a variable for further consideration.

# Stage 5: Interpreting the Factors

- Selecting the factor extraction method – **common** vs. **component** analysis.
- Determining the number of factors to represent the data.

# Interpreting a Factor Matrix:

1. Examine the factor matrix of loadings.
2. Identify the highest loading across all factors for each variable.
3. Assess communalities of the variables.
4. Label the factors.

# Rules of Thumb 3–6

## Interpreting The Factors

- An optimal structure exists when all variables have **high loadings only on a single factor**.
- **Variables that cross-load (load highly on two or more factors) are usually deleted** unless theoretically justified or the objective is strictly data reduction.
- Variables should generally have **communalities of greater than .50** to be retained in the analysis.
- Respecification of a factor analysis can include options such as:
  - deleting a variable(s),
  - changing rotation methods, and/or
  - increasing or decreasing the number of factors.



# Stage 6: Validation of Factor Analysis

- Confirmatory Perspective.
- Assessing Factor Structure Stability.
- Detecting Influential Observations.

# Stage 7: Additional Uses of Factor Analysis Results

- Selecting Surrogate Variables
- Creating Summated Scales
- Computing Factor Scores

# Rules of Thumb 3–7

## Summated Scales

- A summated scale is only as good as the items used to represent the construct. While it may pass all empirical tests, it is useless without theoretical justification.
- Never create a summated scale without first assessing its unidimensionality with exploratory or confirmatory factor analysis.
- Once a scale is deemed unidimensional, its reliability score, as measured by Cronbach's alpha:
  - **should exceed a threshold of .70**, although a .60 level can be used in exploratory research.
  - the threshold should be raised as the number of items increases, especially as the number of items approaches 10 or more.
- With reliability established, validity should be assessed in terms of:
  - **convergent validity** = scale correlates with other like scales.
  - **discriminant validity** = scale is sufficiently different from other related scales.
  - **nomological validity** = scale “predicts” as theoretically suggested.

# Rules of Thumb 3–8

## Representing Factor Analysis In Other Analyses

- The single surrogate variable:
  - Advantages: simple to administer and interpret.
  - Disadvantages:
    - does not represent all “facets” of a factor
    - prone to measurement error.
- Factor scores:
  - Advantages:
    - represents all variables loading on the factor,
    - best method for complete data reduction.
    - Are by default orthogonal and can avoid complications caused by multicollinearity.
  - Disadvantages:
    - interpretation more difficult since all variables contribute through loadings
    - Difficult to replicate across studies.

# Rules of Thumb 3–8 (cont.)

## Representing Factor Analysis In Other Analyses

- Summated scales:
  - Advantages:
    - compromise between the surrogate variable and factor score options.
    - reduces measurement error.
    - represents multiple facets of a concept.
    - easily replicated across studies.
  - Disadvantages:
    - includes only the variables that load highly on the factor and excludes those having little or marginal impact.
    - not necessarily orthogonal.
    - Require extensive analysis of reliability and validity issues.

## Description of HBAT Primary Database Variables

Variable Description	Variable Type
<u>Data Warehouse Classification Variables</u>	
X1      Customer Type	nonmetric
X2      Industry Type	nonmetric
X3      Firm Size	nonmetric
X4      Region	nonmetric
X5      Distribution System	nonmetric
<u>Performance Perceptions Variables</u>	
X6      Product Quality	metric
X7      E-Commerce Activities/Website	metric
X8      Technical Support	metric
X9      Complaint Resolution	metric
X10     Advertising	metric
X11     Product Line	metric
X12     Salesforce Image	metric
X13     Competitive Pricing	metric
X14     Warranty & Claims	metric
X15     New Products	metric
X16     Ordering & Billing	metric
X17     Price Flexibility	metric
X18     Delivery Speed	metric
<u>Outcome/Relationship Measures</u>	
X19     Satisfaction	metric
X20     Likelihood of Recommendation	metric
X21     Likelihood of Future Purchase	metric
X22     Current Purchase/Usage Level	metric
X23     Consider Strategic Alliance/Partnership in Future	nonmetric

# Rotated Component Matrix

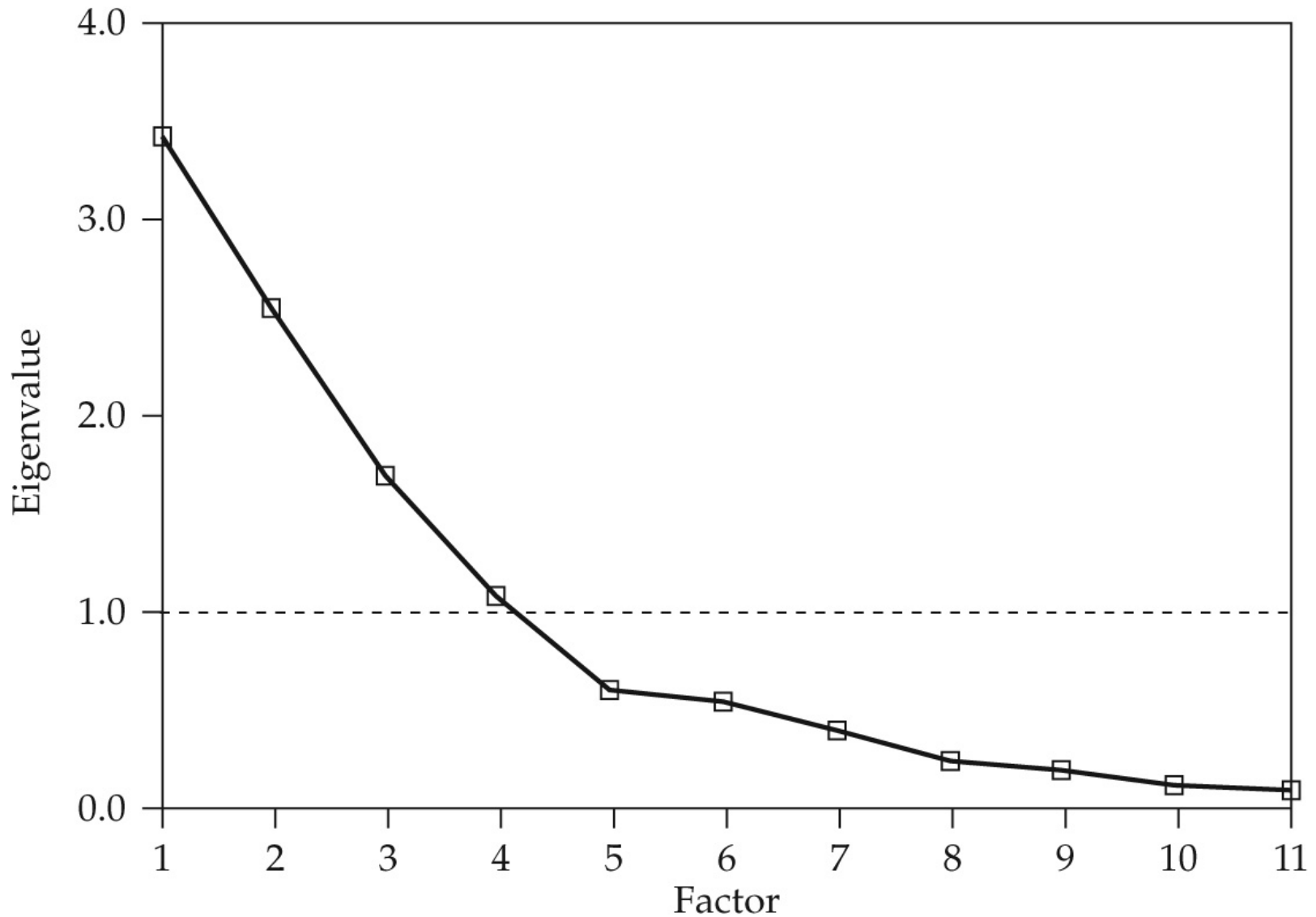
## “Reduced Set” of HBAT Perceptions Variables

	Component				Communality
	1	2	3	4	
<b>X9 – Complaint Resolution</b>	<b>.933</b>				<b>.890</b>
<b>X18 – Delivery Speed</b>	<b>.931</b>				<b>.894</b>
<b>X16 – Order &amp; Billing</b>	<b>.886</b>				<b>.806</b>
<b>X12 – Salesforce Image</b>		<b>.898</b>			<b>.860</b>
<b>X7 – E-Commerce Activities</b>		<b>.868</b>			<b>.780</b>
<b>X10 – Advertising</b>		<b>.743</b>			<b>.585</b>
<b>X8 – Technical Support</b>			<b>.940</b>		<b>.894</b>
<b>X14 – Warranty &amp; Claims</b>			<b>.933</b>		<b>.891</b>
<b>X6 – Product Quality</b>				<b>.892</b>	<b>.798</b>
<b>X13 – Competitive Pricing</b>				<b>-.730</b>	<b>.661</b>
<b>Sum of Squares</b>	<b>2.589</b>	<b>2.216</b>	<b>1.846</b>	<b>1.406</b>	<b>8.057</b>
<b>Percentage of Trace</b>	<b>25.893</b>	<b>22.161</b>	<b>18.457</b>	<b>14.061</b>	<b>80.572</b>

**Extraction Method: Principal Component Analysis.**

**Rotation Method: Varimax.**

# Scree Test for HBAT Component Analysis





# Summary

1. What are the major uses of factor analysis?
2. What is the difference between component analysis and common factor analysis?
3. Is rotation of factors necessary?
4. How do you decide how many factors to extract?
5. What is a significant factor loading?
6. How and why do you name a factor?
7. Should you use factor scores or summated ratings in follow-up analyses?

蕭文龍,  
多變量分析最佳入門實用書--SPSS+LISREL, 第二版,  
碁峰資訊, 2009



PChome線上購物 - 24小時購物

shopping.pchome.com.tw/?mod=item&func=exhibit&IT\_NO=DJAA1L-A41336032&c=A05

線上購物 | 24h購物 | 書店 | 軟體 | 全球購物 運費 大陸RMB¥20 美國US\$10 澳洲US\$10...

購物車 故障通報 廠商合作 PChome

購物滿\$490 免運費

3C 食·日用

週邊 生活  
數位 保養  
NB 衣包錶  
家電 全球購物

音樂CD DVD/BD 語言考試 電腦/攝影 財經企管 文學小說 親子教育 休閒娛樂

自然人文 健康生活 雜誌零售

買不到通知我

顧客中心 查訂單 退貨 客服

## 計算機概論

PCHome > 線上購物 > 24h購物 > 電腦/攝影 > 計算機概論

加入追蹤 故障通報

多變量分析最佳入門實用書--SPSS+LISREL(第二版)(平裝) (商品編號: DJAA1L-A41336032)

24h \$0 (0) 結帳 >

· 全台灣24小時到貨, 遲到給100  
· 非北北基22:00~10:00間下單, 收貨人資訊不完整、新客、大材積商品、ATM或ibon付款者不在此限  
· 365天全年無休, 週末假日照常出貨 [一說明](#)

### 電腦/攝影

- 大專理論用書
  - 計算機概論 大專用書
- 數位攝影
  - 攝影達人專區
  - 影像編修高手
- 影音/繪圖
  - Photoshop專區
  - 影像繪圖
  - 設計師專區
  - 影音剪輯
- 多媒體應用
  - 多媒體網頁設計

多變量分析  
最佳入門實用書 第二版  
SPSS+LISREL

● 本書通過 Scientific Software International (SSI) LISREL 原廠審核通過, 成為 LISREL 原廠推薦的第四本華文書, 相關網址: <http://www.ssicentral.com/cn/books.htm#sem>  
● 本書可作為 Hair (2006) Multivariate Data Analysis 一書的最佳輔助參考書籍

碁峯 GOTOP EGT

蕭文龍

### 多變量分析最佳入門實用書--SPSS+LISREL(第二版)(平裝)

通過Scientific Software International (SSI) LISREL原廠審核通過, 成為LISREL原廠推薦的第四本華文書!

作者: 蕭文龍  
ISBN: 9789861817347/9861817344  
出版社: 碁峯資訊  
出版日期: 2009/06/09

建議售價\$700

分期價: 除不盡餘數於第一期收取

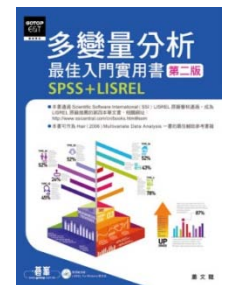
網路價 \$553 賣貴通報

VISA 貨到付款 ibon 說明

3期0利率	每期\$184	27家	12期分期	每期\$48	11家
-------	---------	-----	-------	--------	-----

# 蕭文龍, 多變量分析最佳入門實用書--SPSS+LISREL, 第二版, 碁峰資訊, 2009

- 本書通過Scientific Software International (SSI) LISREL原廠審核通過，成為LISREL原廠推薦的第四本華文書，  
相關網址：<http://www.ssicentral.com/cn/books.html#sem>
- 本書可作為Hair（2006）Multivariate Data Analysis一書的最佳輔助參考書籍
- 從實用的角度出發，完整介紹社會科學概念、統計軟體的運用以及統計分析，協助學習者完成量化的研究及其相關專題或論文。
- 內容整合了社會科學概念、量化研究、量表發展與統計分析。
- 文中納入第二代統計技術，包括結構方程模式(SEM)、LISREL基本操作SEM結構方程模式範例與SEM結構方程模式實例。
- 特別介紹研究流程、論文結構與研究範例、EndNote書目管理軟體使用說明、LISREL和Nvivo軟體的取得與說明。
- 隨書光碟附贈LISREL For Windows學生版



# 蕭文龍, 多變量分析最佳入門實用書--SPSS+LISREL, 第二版, 碁峰資訊, 2009

- Ch01 社會科學的研究與數量方法的基礎
- Ch02 SPSS的基本操作
- Ch03 檢視資料與敘述性統計
- Ch04 相關分析
- Ch05 卡方檢定
- Ch06 平均數比較
- Ch07 因素分析
- Ch08 迴歸分析
- Ch09 區別分析與邏輯迴歸
- Ch10 單變量變異數分析
- Ch11 多變量變異數分析
- Ch12 典型相關



# 蕭文龍, 多變量分析最佳入門實用書--SPSS+LISREL, 第二版, 碁峰資訊, 2009

- Ch13 量表的發展、信度和效度
- Ch14 SEM結構方程模式
- Ch15 LISREL的基本操作
- Ch16 結構方程模式的學習範例
- Ch17 結構方程模式的學習範例進階
- Ch18 SEM結構方程模式實例
- Ch19 聯合分析、多元尺度方法和集群分析
- Ch20 交互作用、中介和調節(干擾)效果之驗證
- Ch21 研究流程、論文結構與研究範例
- 附錄A 統計分配表
- 附錄B ENDNOTE書目管理軟體使用說明
- 附錄C 軟體的取得與說明LISREL



# References

- Joseph F. Hair, William C. Black, Barry J. Babin, Rolph E. Anderson (2009),  
Multivariate Data Analysis, 7th Edition, Prentice Hall
- 蕭文龍 (2009), 多變量分析最佳入門實用書--SPSS+LISREL, 第二版, 碁峰資訊
- 吳明隆 (2006), SPSS 統計應用學習實務：問卷分析與應用統計, 三版, 知城數位科技