Social Media Marketing Research
社會媒體行銷研究

Measurement and Scaling

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/
2012-03-29

1

課程大綱 (Syllabus)

週次　日期　內容（Subject/Topics）

1　101/02/16　Course Orientation of Social Media Marketing
　　　　　　　 Research

2　101/02/23　Social Media: Facebook, Youtube, Blog, Microblog

3　101/03/01　Social Media Marketing

4　101/03/08　Marketing Research

5　101/03/15　Marketing Theories

6　101/03/22　Measuring the Construct

7　101/03/29　Measurement and Scaling

8　101/04/05　教學行政觀摩日 (--No Class--)

9　101/04/12　Paper Reading and Discussion

課程大綱 (Syllabus)

週次　日期　內容（Subject/Topics）

10　101/04/19　Midterm Presentation

11　101/04/26　Exploratory Factor Analysis

12　101/05/03　Paper Reading and Discussion

13　101/05/10　Confirmatory Factor Analysis

14　101/05/17　Paper Reading and Discussion

15　101/05/24　Communicating the Research Results

16　101/05/31　Paper Reading and Discussion

17　101/06/07　Term Project Presentation 1

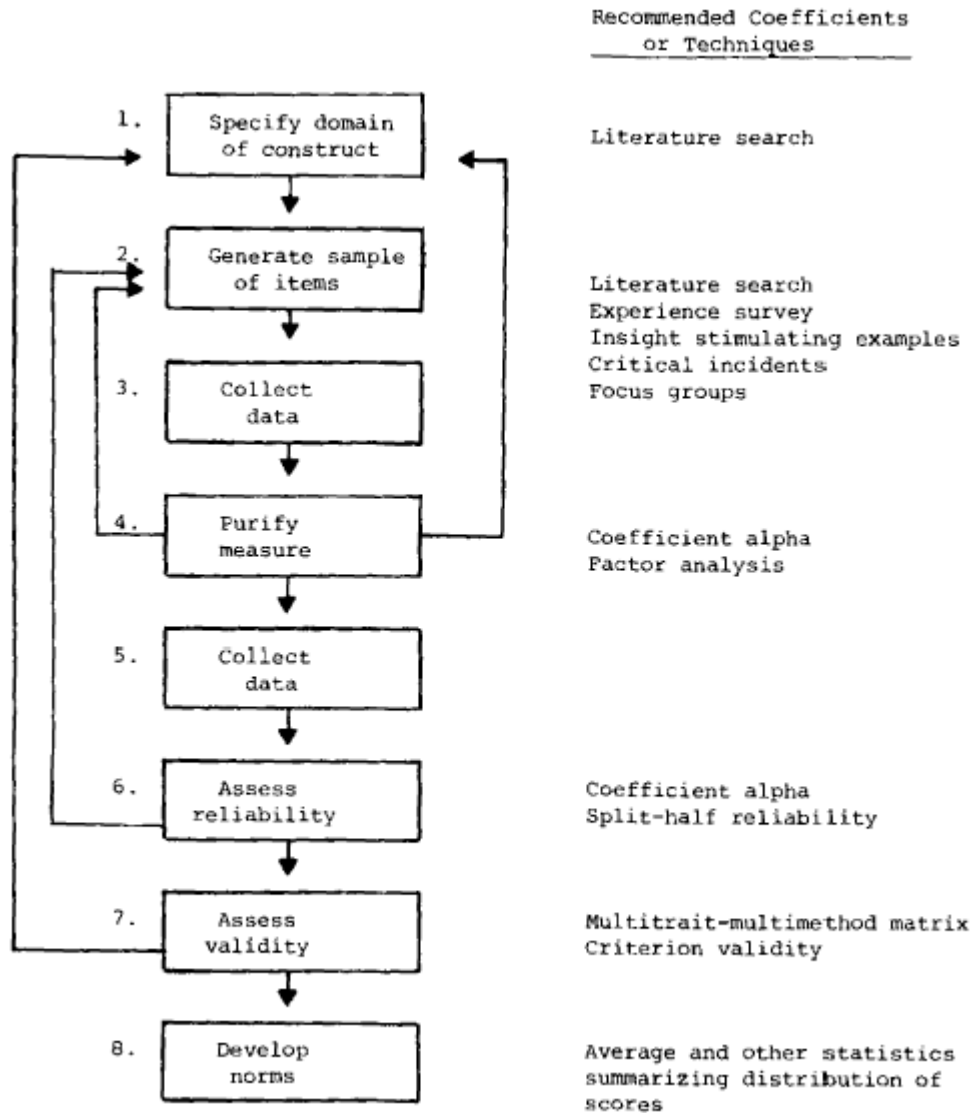18　101/06/14　Term Project Presentation 2

Outline

- A paradigm for developing better measures of marketing constructs

- Current practice in scale development

- The linkage among attitudes, behavior, and marketing effectiveness
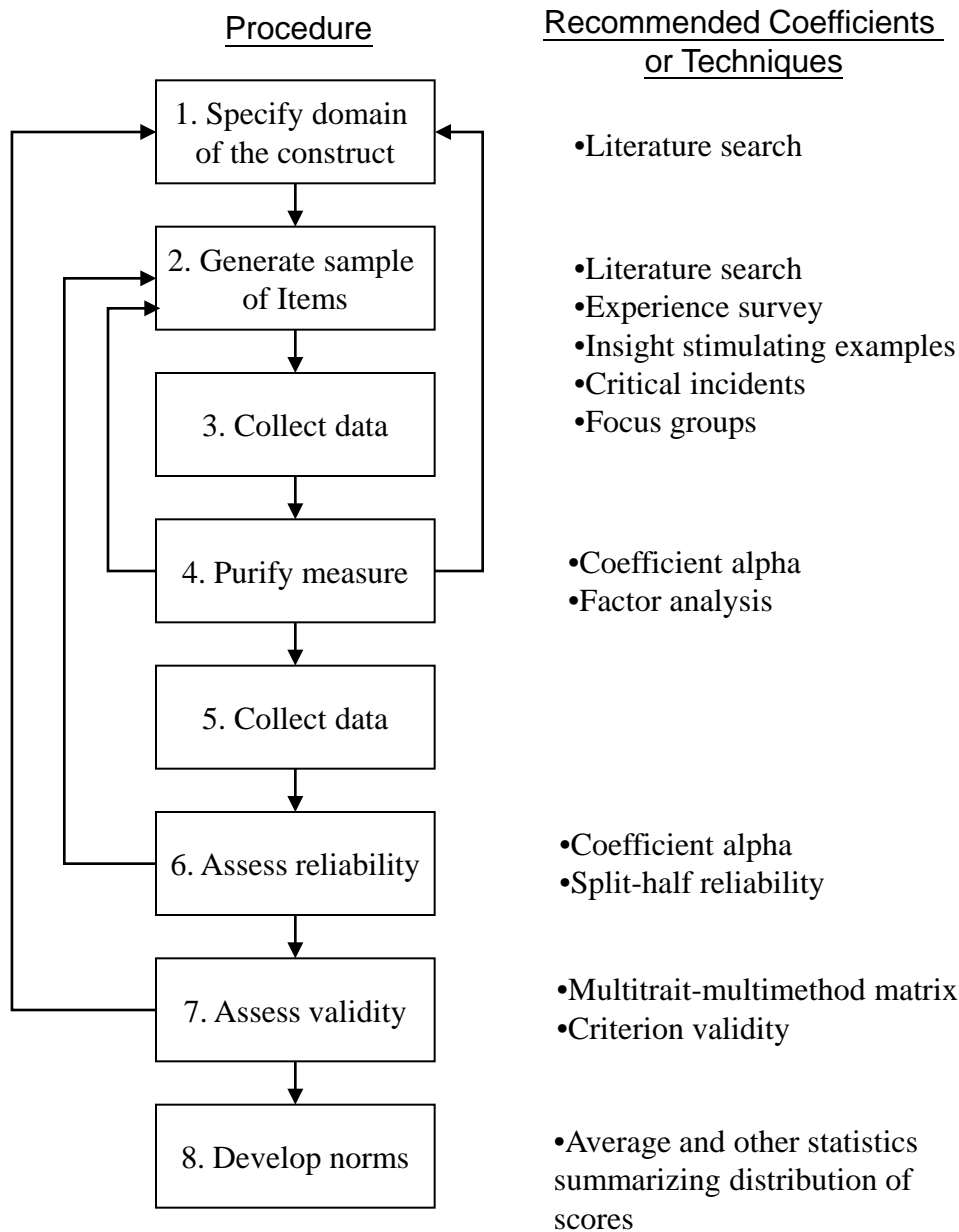
- Measurement Scales

A paradigm for developing better measures of marketing constructs

Churchill, G. A., Jr., (1979),
A paradigm for developing better measures of marketing
constructs.
Journal of Marketing Research, 16(February), 64-73.

Suggested Procedure for Developing Better Measures

Recommended Coefficients or Techniques

1. Specify domain of construct — Literature search

2. Generate sample of items

3. Collect data — Literature search, Experience survey, Insight stimulating examples, Critical incidents, Focus groups

4. Purify measure — Coefficient alpha, Factor analysis

5. Collect data

6. Assess reliability — Coefficient alpha, Split-half reliability

7. Assess validity — Multitrait-multimethod matrix, Criterion validity

8. Develop norms — Average and other statistics summarizing distribution of scores

(Churchill, 1979)(A Paradigm for Developing Better Measures of marketing Constructs)

Suggested Procedure for Developing Better Measures (Churchill, 1979)

Procedure

Recommended Coefficients
or Techniques

1. Specify domain of the construct	•Literature search
2. Generate sample of Items	•Literature search •Experience survey •Insight stimulating examples •Critical incidents •Focus groups
3. Collect data	
4. Purify measure	•Coefficient alpha •Factor analysis
5. Collect data	
6. Assess reliability	•Coefficient alpha •Split-half reliability
7. Assess validity	•Multitrait-multimethod matrix •Criterion validity
8. Develop norms	•Average and other statistics summarizing distribution of scores

(Churchill, 1979)(A Paradigm for Developing Better Measures of marketing Constructs)

The Problem and Approach

- Developing measures which have desirable reliability and validity properties

- The process of measurement of operationalization involves "rules for assigning numbers to objects to represent quantities of attributes".

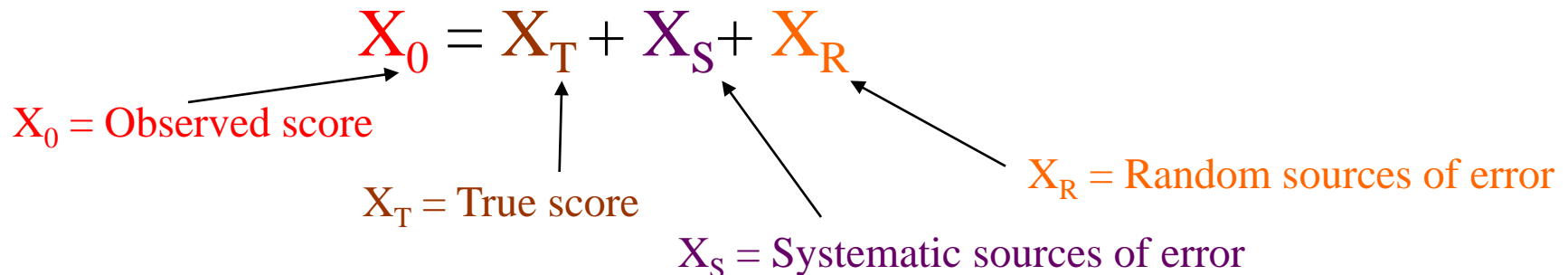- Consider some arbitrary construct, C, such as customer satisfaction.

$$X_0 = X_T + X_S + X_R$$

X_0 = Observed score

X_T = True score

X_S = Systematic sources of error

X_R = Random sources of error

Scale Development
Example from (Davis, 1989)

- Scale Development and Pretest
  - A step-by-step process was used to develop new multi-item scales having high reliability and validity.
  - The conceptual definitions of perceived usefulness and perceived ease of use, stated above, were used to generate 14 candidate items for each construct from past literature.
  - Pretest interviews were then conducted to assess the semantic content of the items. Those items that best fit the definitions of the constructs were retained, yielding 10 items for each construct.
  - Next, a field study (Study 1) of 112 users concerning two different interactive computer systems was conducted in order to assess the reliability and construct validity of the resulting scales.
  - The scales were further refined and streamlined to six items per construct. A lab study (Study 2) involving 40 participants and two graphics systems was then conducted.
  - Data from the two studies were then used to assess the relationship between usefulness, ease of use, and self-reported usage.

1. Specify Domain of the Construct

- Theoretical Definition
  - Perceived Usefulness:
    - The degree to which a person believes that using a particular system would enhance job performance
  - Perceived Ease of Use:
    - The degree to which a person believes that using a particular system would be free of effort.

Example from (Davis, 1989)

2. Generate Sample of Items

- Literature search
- Experience survey
- Insight stimulating examples
- Critical incidents
- Focus groups

2. Generate Sample of Items (Cont.)

Table 1. Initial Scale Items for Perceived Usefulness

1. My job would be difficult to perform without electronic mail.
2. Using electronic mail gives me greater control over my work.
3. Using electronic mail improves my job performance.
4. The electronic mail system addresses my job-related needs.
5. Using electronic mail saves me time.
6. Electronic mail enables me to accomplish tasks more quickly.
7. Electronic mail supports critical aspects of my job.
8. Using electronic mail allows me to accomplish more work than would otherwise be possible.
9. Using electronic mail reduces the time I spend on unproductive activities.
10. Using electronic mail enhances my effectiveness on the job.
11. Using electronic mail improves the quality of the work I do.
12. Using electronic mail increases my productivity.
13. Using electronic mail makes it easier to do my job.
14. Overall, I find the electronic mail system useful in my job.

Table 2. Initial Scale Items for Perceived Ease of Use

1. I often become confused when I use the electronic mail system.
2. I make errors frequently when using electronic mail.
3. Interacting with the electronic mail system is often frustrating.
4. I need to consult the user manual often when using electronic mail.
5. Interacting with the electronic mail system requires a lot of my mental effort.
6. I find it easy to recover from errors encountered while using electronic mail.
7. The electronic mail system is rigid and inflexible to interact with.
8. I find it easy to get the electronic mail system to do what I want it to do.
9. The electronic mail system often behaves in unexpected ways.
10. I find it cumbersome to use the electronic mail system.
11. My interaction with the electronic mail system is easy for me to understand.
12. It is easy for me to remember how to perform tasks using the electronic mail system.
13. The electronic mail system provides helpful guidance in performing tasks.
14. Overall, I find the electronic mail system easy to use.

Example from (Davis, 1989)

4. Purify the Measure

Table 3. Pretest Results: Perceived Usefulness

Old Item #	Item	Rank	New Item #	Cluster
1	Job Difficult Without	13		C
2	Control Over Work	9	2	
3	Job Performance	2	6	A
4	Addresses My Needs	12		C
5	Saves Me Time	11		B
6	Work More Quickly	7	3	B
7	Critical to My Job	5	4	C
8	Accomplish More Work	6	7	B
9	Cut Unproductive Time	10		B
10	Effectiveness	1	8	A
11	Quality of Work	3	1	A
12	Increase Productivity	4	5	B
13	Makes Job Easier	8	9	C
14	Useful	NA	10	NA

Table 4. Pretest Results: Perceived Ease of Use

Old Item #	Item	Rank	New Item #	Cluster
1	Confusing	7		B
2	Error Prone	13		
3	Frustrating	3	3	B
4	Dependence on Manual	9	(replace)	C
5	Mental Effort	5	7	B
6	Error Recovery	10		
7	Rigid & Inflexible	6	5	A
8	Controllable	1	4	A
9	Unexpected Behavior	11		
10	Cumbersome	2	1	A
11	Understandable	4	8	B
12	Ease of Remembering	8	6	C
13	Provides Guidance	12	(replace)	C
14	Easy to Use	NA	10	NA
NA	Ease of Learning	NA	2	NA
NA	Effort to Become Skillful	NA	9	NA

Example from (Davis, 1989)

6. Assess Reliability with New Data

Table 6. Factor Analysis of Perceived Usefulness and Ease of Use Questions: Study 1

Scale Items	Factor 1 (Usefulness)	Factor 1 (Ease of Use)
Usefulness		
1 Quality of Work	.80	.10
2 Control over Work	.86	−.03
3 Work More Quickly	.79	.17
4 Critical to My Job	.87	−.11
5 Increase Productivity	.87	.10
6 Job Performance	.93	−.07
7 Accomplish More Work	.91	−.02
8 Effectiveness	.96	−.03
9 Makes Job Easier	.80	.16
10 Useful	.74	.23
Ease of Use		
1 Cubersome	.00	.73
2 Ease of Learning	.08	.60
3 Frustrating	.02	.65
4 Controllable	.13	.74
5 Rigid & Inflexible	.09	.54
6 Ease of Remembering	.17	.62
7 Mental Effort	−.07	.76
8 Understandable	.29	.64
9 Effort to Be Skillful	−.25	.88
10 Easy to Use	.23	.72

Example from (Davis, 1989)

6. Assess Reliability with New Data (cont.)

Table 7. Factor Analysis of Perceived Usefulness and Ease of Use Items: Study 2

Scale Items		Factor 1 (Usefulness)	Factor 2 (Ease of Use)
Usefulness			
1	Work More Quickly	.91	.01
2	Job Performance	.98	−.03
3	Increase Productivity	.98	−.03
4	Effectiveness	.94	.04
5	Makes Job Easier	.95	−.01
6	Useful	.88	.11
Ease of Use			
1	Easy to Learn	−.20	.97
2	Controllable	.19	.83
3	Clear & Understandable	−.04	.89
4	Flexible	.13	.63
5	Easy to Become Skillful	.07	.91
6	Easy to Use	.09	.91

Example from (Davis, 1989)

7. Assess Construct Validity

- Multitrait-multimethod matrix
- Criterion validity

2. The validity coefficients (3) should be higher than the correlations in the heterotrait-monomethod triangles (2) which suggests that the correlation within a trait measured by different methods must be higher than the correlations between traits which have method in common.

1. Entries in the validity diagonal (3) should be higher than the correlations that occupy the same row and column in the heteromethod block (4). This is a minimum requirement.

MTMM

MULTITRAIT-MULTIMETHOD MATRIX
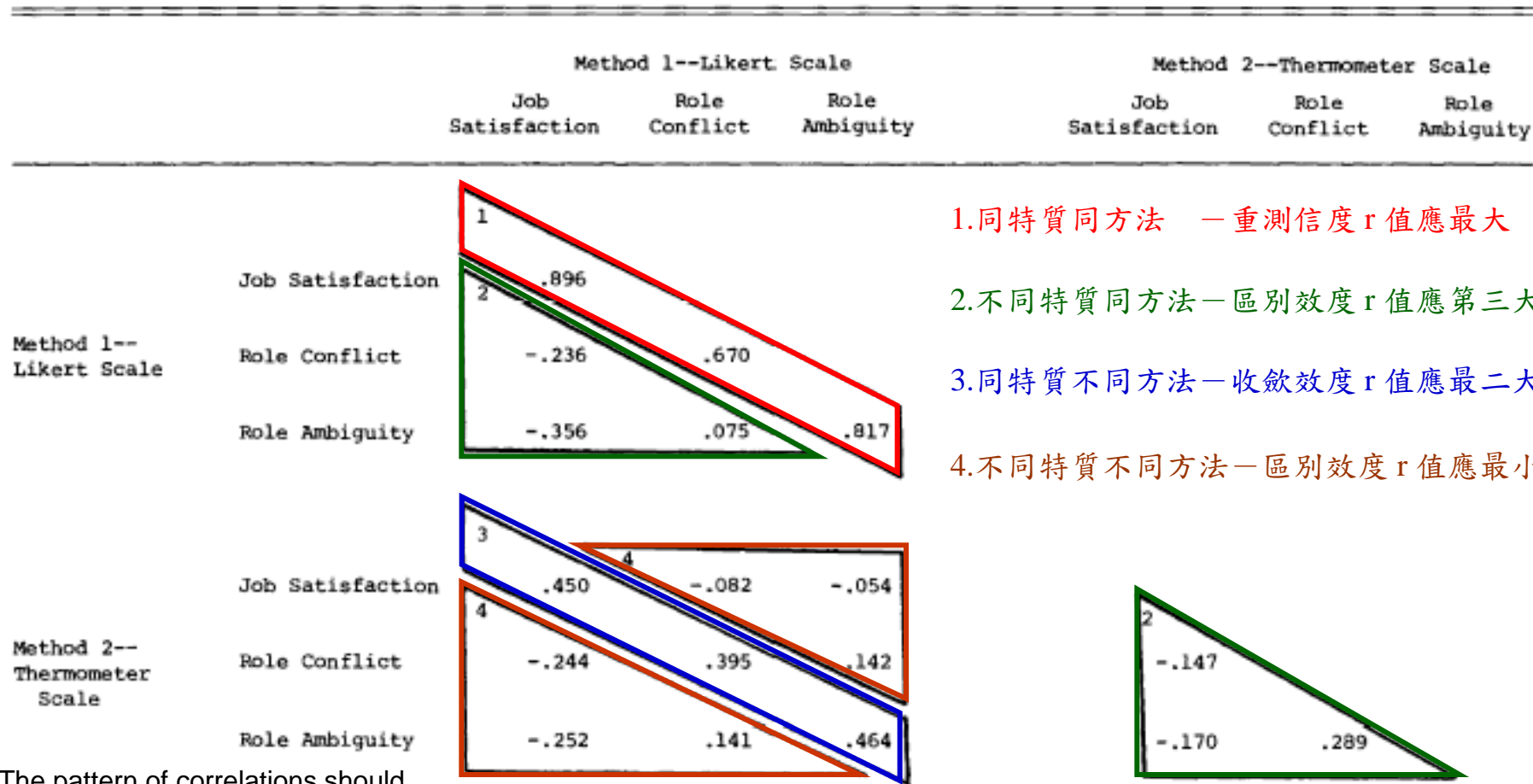
1.同特質同方法　－重測信度 r 值應最大

2.不同特質同方法－區別效度 r 值應第三大

3.同特質不同方法－收斂效度 r 值應最二大

4.不同特質不同方法－區別效度 r 值應最小

3. The pattern of correlations should be the same in all of the heterotrait triangles, e.g., both (2) and (4).

(Churchill, 1979)(A Paradigm for Developing Better Measures of marketing Constructs)

Does the Measure as Expected?
(Churchill, 1979)

- Four separate propositions (Nunnally, 1967, p. 93)
  - 1. The constructs job satisfaction (A) and likelihood of quitting (B) are related.
  - 2. The scale X provides a measure of A.
  - 3. Y provides a measure of B.
  - 4. X and Y correlate positively.
- Only the fourth proposition is directly examined with empirical data.
- To establish that X truly measures A, one must assume that propositions 1 and 3 are correct.
- One must have a good measure for B, and the theory relating A and B must be true.
- The analyst tries to establish the construct validity of a measure by relating it to a number of other constructs and not simply one.

7. Assess Construct Validity

Table 8. Correlations Between Perceived Usefulness, Perceived Ease of Use, and Self-Reported System Usage

	Correlation		
	Usefulness & Usage	**Ease of Use & Usage**	**Ease of Use & Usefulness**
Study 1			
Electronic Mail (n = 109)	.56***	.32***	.56***
XEDIT (n = 75)	.68***	.48***	.69***
Pooled (n = 184)	.63***	.45***	.64***
Study 2			
Chart-Master (n = 40)	.71***	.25	.25
Pendraw (n = 40)	.59***	.47***	.38**
Pooled (n = 80)	.85***	.59***	.56***
Davis, et al. (1989) (n = 107)			
Wave 1	.65***	.27**	.10
Wave 2	.70***	.12	.23**

*** $p < .001$ ** $p < .01$ * $p < .05$

7. Assess Construct Validity (cont.)

Table 9. Regression Analyses of the Effect of Perceived Usefulness and Perceived Ease of Use on Self-Reported Usage

	Independent Variables		
	Usefulness	Ease of Use	R^2
Study 1			
Electronic Mail (n = 109)	.55***	.01	.31
XEDIT (n = 75)	.69***	.02	.46
Pooled (n = 184)	.57***	.07	.38
Study 2			
Chart-Master (n = 40)	.69***	.08	.51
Pendraw (n = 40)	.76***	.17	.71
Pooled (n = 80)	.75***	.17*	.74
Davis, et al. (1989) (n = 107)			
After 1 Hour	.62***	.20***	.45
After 14 Weeks	.71***	− .06	.49

*** $p < .001$ ** $p < .01$ * $p < .05$

Final Measurement Scales for Perceived Usefulness and Perceived Ease of Use

Perceived Usefulness

Using CHART-MASTER in my job would enable me to accomplish tasks more quickly.

likely |_____|_____|_____|_____|_____|_____|_____| unlikely
       extremely quite slightly neither slightly quite extremely

Using CHART-MASTER would improve my job performance.

likely |_____|_____|_____|_____|_____|_____|_____| unlikely
       extremely quite slightly neither slightly quite extremely

Using CHART-MASTER in my job would increase my productivity.

likely |_____|_____|_____|_____|_____|_____|_____| unlikely
       extremely quite slightly neither slightly quite extremely

Using CHART-MASTER would enhance my effectiveness on the job.

likely |_____|_____|_____|_____|_____|_____|_____| unlikely
       extremely quite slightly neither slightly quite extremely

Using CHART-MASTER would make it easier to do my job.

likely |_____|_____|_____|_____|_____|_____|_____| unlikely
       extremely quite slightly neither slightly quite extremely

I would find CHART-MASTER useful in my job.

likely |_____|_____|_____|_____|_____|_____|_____| unlikely
       extremely quite slightly neither slightly quite extremely

Perceived Ease of Use

Learning to operate CHART-MASTER would be easy for me.

likely |_____|_____|_____|_____|_____|_____|_____| unlikely
       extremely quite slightly neither slightly quite extremely

I would find it easy to get CHART-MASTER to do what I want it to do.

likely |_____|_____|_____|_____|_____|_____|_____| unlikely
       extremely quite slightly neither slightly quite extremely

My interaction with CHART-MASTER would be clear and understandable.

likely |_____|_____|_____|_____|_____|_____|_____| unlikely
       extremely quite slightly neither slightly quite extremely

I would find CHART-MASTER to be flexible to interact with.

likely |_____|_____|_____|_____|_____|_____|_____| unlikely
       extremely quite slightly neither slightly quite extremely

It would be easy for me to become skillful at using CHART-MASTER.

likely |_____|_____|_____|_____|_____|_____|_____| unlikely
       extremely quite slightly neither slightly quite extremely

I would find CHART-MASTER easy to use.

likely |_____|_____|_____|_____|_____|_____|_____| unlikely
       extremely quite slightly neither slightly quite extremely

Example from (Davis, 1989)

8 Developing Norms

- A better way of assessing the position of the individual on the characteristic is to compare the person's score with the score achieved by other people.

- Norm quality is a function of both the number of cases on which the average is based and their representativeness.

Summary of Suggested Procedure for Developing Better Measures
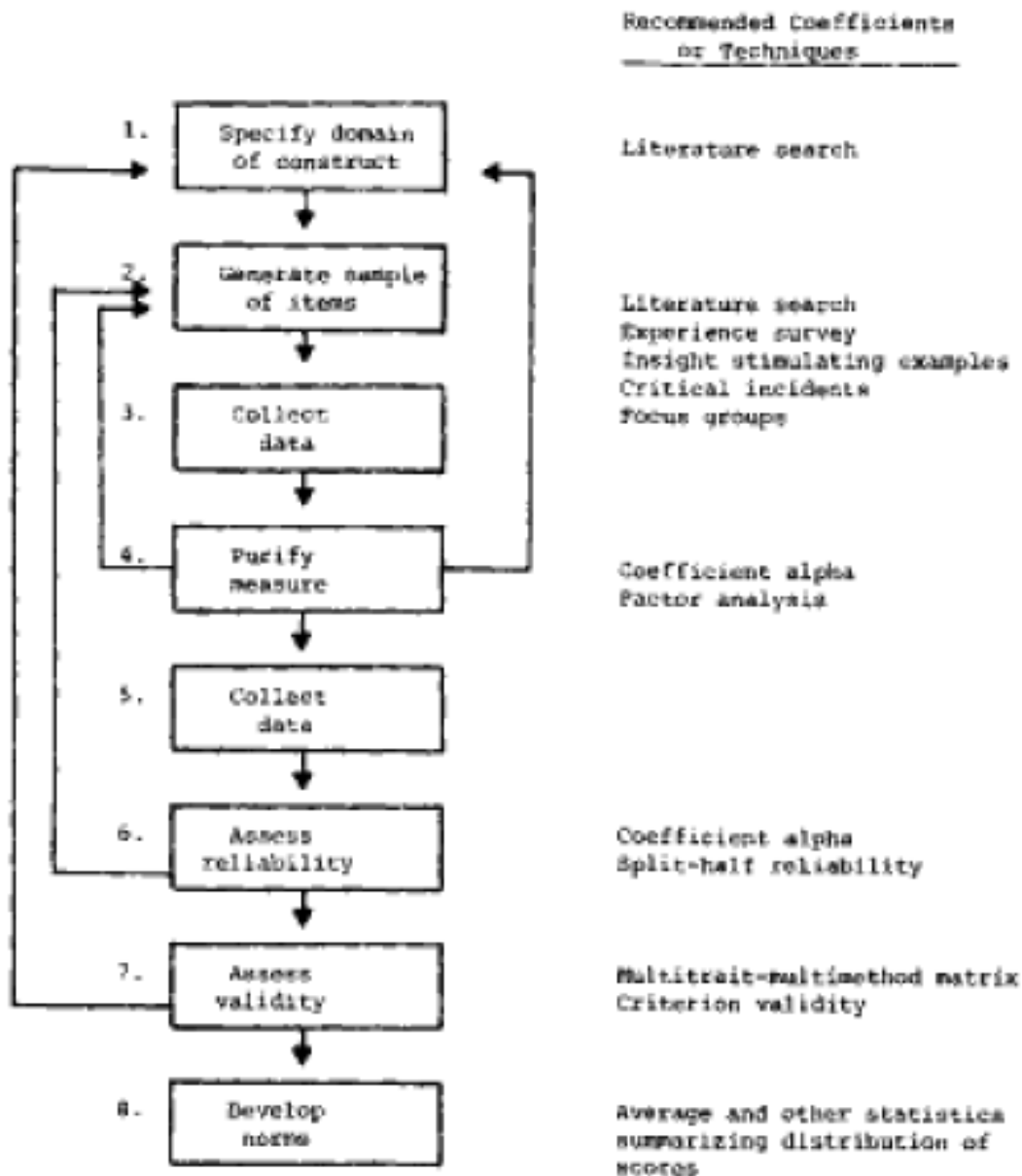**(Churchill, 1979)**

- Researchers doing applied work and practitioners could at least be expected to complete the process through step 4.

- Marketing researchers are already collecting data relevant to steps 5-8.
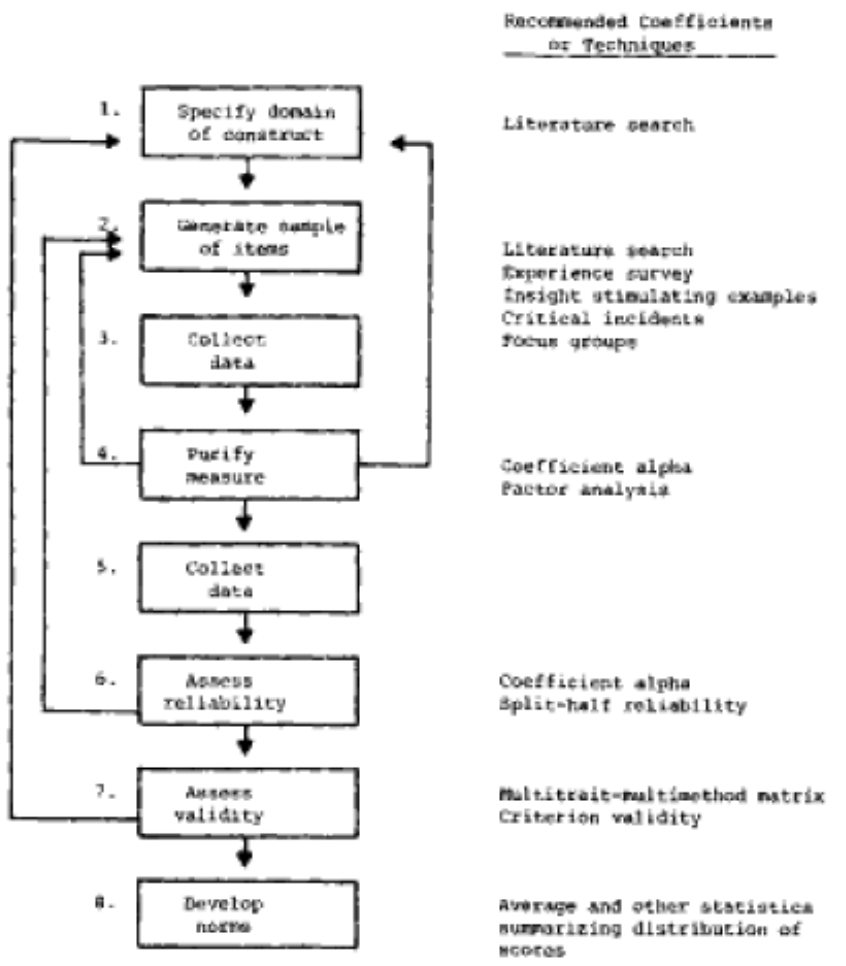
Current Practice in Scale Development

- Churchill, G. A., Jr., (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research, 16*(February), 64-73.
- Gerbing, D. W., & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. Journal of Marketing Research, 25(2), 186-192.
- DeVellis, R. F. (1991). *Scale development: Theory and applications.* Newbury Park, CA: Sage Publications.
- Spector, P. E. (1992). *Summated rating scale construction: An introduction.* Newbury Park, CA: Sage Publications.
- Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). *Scaling procedures: Issues and applications.* Thousand Oaks, CA: Sage Publications.
- Clark R. A. (2006), Consumer Independence: Conceptualization, Measurement and Validation of a Previously Unmeasured Social Response Tendency, Ph.D. Dissertation, College of Business of The Florida State University.
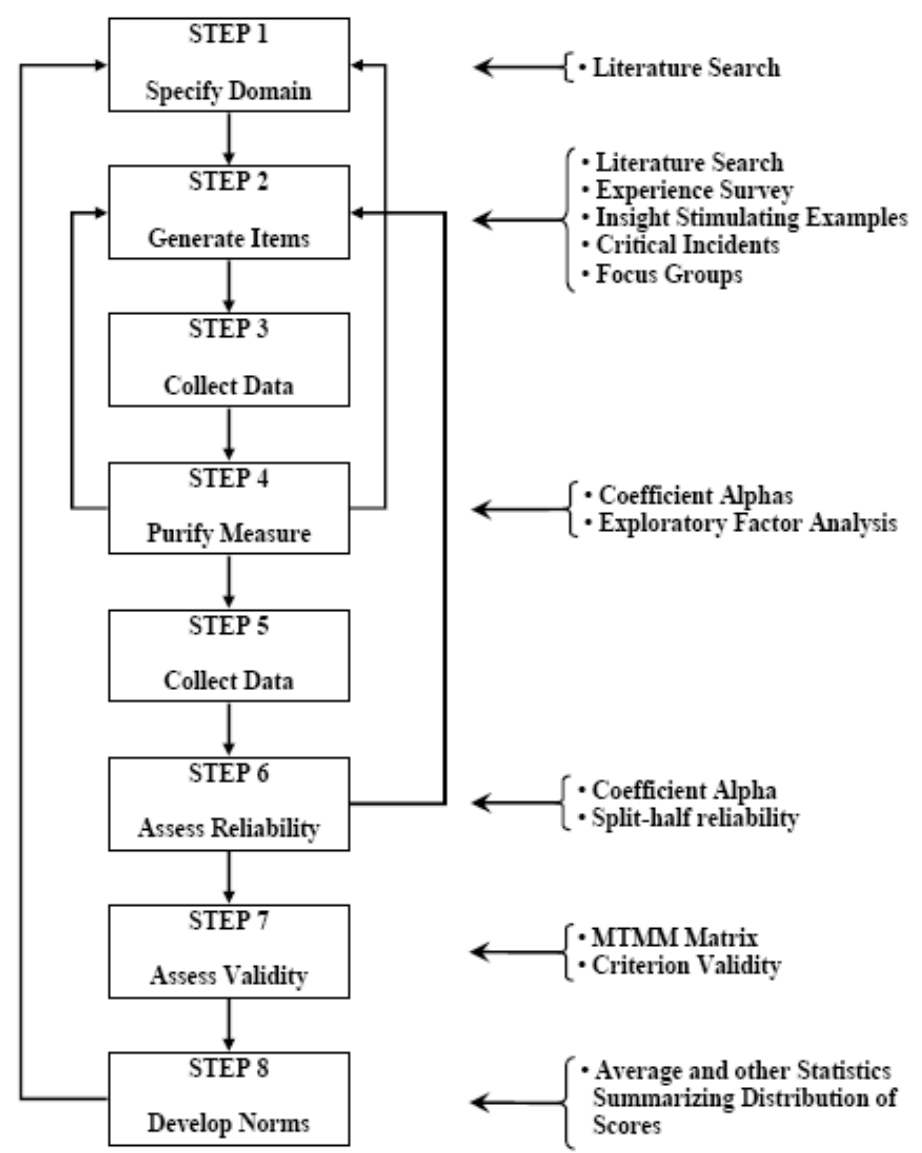
Suggested Procedure for Developing Better Measures (Churchill, 1979)

Recommended Coefficients
or Techniques

1.	Specify domain of construct	Literature search
2.	Generate sample of items	Literature search Experience survey Insight stimulating examples Critical incidents Focus groups
3.	Collect data	
4.	Purify measure	Coefficient alpha Factor analysis
5.	Collect data	
6.	Assess reliability	Coefficient alpha Split-half reliability
7.	Assess validity	Multitrait-multimethod matrix Criterion validity
8.	Develop norms	Average and other statistics summarizing distribution of scores

(Churchill, 1979)(A Paradigm for Developing Better Measures of marketing Constructs)

Current Practice in Scale Development

Recommended Coefficients or Techniques

1. Specify domain of construct — Literature search
2. Generate sample of items — Literature search, Experience survey, Insight stimulating examples, Critical incidents, focus groups
3. Collect data
4. Purify measure — Coefficient alpha, Factor analysis
5. Collect data
6. Assess reliability — Coefficient alpha, Split-half reliability
7. Assess validity — Multitrait-multimethod matrix, Criterion validity
8. Develop norms — Average and other statistics summarizing distribution of scores

(Churchill, 1979)

STEP 1 Specify Domain — • Literature Search
STEP 2 Generate Items — • Literature Search, • Experience Survey, • Insight Stimulating Examples, • Critical Incidents, • Focus Groups
STEP 3 Collect Data
STEP 4 Purify Measure — • Coefficient Alphas, • Exploratory Factor Analysis
STEP 5 Collect Data
STEP 6 Assess Reliability — • Coefficient Alpha, • Split-half reliability
STEP 7 Assess Validity — • MTMM Matrix, • Criterion Validity
STEP 8 Develop Norms — • Average and other Statistics Summarizing Distribution of Scores
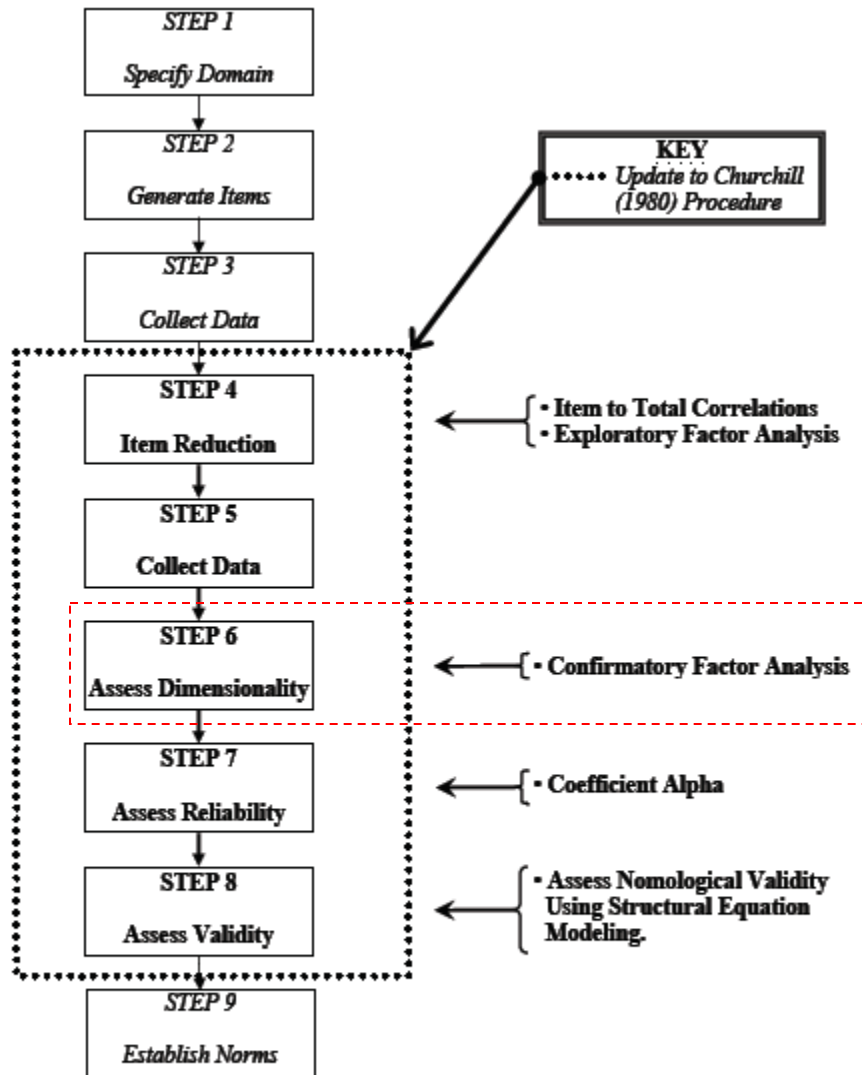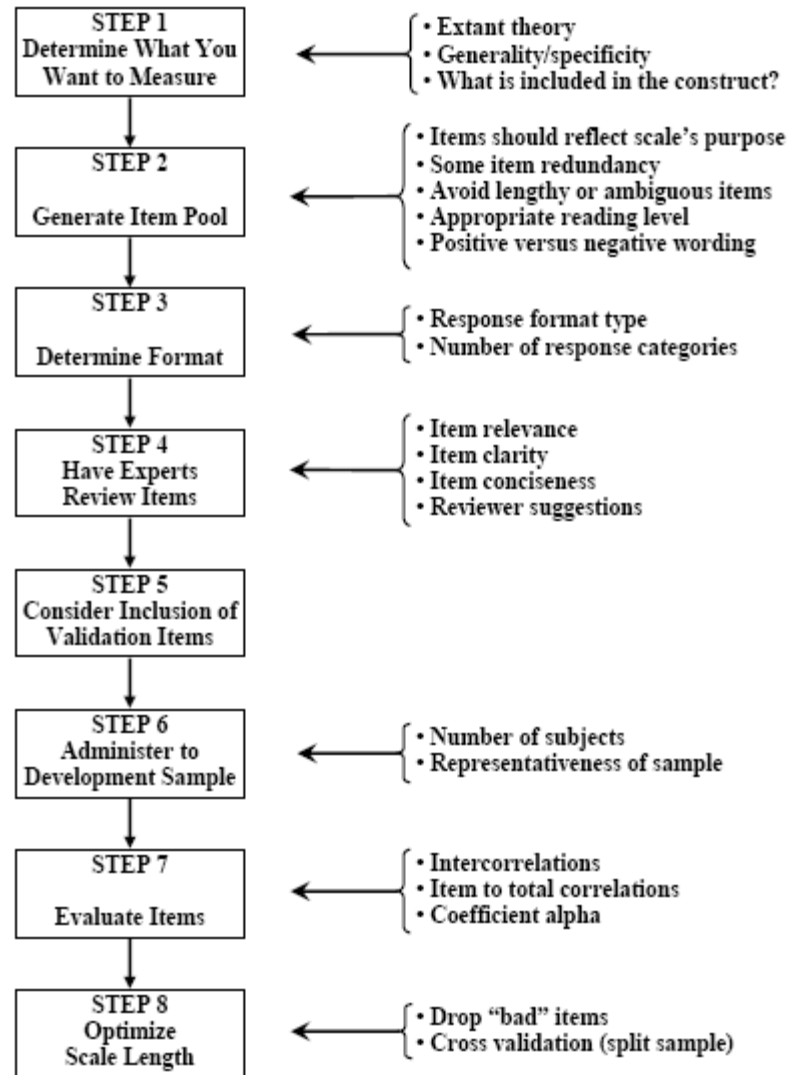
Figure 3.1
Churchill's (1980) Scale Development Procedure
Adapted from Churchill (1979)

(Gerbing & Anderson, 1988)

Figure 3.2
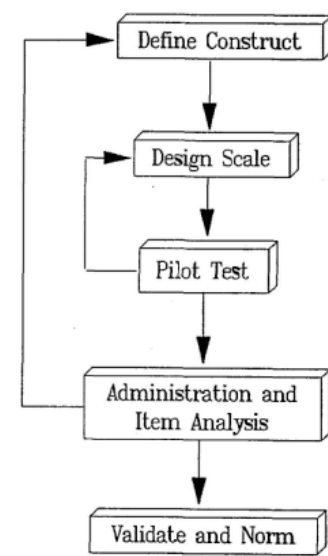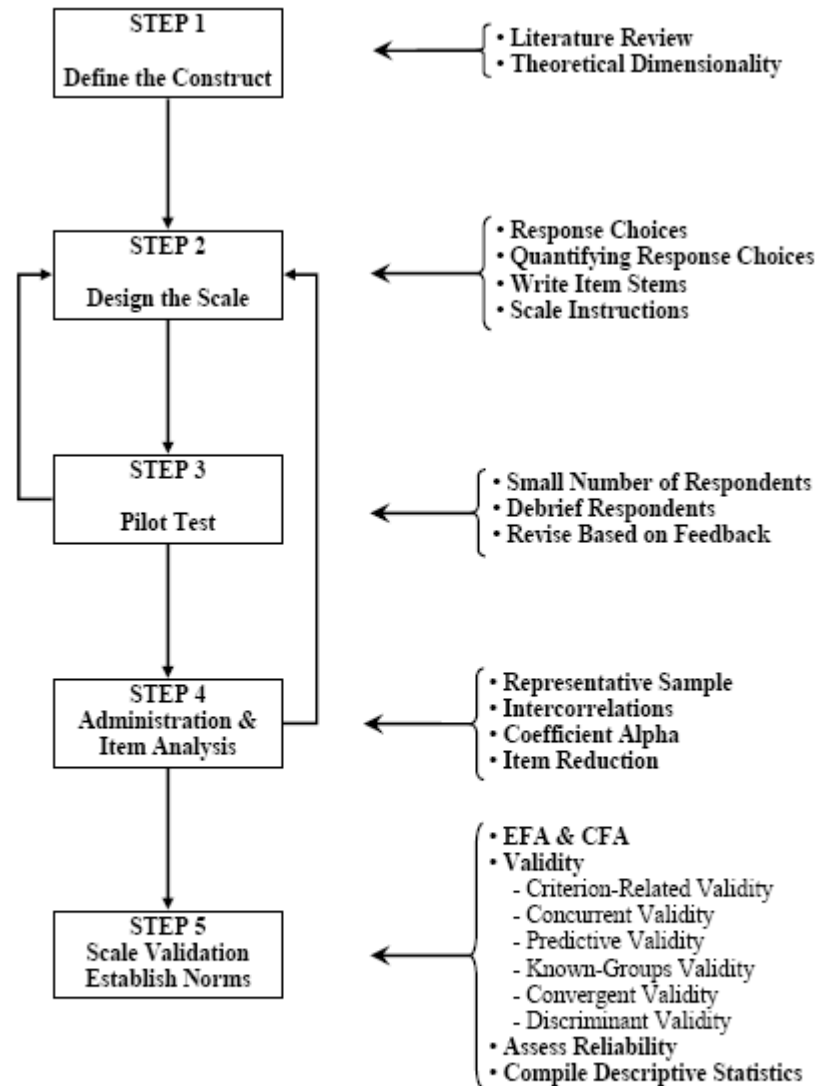**Gerbing & Anderson's (1988) Updated Paradigm**

(DeVellis, 1991)

STEP 1
Determine What You Want to Measure
- Extant theory
- Generality/specificity
- What is included in the construct?

STEP 2
Generate Item Pool
- Items should reflect scale's purpose
- Some item redundancy
- Avoid lengthy or ambiguous items
- Appropriate reading level
- Positive versus negative wording

STEP 3
Determine Format
- Response format type
- Number of response categories

STEP 4
Have Experts Review Items
- Item relevance
- Item clarity
- Item conciseness
- Reviewer suggestions

STEP 5
Consider Inclusion of Validation Items

STEP 6
Administer to Development Sample
- Number of subjects
- Representativeness of sample

STEP 7
Evaluate Items
- Intercorrelations
- Item to total correlations
- Coefficient alpha

STEP 8
Optimize Scale Length
- Drop "bad" items
- Cross validation (split sample)

Figure 3.3
DeVellis's (1991) Scale Development Approach

(Spector, 1992)

| STEP 1
Define the Construct | ← | { • Literature Review
• Theoretical Dimensionality |

| STEP 2
Design the Scale | ← | { • Response Choices
• Quantifying Response Choices
• Write Item Stems
• Scale Instructions |

| STEP 3
Pilot Test | ← | { • Small Number of Respondents
• Debrief Respondents
• Revise Based on Feedback |

| STEP 4
Administration &
Item Analysis | ← | { • Representative Sample
• Intercorrelations
• Coefficient Alpha
• Item Reduction |

| STEP 5
Scale Validation
Establish Norms | ← | { • EFA & CFA
• Validity
   - Criterion-Related Validity
   - Concurrent Validity
   - Predictive Validity
   - Known-Groups Validity
   - Convergent Validity
   - Discriminant Validity
• Assess Reliability
• Compile Descriptive Statistics |

Figure 3.4
Spector's (1992) Summated Rating Scale Development Procedure
Adapted from Spector (1992)

Major Steps to Developing
a Summated Rating Scale
(Spector, 1992, p.8)

(Netemeyer et al., 2003)

STEP 1
Construct Definition
- Delimit construct domain
- Theoretical Definition
- Theoretical Dimensionality

STEP 2
Generating and Judging Items
- Domain Sampling
- Generate an Item Pool
  - Item Sources
  - Item Writing
  - Number of Items
- Judge Items
  - Content Validity
  - Face Validity

STEP 3
Designing & Conducting Studies to Develop a Scale
- Pilot Test
- Initial Validation Studies
  - Include Constructs
  - EFA
  - Coefficient Alpha
  - Item to Total Correlations
  - Item Reduction

STEP 4
Administration & Item Analysis
- EFA & Item Statistics
  - Item Reduction
- CFA
  - Fit Indices
  - Significant Parameter Estimates
  - Composite (construct) Reliability
  - Average Variance Extracted
  - Residuals & Modification Indices
  - Measurement Invariance
- Validity
  - Convergent & Discriminant
  - Predictive & Concurrent
- Norms
  - Means & Standard Deviations

Figure 3.5
Netemeyer, Bearden, & Sharma's (2003) Scaling Procedure

(Rossiter, 2002)

| STEP 1 Construct Definition | • Write Initial Definition
 - Object
 - Attribute
 - Rater Entity |

| STEP 2 Object Classification | • Open-ended Interviews
• Classify Object
 - Concrete Singular
 - Abstract Collective
 - Abstract Formed
• Generate Object Item Parts |

| STEP 3 Attribute Classification | • Open-ended Interviews
• Classify Attribute
 - Concrete
 - Formed
 - Eliciting
• Generate Attribute Item Parts |

| STEP 4 Rater Identification | • Identify Rater
 - Individual
 - Group
 - Expert
• Estimate Reliability |

| STEP 5 Scale Formation | • Combine Object & Attribute Parts
• Select Appropriate Rating Scales
• Pre-Test Each Item
 - Comprehension
 - Dimensionality (if eliciting)
• Randomize Order |

| STEP 6 Enumeration | • Derive Total Scale Score
 - Using Indexes & Averages
• Transform into Meaningful Range
• Report an Estimate of Precision
 - Reliability |

Figure 3.6
Rossiter's (2002) C-OAR-SE Marketing Scale Procedure
Adapted from Rossiter (2002)

C-OAR-SE procedure

- Rossiter (2002) laments that the current scale paradigm places too much emphasis on empiricism (i.e., factor analysis and reliability), which leads deletion of conceptually necessary items and retention of conceptually inappropriate items.

- The emphasis in the C-OAR-SE procedure is on content validity (Rossiter, 2002).

(Clark, 2006)
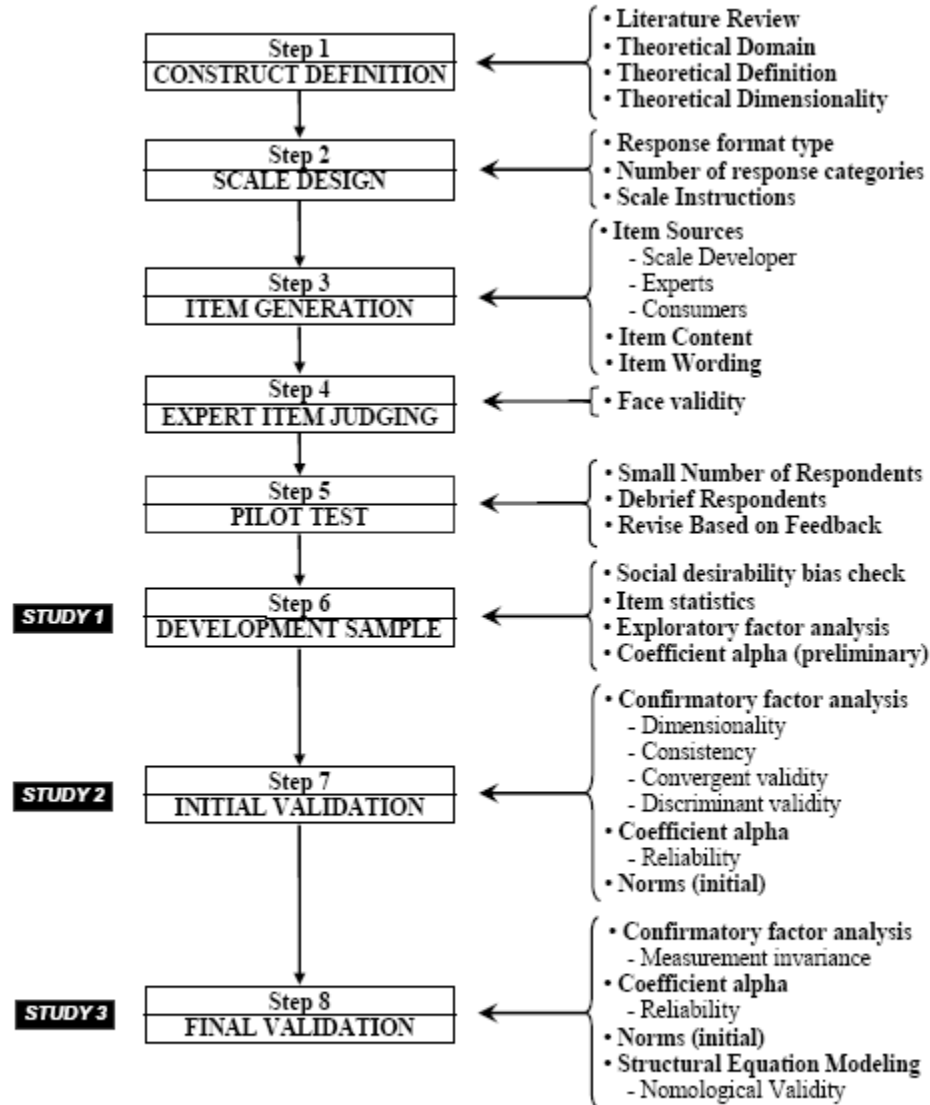
Figure 3.7
Amalgamated Scale Development Procedure

研究流程	研究方法與工具	研究內容

研究流程

```
  ┌─────────────────┐
  │  1. 構念定義      │◄──┐
  └─────────────────┘   │
          │             │
  ┌─────────────────┐   │
  │  2. 問項發展      │   │
  └─────────────────┘   │
          │             │
  ┌─────────────────┐   │
  │  3. 資料蒐集      │   │
  └─────────────────┘   │
          │             │
  ┌─────────────────┐   │
  │  4. 量表精鍊      │   │
  └─────────────────┘   │
          │             │
  ┌─────────────────┐   │
  │  5. 資料再蒐集    │   │
  └─────────────────┘   │
          │             │
  ┌─────────────────┐   │
  │  6. 量表再精鍊    │   │
  └─────────────────┘   │
          │             │
  ┌─────────────────┐   │
  │  7. 效度評估      │   │
  └─────────────────┘   │
          │             │
  ┌─────────────────┐   │
  │  8. 發展常模      │   │
  └─────────────────┘
```
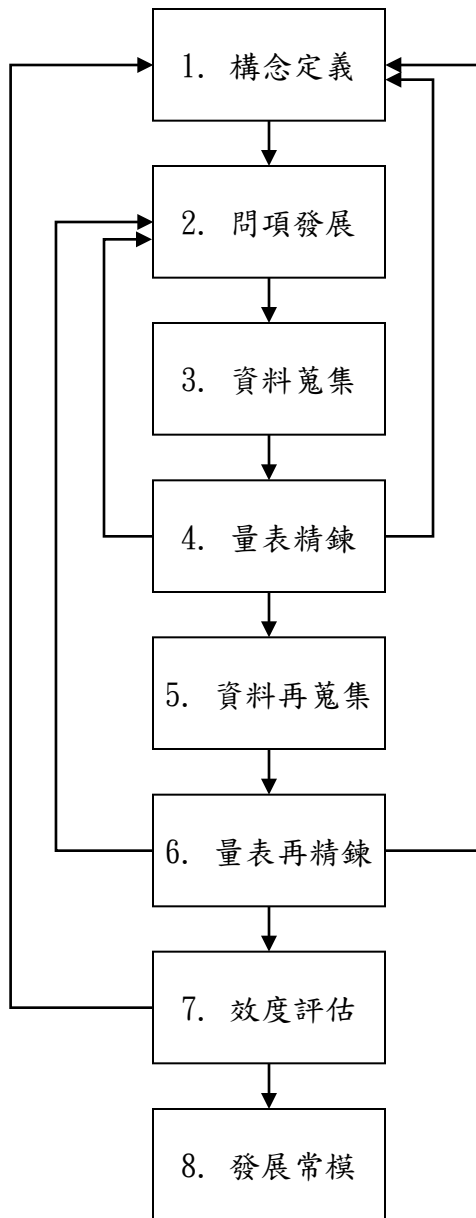
研究方法與工具

‧文獻探討

‧文獻蒐尋
‧經驗調查
‧內容效度比率(CVR)

‧抽樣

‧Cronbach’s α係數
‧相關係數矩陣
‧Item-to-Total相關法

‧抽樣

‧因素分析
‧Cronbach’s α係數
‧Item-to-Total相關法

‧相關係數矩陣
‧多特質多方法矩陣(MTMM)
‧Pearson積差相關係數

‧中位數
‧百分位數
‧標準差
‧平均數
‧期望常態分配

研究內容

•領域界定
•歸納構念之關係面向
•構念之定義

•發展問項集合(初始問項)
•決定量表格式
•確保內容效度
•加入效度評估問項

•決定抽樣方法
•決定樣本規模
•針對小樣本進行預試

•信度與構念效度分析
•刪除不良問項確保構念效度

•決定抽樣方法
•決定樣本規模
•針對大樣本進行預試

•信度與構念效度分析
•刪除不良問項確保構念效度

•驗證內容效度
•驗證構念效度
•驗證法理效度

•發展測量評估標準
•樣本分數之統計分配

研究流程 | 研究方法與工具 | 研究內容

研究流程

1. 構念定義
2. 問項發展
3. 資料蒐集
4. 量表精鍊
5. 資料再蒐集
6. 信度評估
7. 效度評估
8. 發展常模
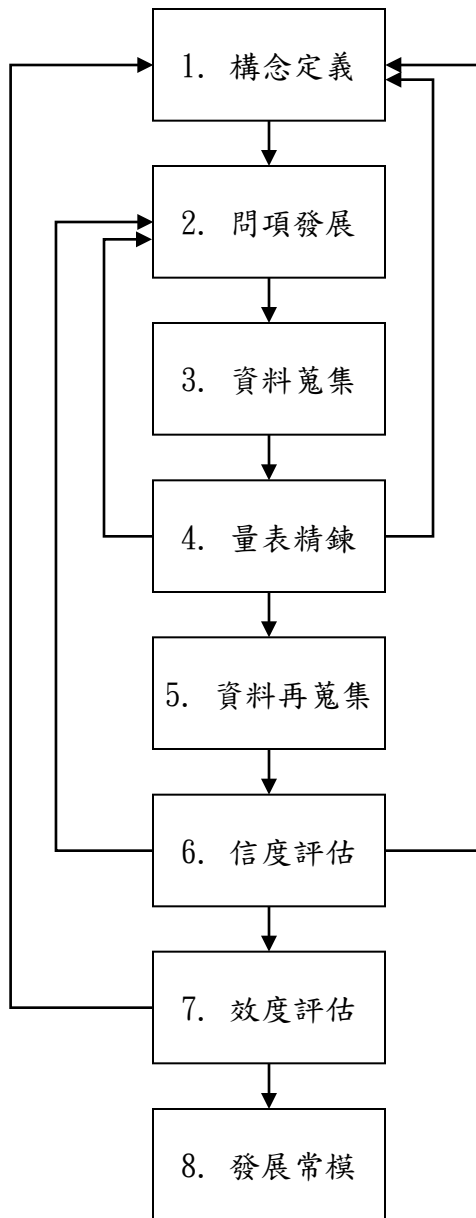
研究方法與工具

•文獻探討

•文獻蒐尋
•經驗調查
•專家意見
•焦點群體
•內容效度比率(CVR)
•表面效度

•抽樣

•項目分析(Item Analysis)
•探索性因素分析 (EFA)
•Cronbach's α係數
•相關係數矩陣
•Item-to-Total相關法

•抽樣

•探索性因素分析 (EFA)
•Cronbach's α係數
•Item-to-Total相關法
•驗證性因素分析(CFA)

•相關係數矩陣
•多特質多方法矩陣(MTMM)
•Pearson積差相關係數
•驗證性因素分析(CFA)(SEM)

•中位數
•百分位數
•標準差
•平均數
•期望常態分配

研究內容

•領域界定
•歸納構念之關係面向
•構念之定義

•發展問項集合(初始問項)
•決定量表格式
•確保內容效度
•加入效度評估問項

•決定抽樣方法
•決定樣本規模
•針對小樣本進行預試

•信度與構念效度分析
•刪除不良問項確保構念效度

•決定抽樣方法
•決定樣本規模
•針對大樣本進行預試

•信度與構念效度分析
•刪除不良問項確保構念效度

•驗證內容效度
•驗證構念效度
•驗證法理效度

•發展測量評估標準
•樣本分數之統計分配

(adapted from 賴榮裕，2006; Netemeyer et al., 2003; Spector, 1992; DeVellis, 1991; Gerbin & Anderson, 1988; Churchill Jr., 1979)

Summary of Best practices for scale development

- Follow the paradigm for developing better measures (Churchll, 1978; Gerbing, D. W., & Anderson) and best practices for scale development (Netemeyer et al., 2003; Spector, 1992; DeVellis, 1991).

The linkage among attitudes, behavior, and marketing effectiveness

Attitudes and Linkage

- Attitude defined:
  - Enduring organization of motivational, emotional, perceptual, and cognitive processes with respect to some aspect of a person's environment.

  - Level of Customer Involvement
  - Attitude Measurement & Strength
  - Effects of Other People & Brands
  - Situational Factors

Measurement Scales

- Scaling defined:
  - Procedures for assigning numbers (or other symbols) to properties of an object in order to impart some numerical characteristics to the properties in question.

Measurement Scales

- Scaling Approaches:
  - Unidimensional:
    - Measures only one attribute of a concept, respondent, or object.

  - Multidimensional:
    - Measures several dimensions of a concept, respondent, or object.

Measurement Scales

- Types of Scales:
  - Noncomparative Scale:
    - Scales in which judgment is made without reference to another object, concept, or person.

  - Comparative Scale:
    - Scales in which one object, concept, or person is compared with another on a scale.

Graphic Rating Scales

- Measurement scales that include a graphic continuum, anchored by two extremes.
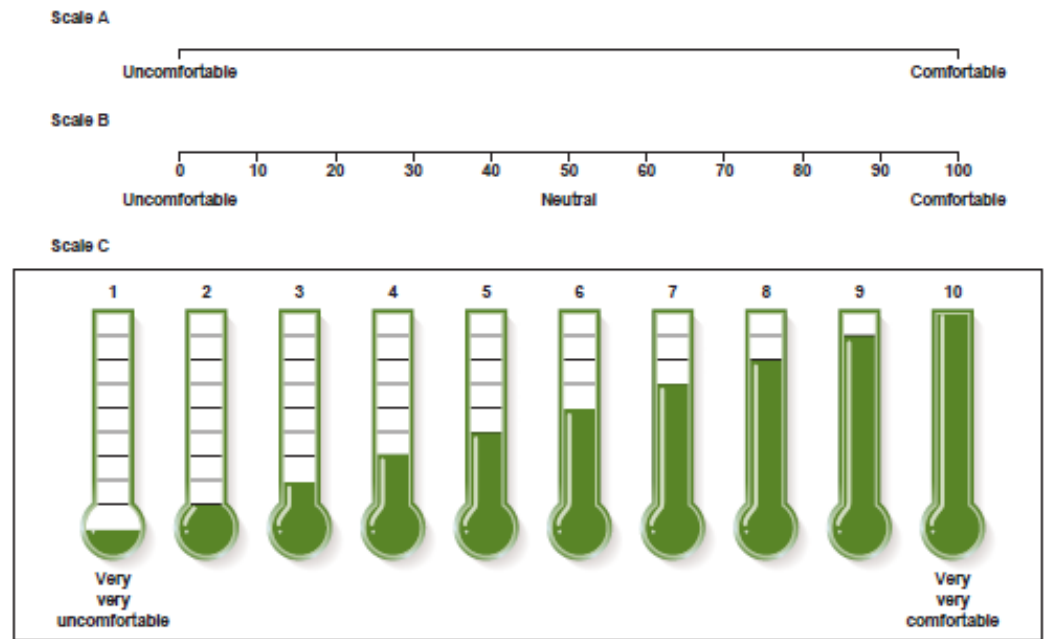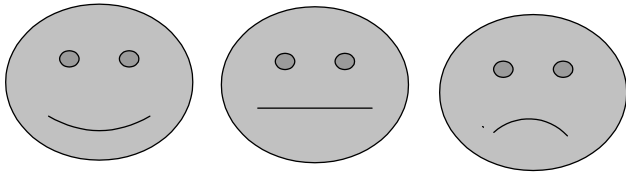
Graphic Rating Scales
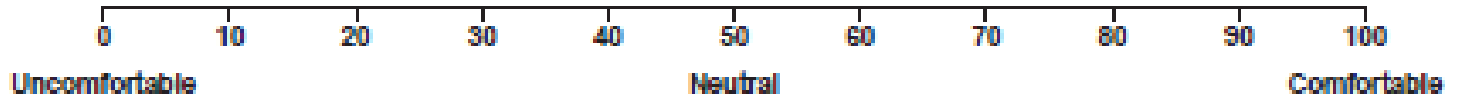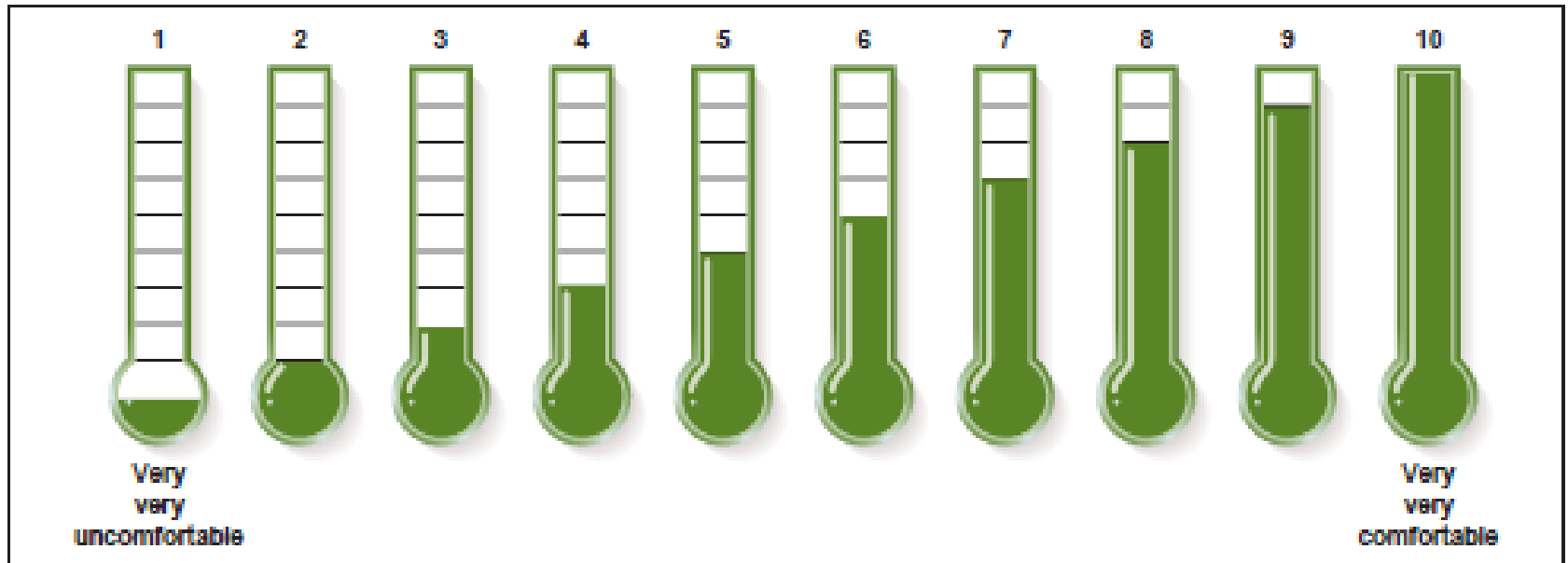
Itemized Rating Scales

- The respondent selects an answer from a limited number of ordered categories.

Odd Scale

Important   **Not Important**

1   **2**   **3**   **4**   **5**

Even Scale

Important   **Not Important**

1   **2**   **3**   **4**   **5**   **6**

Itemized Rating Scales

Exhibit 11.2

Itemized Rating Scales Used in Internet and Mall Surveys

If offered, how likely would you be to use the following areas on this site?

Scale A

a. Auctions
Not at all likely to use   ○1   ○2   ○3   ○4   ○5   ○6   ○7   Extremely likely to use

b. Fee-based education tools
Not at all likely to use   ○1   ○2   ○3   ○4   ○5   ○6   ○7   Extremely likely to use

c. Event registration
Not at all likely to use   ○1   ○2   ○3   ○4   ○5   ○6   ○7   Extremely likely to use

d. Online shopping markets
Not at all likely to use   ○1   ○2   ○3   ○4   ○5   ○6   ○7   Extremely likely to use

e. Recruiting
Not at all likely to use   ○1   ○2   ○3   ○4   ○5   ○6   ○7   Extremely likely to use

f. Research subscription
Not at all likely to use   ○1   ○2   ○3   ○4   ○5   ○6   ○7   Extremely likely to use

g. Trading community
Not at all likely to use   ○1   ○2   ○3   ○4   ○5   ○6   ○7   Extremely likely to use

h. Training/seminars
Not at all likely to use   ○1   ○2   ○3   ○4   ○5   ○6   ○7   Extremely likely to use

Itemized Rating Scales

What factors influence your choice of music Web sites? (Rate the importance of each item.)

	Not at All Important				Very Important
Customer benefits or rewards for shopping	○	○	○	○	○
Customer service or delivery options	○	○	○	○	○
Ease of use of Web site	○	○	○	○	○
Low prices	○	○	○	○	○
Real-time audio sampling of CDs	○	○	○	○	○
Reviews and artist information	○	○	○	○	○

Scale D

How interested would you be in obtaining additional information about this customer relationship management solution for your business?

○ Extremely interested   ○ Somewhat interested   ○ Not at all interested

○ Very interested   ○ Not very interested

How likely is it that your business will invest in this type of customer relationship management solution within the next 12 months?

○ Extremely likely   ○ Somewhat likely   ○ Not at all likely

○ Very likely   ○ Not very likely

Itemized Rating Scales

Submitting a Request for a Hotel Reservation

We'd like to get your feedback regarding your experience in submitting a request for a hotel reservation at our Web site today. Please rate your satisfaction with each of the following aspects of fasthotels.com based on your experience this visit.

	Very Satisfied				Very Dissatisfied
	1	2	3	4	5
Ability to access the offer page	○	○	○	○	○
Ability to locate hotel information	○	○	○	○	○
Ability to locate city information	○	○	○	○	○
Clarity of how the bonus program works	○	○	○	○	○
Clarity of the purchase agreement	○	○	○	○	○

Please rate the extent to which you are satisfied that Fasthotels.com has communicated each of the following to you during this visit:

	Very Satisfied				Very Dissatisfied
	1	2	3	4	5
Your hotel reservation is/will be nonchangeable	○	○	○	○	○
Your hotel reservation is/will be nonrefundable	○	○	○	○	○

How satisfied would you say you were with this visit to Fasthotels.com?

○ Very satisfied
○ Satisfied
○ Somewhat satisfied
○ Neither satisfied nor dissatisfied
○ Somewhat dissatisfied
○ Dissatisfied
○ Very dissatisfied

Itemized Rating Scales

EXHIBIT 11.3	Selected Itemized Rating Scales				
Characteristic of Interest	**Rating Choices**				
Purchase Intent	Definitely will buy	Probably will buy	Probably will not buy	Definitely will not buy	
Level of Agreement	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
Quality	Very good	Good	Neither good nor bad	Fair	Poor
Dependability	Completely dependable	Somewhat dependable	Not very dependable	Not dependable at all	
Style	Very stylish	Somewhat stylish	Not very stylish	Completely unstylish	
Satisfaction	Completely satisfied	Somewhat satisfied	Neither satisfied nor dissatisfied	Somewhat dissatisfied	Completely dissatisfied
Cost	Extremely expensive	Expensive	Neither expensive nor inexpensive	Slightly inexpensive	Very inexpensive
Ease of Use	Very easy to use	Somewhat easy to use	Not very easy to use	Difficult to use	
Color Brightness	Extremely bright	Very bright	Somewhat bright	Slightly bright	Not bright at all
Modernity	Very modern	Somewhat modern	Neither modern nor old-fashioned	Somewhat old-fashioned	Very old-fashioned

One Stage vs. Two Stage

Traditional One-Stage Format

"How effective do you believe Senator Foghorn is in having your money stay in the community?"

Very effective	Somewhat effective	Somewhat ineffective	Very ineffective	Don't know
4	3	2	1	0

Two-Stage Format

"How effective do you believe Senator Foghorn is in having your money stay in the community?"

How effective?	Would that be very or somewhat?
☐ Effective	☐ Very
☐ Ineffective	☐ Somewhat
☐ No opinion	

Advice for analyzing rating scales is given in the Practicing Marketing Research box below.

Rank Order Scale

Uses Comparative Scaling:

Put these fast food chains in order of preference:

- McDonalds
- Burger King
- Taco Bell

Rank Order Scale

Series of Rank-Order Scales Used to Evaluate Eye Shadows and Car Resale Values

Eye Shadow Scales

Please rank the following eye shadows, with 1 being the brand that best meets the characteristic being evaluated and 6 the worst brand on the characteristic being evaluated. The six brands are listed on card C. (HAND RESPONDENT CARD C.) Let's begin with the idea of having high-quality compacts or containers. Which brand would rank as having the highest quality compacts or containers? Which is second? (RECORD BELOW.)

	Q.48. Having High-Quality Container	Q.49. Having High-Quality Applicator	Q.50. Having High-Quality Eye Shadow
Avon	_____	_____	_____
Cover Girl	_____	_____	_____
Estee Lauder	_____	_____	_____
L'Oreal	_____	_____	_____
Natural Wonder	_____	_____	_____
Revlon	_____	_____	_____

Q-Sorting

- Q-sorting is basically a sophisticated form of rank ordering.

- A respondent is given cards listing a set of objects—such as verbal statements, slogans, product features, or potential customer services—and asked to sort them into piles according to specified rating categories.

- Q-sorts usually contain a large number of cards—from 60 to 120 cards.

- For statistical convenience, the respondent is instructed to put varying numbers of cards in several piles, the whole making up a normal statistical distribution.

Here is a Q-sort distribution of 90 items:

Excellent Feature									Poor Feature	
3	4	7	10	13	16	13	10	7	4	3
10	9	8	7	6	5	4	3	2	1	0

This is a rank-order continuum from Excellent Feature (10) to Poor Feature (0), with varying degrees of approval and disapproval between the extremes.

Q-Sorting

Here is a Q-sort distribution of 90 items:

Excellent Feature										Poor Feature
3	4	7	10	13	16	13	10	7	4	3
10	9	8	7	6	5	4	3	2	1	0

This is a rank-order continuum from Excellent Feature (10) to Poor Feature (0), with varying degrees of approval and disapproval between the extremes.

Paired Comparison

"Which drink do you prefer:"

___Coke
___Pepsi
_____

___Coke
___Sprite
_____

___Pepsi
___Sprite

Exhibit 11.5

Paired Comparison Scale for Sun Care Products

Here are some characteristics used to describe sun care products in general. Please tell me which characteristic in each pair is more important to you when selecting a sun care product.

a. Tans evenly	b. Tans without burning
a. Prevents burning	b. Protects against burning and tanning
a. Good value for the money	b. Goes on evenly
a. Not greasy	b. Does not stain clothing
a. Tans without burning	b. Prevents burning
a. Protects against burning and tanning	b. Good value for the money
a. Goes on evenly	b. Tans evenly
a. Prevents burning	b. Not greasy

Paired Comparison

Exhibit 11.5

Paired Comparison Scale for Sun Care Products

Here are some characteristics used to describe sun care products in general. Please tell me which characteristic in each pair is more important to you when selecting a sun care product.

a. Tans evenly	b. Tans without burning
a. Prevents burning	b. Protects against burning and tanning
a. Good value for the money	b. Goes on evenly
a. Not greasy	b. Does not stain clothing
a. Tans without burning	b. Prevents burning
a. Protects against burning and tanning	b. Good value for the money
a. Goes on evenly	b. Tans evenly
a. Prevents burning	b. Not greasy

Constant Sum Scale

What features do you want in a car?

Sun roof   _____
Leather   _____
ABS Breaks _____
CD Player   _____

Total   100 points

Constant Sum Scale

Exhibit 11.6
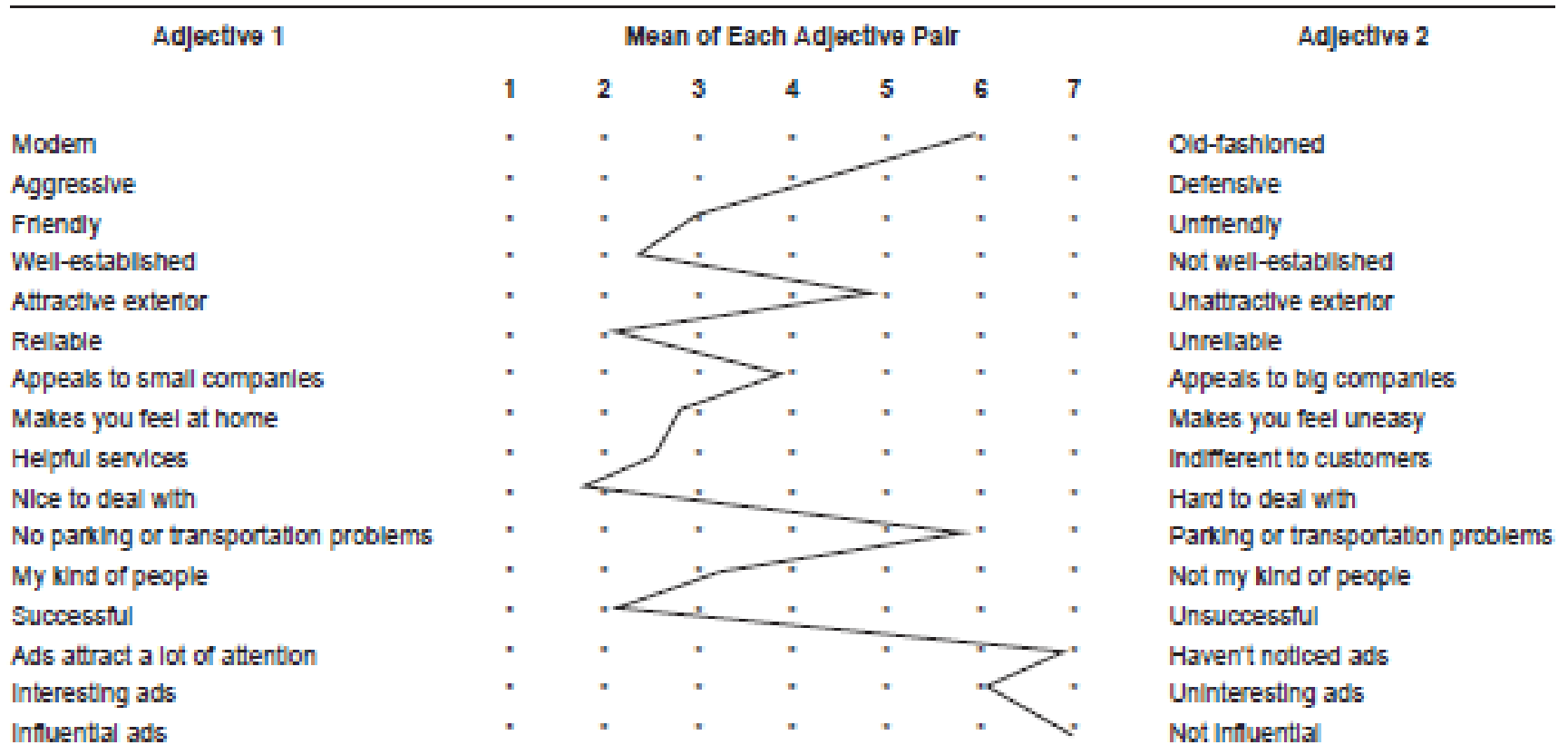
Constant Sum Scale Used in Tennis Sportswear Study

Below are seven characteristics of women's tennis sportswear. Please allocate 100 points among the characteristics such that the allocation represents the importance of each characteristic to you. The more points that you assign to a characteristic, the more important it is. If the characteristic is totally unimportant, you should not allocate any points to it. When you've finished, please double-check to make sure that your total adds to 100.

Characteristics of Tennis Sportswear	Number of Points
Is comfortable to wear	_____
Is durable	_____
Is made by well-known brand or sports manufacturers	_____
Is made in the United States	_____
Has up-to-date styling	_____
Gives freedom of movement	_____
Is a good value for the money	_____
	100 points

Semantic Differential Scale

Exhibit 11.7

Semantic Differential Profile of an Arizona Savings and Loan Association

Adjective 1	1	2	3	4	5	6	7	Adjective 2
Modern								Old-fashioned
Aggressive								Defensive
Friendly								Unfriendly
Well-established								Not well-established
Attractive exterior								Unattractive exterior
Reliable								Unreliable
Appeals to small companies								Appeals to big companies
Makes you feel at home								Makes you feel uneasy
Helpful services								Indifferent to customers
Nice to deal with								Hard to deal with
No parking or transportation problems								Parking or transportation problems
My kind of people								Not my kind of people
Successful								Unsuccessful
Ads attract a lot of attention								Haven't noticed ads
Interesting ads								Uninteresting ads
Influential ads								Not influential

Staple Scale

Exhibit 11.8

Stapel Scale Used to Measure a Retailer's Web Site

+5	+5
+4	+4
+3	+3
+2	+2
+1	+1
Cheap Prices	Easy to Navigate
−1	−1
−2	−2
−3	−3
−4	−4
−5	−5

Likert Scale

Likert Scales Used by an Internet Game Site

Scale A

How did you feel about the registration process when you became a new user?

	Strongly disagree	Somewhat disagree	Neutral	Somewhat agree	Strongly agree
The registration was simple.	○	○	○	○	○
The registration questions were "nonthreatening."	○	○	○	○	○
Registration here will protect my privacy.	○	○	○	○	○
The registration did not take a long time to complete.	○	○	○	○	○
The registration informed me about the site.	○	○	○	○	○

Purchase Intent Scales

Exhibit 11.10

Purchase Intent Scale and Related Questions for In-Home Product Placement of Fly Traps

21. If a set of three traps sold for approximately $3.00 and was available in the stores where you normally shop, would you:

	(51)
definitely buy the set of traps	1
probably buy	2
probably not buy	3
definitely not buy	4

22. Would you use the traps (a) instead of or (b) in addition to existing products?

	(52)
instead of	1
in addition to	2

23. Would you recommend this product to your friends?

	(53)
definitely	1
probably	2
probably not	3
definitely not	4

Multiple Choice Scale

- Multiple response ⟶ *Check all that apply*
- Single response ⟶ *Check only one*
- Controlled response ⟶ *Check the top three*

Net Promoter Score (NPS):

Begins with a 10-point scale on likelihood to recommend. Next, the difference between promoters and dissuaders is computed.

How to Select a Scale
Things to Consider

1. The Nature of the Construct Being Measured

2. Type of Scale and Number of Scale Categories

3. Balanced vs. Nonbalanced

   – Balanced:

     • Scales with equal numbers of positive & negative categories.

   – Nonbalanced:

     • Scales weighted towards one end or the other of the scale.

4. Forced vs. Nonforced

   – Having an odd vs. even number of response choices.

Attitude Measures and Management Decision Making

- Determinant Attitudes
  - A key component to intentions
  - Those customer attitudes most closely related to preferences or to actual purchase decisions.

Types of Questioning

- Direct vs. Indirect
  - Observation

Summary

- A paradigm for developing better measures of marketing constructs

- Current practice in scale development

- The linkage among attitudes, behavior, and marketing effectiveness

- Measurement Scales

References

- McDaniel & Gates (2009), Marketing Research, 8th Edition, Wiley
- Nunnally, J. C. (1978), *Psychometric theory.* (2nd ed.). New York: McGraw Hill.
- Nunnally, J. C., & Bernstein, I. H. (1994), *Psychometric theory.* (3rd ed.). New York: McGraw Hill.
- Churchill, G. A., Jr., (1979), A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research, 16*(February), 64-73.
- Gerbing, D. W., & Anderson, J. C. (1988), An updated paradigm for scale development incorporating unidimensionality and its assessment. Journal of Marketing Research, 25(2), 186-192.
- DeVellis, R. F. (1991), *Scale development: Theory and applications.* Newbury Park, CA: Sage Publications.
- Spector, P. E. (1992), *Summated rating scale construction: An introduction.* Newbury Park, CA: Sage Publications.
- Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003), *Scaling procedures: Issues and applications.* Thousand Oaks, CA: Sage Publications.
- Clark R. A. (2006), Consumer Independence: Conceptualization, Measurement and Validation of a Previously Unmeasured Social Response Tendency, Ph.D. Dissertation, College of Business of The Florida State University.

- Davis, F. D. (1989), Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319-340.

- 賴榮裕(2006)，衡量員工電子化企業預備度之研究，
臺灣大學資訊管理學研究所未出版博士論文