

## LBS Research Online

[Jochen Wirtz](#)

Consumer satisfaction with services

Thesis

This version is available in the LBS Research Online repository: <https://lbsresearch.london.edu/id/eprint/2435/>

[Wirtz, Jochen](#)

(1992)

*Consumer satisfaction with services.*

Doctoral thesis, University of London: London Business School.

DOI: <https://doi.org/10.35065/TTDL7835>

---

Users may download and/or print one copy of any article(s) in LBS Research Online for purposes of research and/or private study. Further distribution of the material, or use for any commercial gain, is not permitted.

# **Consumer Satisfaction with Services**

**Jochen Wirtz**

**Submitted in fulfilment of the  
requirement of the degree of  
Doctor of Philosophy**

**London Business School**

**1991**

## Acknowledgements

I am most grateful and deeply indebted to my supervisor, Dr. John Bateson, who has provided invaluable support throughout the past four years. The completion of the thesis is directly attributable to John's guidance, constructive criticism and encouragement and his enthusiasm for academic research. I feel very privileged and lucky having had John as my supervisor.

A special word of thanks to Professor Andrew Ehrenberg, Dr. Mark Uncles and John Williams, who provided stimulating feedback and comments on drafts of the thesis. Their criticism and encouragement were very welcome and contributed greatly to the quality of the final version.

Furthermore, Mark has always found time to give advice on any research and career related issues. He also encouraged me to give two presentations to our marketing department and a presentation at the TIMS College on Marketing - Special Interest Conference on Services Marketing. Each presentation contributed greatly to the development of my ideas and research skills.

I would like to thank Professor Craig Smith, Vanderbilt University, for reading a draft of the review chapter on the affect literature and discussing it with me. My thanks are due also to Professor Richard Oliver and Professor Roland Rust, both from Vanderbilt University, who found the time to discuss my ideas on consumer satisfaction.

When holding a visiting appointment at London Business School in 1989/90, Professor James Carman, University of California - Berkeley, provided constructive criticism at the design stage of the study which I gratefully acknowledge.

Professor Brian Everitt at the Institute of Psychiatry in London, and Dr. Chris McManus at University College London helped me understand and use Lisrel which was important for the testing of the models presented in the thesis.

Many thanks to Nick Brown, a fellow PhD student, for writing the software for the homebanking simulator with much skill, creativity and speed.

Professor Michael Beesley provided a constant stream of interesting and challenging consulting projects without which the financing of the thesis would have been far more difficult. Thank you for looking after "your team" so well and introducing us to the intricacies of consultancy work.

I thank my fellow PhD students whose friendship contributed immensely to the enjoyment of the past years and helped me through the frustrating periods which are unavoidable when doing a PhD. Katie Truss, my office mate, filled our office with laughter and cheerfulness, and was asked, because of her exceptional language skills, to comment on numerous letters and documents that were produced during the last few years. Michael Klugherz, with whom I shared the same bench in grammar school, turned up at London Business School in my second year - what a surprise for both of us. Katie and Michael, thank you both very much for your friendship.

Finally, I would like to thank the "two women" in my life. First, my mother whose emotional and material support facilitated and encouraged my going back to school after three years of employment to do my A-levels, an undergraduate degree and finally come to London Business School. And to Jeannette, who read all about services marketing and listened to the presentations I had to practise. More than for her direct help with the thesis I want to thank Jeannette for her emotional support when things seemed to be going entirely wrong and for sharing the joy and happiness when things were going well. And thank you for giving me a pressing reason to finish in 'July 91' so that we can start our new life together in Singapore as soon as possible. Mami and Jeannette, thank you for your love which I never want to miss.

While many have contributed to the completion of this thesis, I have to admit that no research can be perfect, and I alone remain responsible for any unintended errors and weaknesses.

## Abstract: Consumer Satisfaction with Services

The concept of consumer satisfaction is a central issue in marketing and much research has been conducted. However, hardly any of this research into satisfaction has focused on the typical features of services. This thesis focuses on two of them: (1) the performance heterogeneity inherent in service encounters, and (2) their experiential nature.

Extensive research on performance heterogeneity has been conducted on the pre-choice process. Examples are the literature on decision making under uncertainty, uncertain attributes in multiattribute models and perceived risk. This thesis however, looks for the first time at the role of heterogeneity in the satisfaction process.

The main feature of services is their experiential nature. Several researchers have suggested that the commonly applied cognition based multiattribute and disconfirmation models are inadequate for capturing the experiential nature of the service encounter. Therefore, alternative models that use affect as a mediating variable between stimuli, cognitive processes and subsequent behaviour have recently been proposed. A review of the psychology and services marketing literature suggest that Russell's (1980) circumplex model is currently the best conceptualization of affect. In this thesis, it is examined whether Russell's model can be applied to service experiences, and whether affect can be included in satisfaction models with services to better capture their experiential nature.

Consumption experience with a PC-based homebanking service was simulated in a laboratory. A 2 (expected mean performance) x 2 (expected variance in performance) x 2 (actual level of performance) factorial design was employed. Measures of disconfirmation, affect and satisfaction were taken.

The results on the heterogeneity issue were surprising. All hypotheses developed had to be refuted, and alternative interpretations of the data were proposed. On the other hand, the data clearly support the inclusion of affect in satisfaction models. The implications of the findings on marketing theory and services management are discussed.

# Summary Table of Contents

	Page
Chapter I: Introduction	16
Chapter II: Consumer Satisfaction Theory Applied to Services	36
Chapter III: The Impact of Heterogeneous Service Performances on Consumer Satisfaction	85
Chapter IV: The Concept of Affect: Research and Theoretical Developments in Psychology	114
Chapter V: The Concept of Affect Applied to Satisfaction Theory with Services	153
Chapter VI: An Experimental Study: Methodological Issues and Procedures	174
Chapter VII: Data Analysis	249
Chapter VIII: Conclusions and Implications	333
Appendices	366
References	408

# Table of Contents

	Page
<b>Chapter I: Introduction</b>	<b>16</b>
1 Why Study Consumer Satisfaction?	16
2 Why Study Services and Their Features?	17
3 The Experiential Nature of Services - Emerging Perspectives on the Service Encounter	22
3.1 Environmental Psychology Approach	26
3.2 Perceived Control Theory	28
3.3 Conclusions	32
4 Objectives and Thesis Outline	33
<b>Chapter II: Consumer Satisfaction Theory Applied to Services</b>	<b>36</b>
1 Consumer Satisfaction Models	36
1.1 Disconfirmation Models and Their Components	37
1.1.1 Comparison Standards	39
1.1.1.1 Expectations	39
1.1.1.2 Ideal Performance	42
1.1.1.3 Value-Precept Disparity	42
1.1.1.4 Experience-Based Standard	45
1.1.1.5 Normative Deficit	46
1.1.1.6 Multiple Standards	47
1.1.1.7 Summary and Conclusions	48
1.1.2 Perceived Performance	51
1.1.3 Confirmation/Disconfirmation	56
1.1.3.1 Theories About the Disconfirmation Process	56
1.1.3.2 Subtractive Versus Subjective Disconfirmation	59

1.1.4	Satisfaction	61
1.1.5	Operationalization of the Components with Multiattribute-Type Models	63
1.2	Equity Theory	65
1.3	Empirical Comparisons of Alternative Satisfaction Models	67
1.4	Summary	71
2	Services Specific Features in Disconfirmation Models	73
2.1	Ambiguous Attributes Foster Consistency Effects	74
2.2	Expectations Can be Managed in the Service Encounter	76
2.3	Attribution Processes Can be Managed in the Service Encounter	77
2.4	Service Performances are Heterogeneous	79
2.5	Summary	80
3	Conclusions, Further Research and Development of Hypothesis I	82
 <b>Chapter III: The Impact of Heterogeneous Service Performances on Consumer Satisfaction</b>		 <b>85</b>
1	Why is Heterogeneity Important - An Operations Management Perspective	86
2	Performance Heterogeneity in the Choice Process - A Consumer Behaviour Perspective	89
2.1	The Impact of Performance Heterogeneity on Perceived Risk	90
2.2	Performance Heterogeneity Included in Multiattribute Models	94
2.3	Summary and Conclusions	98
3	Decision Science Literature on Variance of Expected Outcomes	99
3.1	Decision Making Under Risk - Normative and Descriptive Models	100
3.2	The Perception of Variability	105
3.3	Summary and Conclusions	107
4	Summary and Hypotheses Development on the Impact of Performance Heterogeneity on the Evaluation Stage - Hypotheses II & III	110

<b>Chapter IV: The Concept of Affect: Research and Theoretical Developments in Psychology</b>	<b>114</b>
<b>1 The Concept of Affect</b>	<b>115</b>
1.1 Definition of Affect	116
1.2 Delineation Between Affect and Cognition	117
1.3 Conceptualizations of Affect	119
1.3.1 Classifications Based on Facial Expression Research	119
1.3.2 Russell's Circumplex Model of Affect	123
1.3.3 Plutchik's Eight Basic Emotions	129
1.3.4 Comparison of the Concepts	133
1.4 Conclusions	137
<b>2 Theories on the Causes of Affect</b>	<b>139</b>
2.1 The Independent Theory	141
2.1.1 Mere Exposure Research	141
2.1.2 Unconditioned Direct Affective Responses	142
2.1.3 Conditioned Affective Responses	143
2.2 Cognition-Affect Sequence Theories	143
2.2.1 Schema-Triggered Affect	144
2.2.1.1 Interruption of Perceptual Schema	145
2.2.1.2 Social Schema and Affect	146
2.2.1.3 Affect-Charged Schemata	147
2.2.2 Appraisal Triggered Affect	148
2.2.2.1 Assessment of Consequences	149
2.2.2.2 Attribution of Causes	149
2.2.2.3 Comparison with Moral Standards	150
2.3 Conclusions	151
<b>3 Summary</b>	<b>152</b>

<b>Chapter V: The Concept of Affect Applied to Satisfaction Theory with Services</b>	<b>153</b>
1 An Introduction to the Role of Affect in Consumer Behaviour	154
2 Affect as an Important Variable in Satisfaction Models	155
2.1 Confirmation/Disconfirmation as a Direct Determinant of Affect	156
2.2 Affect as Direct Causal Antecedent of Satisfaction	157
3 Conceptualizations of Affect Applied to the Service Encounter	162
4 Inclusion of Affect in Disconfirmation Models - Hypotheses IV to VII	164
5 Summary	173
<b>Chapter VI: An Experimental Study: Methodological Issues and Procedures</b>	<b>174</b>
1 Four Basic Types of Validity	176
2 The Relative Importance of the Types of Validity	179
3 Research Designs	182
3.1 True Experiments	182
3.2 Quasi Experiments	184
3.3 Surveys	186
3.4 Selected Research Design	187
4 Research Settings	191
4.1 Trade-Offs Between the Field and Laboratory Setting	191
4.2 A Laboratory Setting is Chosen and a Homebanking Service Simulated	193
4.2.1 Performance of a Homebanking Service is Controllable and Replicable	194
4.2.2 Subjects Not Having Previous Experiences With Homebanking Could be Selected	195
4.2.3 Time is an Important Attribute of Homebanking with Value-free Perceived Performance Measures and Low Ambiguity	196

5	Subjects	201
5.1	Three Potential Objectives and Subject Selection	201
5.2	Subject Selection and Recruitment	202
6	Methodological Procedures	204
6.1	Design and Production of the Homebanking Service System	204
6.2	Experimental Sessions	207
6.3	Reason, System Description and Instructions Given to Subjects	209
6.4	Manipulation of the Independent Variables	211
6.5	Measures	218
6.5.1	Measures of Expectation, Perceived Performance and Disconfirmation	219
6.5.2	Satisfaction Measures	228
6.5.3	Measures of Affect	242
6.5.4	Measures of the Demographic Background and Other Individual Differences of Subjects	246
7	Summary	247
 <b>Chapter VII: Data Analysis</b>		<b>249</b>
1	Independent and Dependent Variables Included in the Data Analysis	250
2	Analyses Conducted Prior to Hypotheses Testing	255
2.1	Usable Questionnaires and Demographic Background of Subjects	255
2.2	Number of Subjects Per Experimental Condition and Randomization Checks	258
2.3	Reliability of Measures	260
2.4	Convergent and Discriminant Validity	265
2.5	Manipulation Check	271
2.5.1	Correlation Analysis	271
2.5.2	Cross-Tabulations and Analyses of Variance	274
2.6	Summary	280

<b>3</b>	<b>Hypotheses Testing I</b>	<b>281</b>
3.1	Correlation Analysis (Hypotheses I, IV & VI)	284
3.2	The Impact of Variance Expectations on Disconfirmation (Hypotheses II & III)	286
3.3	The Impact of Disconfirmation on Arousal (Hypothesis V)	292
3.4	The Effect of the Interaction of Arousal and Pleasure on Satisfaction (Hypothesis VII)	294
3.5	Summary	299
<b>4</b>	<b>Hypotheses Testing II: Structural Equation Analysis</b>	<b>301</b>
4.1	An Introduction to Structural Equation Modelling	302
4.2	Specification of the Hypothesized Model	307
4.3	Fitting of the Lisrel Model	312
4.4	Discussion of the Results	317
4.5	Test of Rival Models	324
4.6	Summary	331
	<b>Chapter VIII: Conclusions and Implications</b>	<b>333</b>
<b>1</b>	<b>The Impact of Expected Performance Heterogeneity on Disconfirmation</b>	<b>334</b>
1.1	Summary of Findings - Hypotheses II & III	335
1.2	Managerial Implications	336
1.3	Further Research	339
<b>2</b>	<b>The Role of Affect in the Satisfaction Process</b>	<b>341</b>
2.1	Summary of Findings - Hypotheses IV to VII	344
2.2	Further Research: The Role of Arousal in the Satisfaction Process	346
2.3	Further Research: Opening Up the Disconfirmation Paradigm	348
2.4	Managerial Implications	355
<b>3</b>	<b>The Role of Perceived Performance in the Satisfaction Process</b>	<b>357</b>
3.1	Summary of Findings - Hypothesis I	358
3.2	Research Implications	358
<b>4</b>	<b>Limitations</b>	<b>361</b>
<b>5</b>	<b>Summary</b>	<b>364</b>

<b>Appendices</b>	<b>366</b>
A. A Sample Leaflet Used for the Recruitment of Subjects	366
B. Introductory Presentation to Subjects	368
C. Slide: A Potential User of the Homebanking Service	371
D. A Sample Questionnaire	373
E. Selection of Screens of the Homebanking System	389
F. The Printout of the Lisrel Solution for the Final Model $M_3$	395
G. Further Lisrel Analysis for Testing Hypothesis I	400
H. Manipulation Checks as Key Variables in the Lisrel Model	405
<b>References</b>	<b>408</b>

# List of Tables

Table	Page
1-1 Typical Service Features and Resulting Marketing Problems	21
4-1 Izard's (1977) Taxonomy of Affective Experiences	122
4-2 Plutchik's (1980) Eight Basic Emotions	129
4-3 Comparison of Different Taxonomies of Affect	135
6-1 Major Empirical Studies on Consumer Satisfaction Using Disconfirmation-of-Expectations Models and/or Affect	189
6-2 Chronological Overview of an Experimental Session	208
6-3 System Response Time Guidelines from MIL-STD-1472C	213
6-4 Measures of Expected Performance	220
6-5 Measures of Expected Variance of Performance	221
6-6 Perceived Performance Measures	223
6-7 Disconfirmation-of-Expectation Measures	226
6-8 Attribute Specific Measures	235
6-9 Single-Item Satisfaction Measures	238
6-10 Multi-Item Satisfaction Measures	240
6-11 Russell and Pratt's Measure of Affect Adapted to Homebanking Services	244
6-12 Mehrabian's Measure of Affect Adapted to Homebanking	245
7-1 Operationalizations of the Treatment and Dependent Variables	251
7-2 Description of the Distribution of all Dependent Variables	254
7-3 Demographic Background of the Subjects	256
7-4 Number of Subjects Per Experimental Condition	259
7-5 Cronbach Alpha of All Multi-Item Scales	264
7-6 Pearson Correlation Coefficients	270
7-7 Pearson Correlation Coefficients Between Experimental Conditions and Perception Measures	273
7-8 Mean Ratings of Expected Performance Across Experimental Conditions	274
7-9 Mean Ratings of Expected Variance Across Experimental Conditions	276

7-10	Mean Ratings of Perceived Performance Across Experimental Conditions	279
7-11	Measures Selected for Hypothesis Testing	280
7-12	Hypotheses to be Tested	282
7-13	Pearson Correlation Coefficients for Overall Measures	285
7-14	Anova of Experimental Conditions on Overall Disconfirmation	287
7-15	Anova of Experimental Conditions on Response Time Disconfirmation	289
7-16	T-Tests on the Impact of the Variance Conditions on Response Time Disconfirmation	291
7-17	Tabulation of Arousal with Disconfirmation	293
7-18	Test of the Hypothesized Interaction Effect of an Interaction Between Pleasure and Arousal on Satisfaction	296
7-19	Summary of All Notations Used in the Lisrel Analysis	306
7-20	Specifications for the Hypothesized Model in Figure 7-2	311
7-21	Goodness-of-Fit Indicators for Alternative Models	316
7-22	Estimated Parameters for the Final Model $M_3$	322
7-23	Goodness-of-Fit Indicators for Rival Models	330
A-1	Specifications of the Variables for Model $M_{b1}$ with Included Manipulation Checks	400
A-2	Estimated Parameters for the Rival Model $M_{b12}$	401

# List of Figures

Figure		Page
1-1	A Model of the Service Encounter	23
1-2	The Revised Mehrabian and Russell (1974) Model of Environmental Responses	27
1-3	The Perceived Control Perspective	31
2-1	Disconfirmation Models	38
2-2	The Disconfirmation-of-Expectations Model	41
2-3	Perceived Performance as a Direct Determinant of Satisfaction?	53
3-1	A Typical Prospect Theory Value Function	103
4-1	Russell's Circumplex Model of Affect	127
4-2	An Empirical Plot of the Affective Quality of Different Environments	128
4-3	Plutchik's (1980) Structural Model of Affect	132
5-1	Incorporating Affect and Cognition in Satisfaction Models	165
5-2	The Hypothesized Relationship Between Pleasure, Arousal and Satisfaction	171
5-3	Including Russell's (1980) Conceptualization of Affect in Satisfaction Models	172
7-1	Plot of the Effects of Arousal and Pleasure on Satisfaction	297
7-2	Hypothesized Model	309
7-3	Final Model	321
8-1	Inclusion of Experiential Variables in Satisfaction Models	354
A-1	Hypothesized Model (including Manipulation Checks)	403
A-2	Final Model (including Manipulation Checks)	404
A-3	Rival Hypothesized Model	406
A-4	Rival Fitted Model	407

# Chapter I

## Introduction

This thesis deals with the extension of the standard satisfaction model in consumer behaviour so as to include two features typical of services, namely heterogeneous performance expectations and the experiential nature of the service encounter. This first chapter provides an overview showing where the thesis fits in relation to the services marketing literature. A thesis outline with a rationale for subsequent chapters is given at the end of this chapter.

### 1 Why Study Consumer Satisfaction?

Consumer satisfaction is a central issue in marketing thought and practice. Satisfaction is a major outcome of marketing activity and serves to link processes of decision making and consumption with postpurchase phenomena such as attitude change, complaining behaviour and word-of-mouth, repeat purchase and brand loyalty (Bearden and Tell 1983, Oliver 1980a). The centrality of consumer satisfaction is reflected by its inclusion in the marketing concept, which focuses on profit generation through "determining the needs and wants of target markets and delivering the desired satisfaction" (Kotler 1989, p. 14). The need to translate this philosophical marketing concept statement into operational guidelines has kindled extensive model building, development of measures and empirical research

in the area of consumer satisfaction since the early 1970s. In spite of the intensive research in this area over the past 20 years, many issues still remain unanswered.

## 2 Why Study Services and Their Features?

The service sector has been increasing in importance in most economies. In particular, the service sector has come to provide the bulk of new employment and has absorbed all the jobs shed by agriculture, mining and manufacturing (Riddle 1986). Today, in most industrialized economies, the service sector has outgrown other sectors in terms of both employment and value of output (Heskett 1986). The growing importance of services has also been reflected in the increasing interest in the marketing problems of service firms (Bateson 1989). In fact, services marketing has become a recognized and accepted subset of the marketing discipline (Kotler 1989; Zeithaml, Parasuraman and Berry 1985).

Services have four main features that greatly affect the marketing activities of service firms (Kotler 1989). These features, which are consistently cited in the services marketing literature, are: intangibility, inseparability of production and consumption, heterogeneity of performance and perishability (for a review refer to Zeithaml, Parasuraman and Berry 1985).

First, services are intangible. Because services are performances, rather than objects, they cannot be seen, heard, smelled, tasted or touched in the same manner in which

goods can be sensed (Bateson 1989, Levitt 1980). The consequence of this intangibility is that consumers have only a few reference points for assessing the quality of services. This, in turn, leads consumers to rely on cues associated with the service production process (procedures, mechanisms and flow of activities), the participants (service employees and other customers present in the service setting) and the physical evidence (the service environment and all tangible cues). In fact, Bitner (1990) suggests that the marketing mix for services, consisting of the four traditional elements (product, price, place and promotion), should be expanded to include these three new ones (process, participants and physical evidence). Although these three elements could be encompassed within the traditional mix, Bitner argues that separating them out draws attention to the factors that are of particular importance to service firms.

Second, services are inseparable, meaning that they are typically produced and consumed at the same time. This has a number of important implications. (1) Consumers form an integral part of the service production process. Their participation may be active or passive, but they are always present (Bateson 1989). In many cases, the input of the consumer is crucial. For example, for a haircut one has to make an appointment, arrive at the hairdressing salon in time and communicate to the hairdresser what kind of style one would like to have, before the actual haircut can be carried out. Mills and Moberg (1982) even regard consumers as "partial employees" of service organizations, and propose increasing productivity through more consumer participation in the production process (refer also to Bateson 1985; Lovelock and Young 1979). (2) Services are time and place

dependent and this has important implications for capacity management (the service has to be produced while the customer is present) and productivity (the "factory" has to be close to the customer which makes centralized production difficult) (Bateson 1989). (3) Real-time production and consumption make quality control difficult, because bad services cannot be rejected before they get to the consumer (DeSouza 1989, Parasuraman et al 1985). As with any manufacturing operation, it is impossible for a service production to be 100 per cent successful all of the time, which means consumers will occasionally be exposed to inferior quality services (Bitner 1987).

Third, services are heterogeneous. In contrast with products, services show a high degree of unit-to-unit variability, because of the problems of quality control referred to earlier, and the reliance on the technical and social capabilities of service employees. The performance of service employees can vary considerably from one employee to another within the same firm and, even for the same individual, the performance can fluctuate drastically over time (Krughoff 1981).

Solomon and Surprenant (1985) applied role theory to the service encounter and suggest that the congruity between employee and consumer behaviour is an important determinant of the service production process. As role behaviour varies according to the individual's role expectations and capabilities (Solomon and Surprenant 1985), one can infer that not only employee but also consumer behaviour contributes to the heterogeneity of services.

The review of the satisfaction literature in Chapter II showed that no research had been conducted to understand the impact of heterogeneous performance expectations on the satisfaction process. As this seemed to be an important and interesting issue in services marketing, it was selected as one focal point of this thesis and was examined for the first time in the context of consumer satisfaction. The effects of performance heterogeneity on consumer behaviour and operations management are reviewed in Chapter III, and two hypotheses are developed on its role in the satisfaction process.

Fourth, services are perishable. Although perishability is a direct consequence of the inseparability feature of services, researchers have often referred to it separately because of its important implications. Perishability means that services cannot be stored, which is linked to problems such as capacity management and operations efficiency (Bateson 1989).

It has been suggested in the literature, that these four features of services lead to specific problems for service firms. Table 1-1 summarises the problems that have frequently been addressed by researchers.

In addition to the four features presented, a fifth feature, which can be viewed as a direct consequence of the intangibility and inseparability of services, is their experiential nature (Bateson 1989, p. 6; Hui 1988, p. 14-7). This feature is discussed in more detail in the next section, as one of the key issues in this thesis.

**Table 1-1: Typical Service Features and Resulting Marketing Problems**

---

<b>Typical Service Features</b>	<b>Resulting Marketing Problems</b>
<b>Intangibility</b>	<ul style="list-style-type: none"><li>- Cannot protect service through patents</li><li>- Cannot readily display or communicate services</li><li>- Prices are difficult to set</li><li>- Consumers have only few reference points to assess service quality</li></ul>
<b>Inseparability</b>	<ul style="list-style-type: none"><li>- Consumer is involved in the production process</li><li>- Other consumers are involved in production</li><li>- Services are time and place dependent</li><li>- Quality control is difficult</li></ul>
<b>Heterogeneity</b>	<ul style="list-style-type: none"><li>- Standardization of the service is difficult</li></ul>
<b>Perishability</b>	<ul style="list-style-type: none"><li>- Services cannot be stored</li></ul>

---

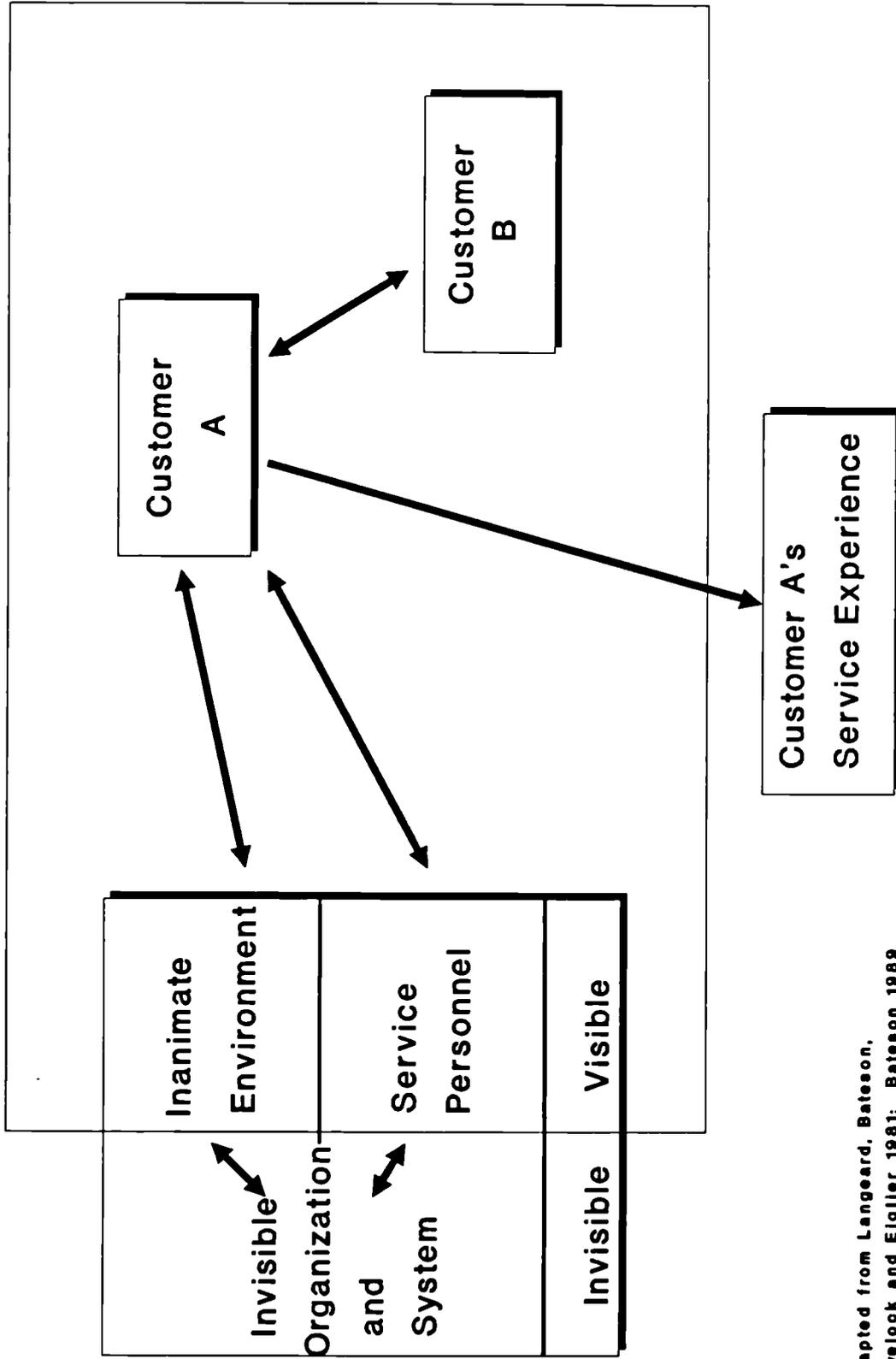
(Adapted from Zeithaml, Parasuraman and Berry 1985)

### 3 The Experiential Nature of Services - Emerging Perspectives on the Service Encounter

The term "service encounter" was coined to describe the moment of interaction between the customer and the firm (Czepiel, Solomon and Surprenant 1985) or the time during which a consumer directly interacts with a service (Shostack 1985). This term "service encounter" therefore includes all the interactions between a customer and the components of a service production and consumption process. In many cases a discrete encounter constitutes the service itself from the consumer's point of view (Bitner 1990). Eiglier and Langeard (1977) have developed a model of the service encounter (Figure 1-1), which captures the interactive nature of the service production and consumption process, and the components of the service encounter. The model emphasises that services are the output of a multitude of interactions between the customer and service personnel, the service environment, the physical support provided by the service organization, and with other customers in the service setting.

One obvious implication of this interactive perspective is that services are experiential in nature. In fact, the service experience per se can be regarded as the major output or product of service organizations (Hui 1988). As Shorr, then executive vice president-marketing of Holiday Inns Inc, expressed it: "What I am selling, in terms of what people are buying, is a hotel experience (in Knisely 1989)."

Figure 1-1: A Model of the Service Encounter



Adapted from Langeard, Bateson, Lovelock and Egitler 1981; Bateson 1989

It is of paramount importance for service providers that they understand and manage their customers' experiences. As illustrated in the following example, this need is far more pressing for service firms than for product marketers. The consumption experience of eating a chocolate bar while being surrounded by rushing people, traffic noise and diesel fumes is completely different from the experience of eating a chocolate bar while enjoying the sun in a deck-chair in Regent's Park and listening to the band. The chocolate manufacturer has virtually no possibility of influencing this consumption process.

A service provider on the other hand, has considerable opportunities to manage the multitude of interactions that together make up a consumption experience: the interactions between the customer and service employees, those between the customer and the service setting, and those between customers. For example, the service experience can be managed by designing and managing the interactive production process (Shostack 1984), by selecting, training and managing the service employees (Berry 1981, Bitner 1990, Bowen and Scheider 1985, Schneider 1980, Shamir 1980), by designing and maintaining the service environment (Bitner 1990, Bruner 1990, Donovan and Rossiter 1982, Milliman 1982), and by selectively targeting, socializing and educating customers (Lovelock 1981, Mills and Moberg 1982).

For better understanding of consumer behaviour and guiding management, it is important to capture this experiential nature of services. However, most of the research has focused only on the cognitive processes driving consumer behaviour.

In the context of consumer satisfaction as well, the disconfirmation and multiattribute models that have been used (reviewed in Chapter II) are predominantly cognition based, and most of the research only looks at the cognitive processes leading to satisfaction. These information processing perspectives do not capture the experiential nature of services adequately (Hui 1988), and as Holbrook and Hirschman (1982, p. 132) put it, the "fantasies, feelings and fun" part of many service experiences are ignored.

As a response to the shortcomings of cognition based models in explaining consumer behaviour in the service encounter, theorists have recently proposed alternative perspectives which try to capture the experiential nature of services. This work has been based on a comprehensive model of human-environment relations (Mehrabian and Russell 1974; Russell 1978, 1980), which is considered a leading model in the broad area of environmental psychology. Applied to the service encounter they are called the environmental psychology approach (Donovan and Rossiter 1982) and perceived control theory (Bateson 1985b, Hui 1988). Both of these perspectives use affect (the way people feel while consuming a service) as a mediating variable between stimuli, cognitive processes and response behaviour. Both are reviewed in the next two sections.

### 3.1 Environmental Psychology Approach

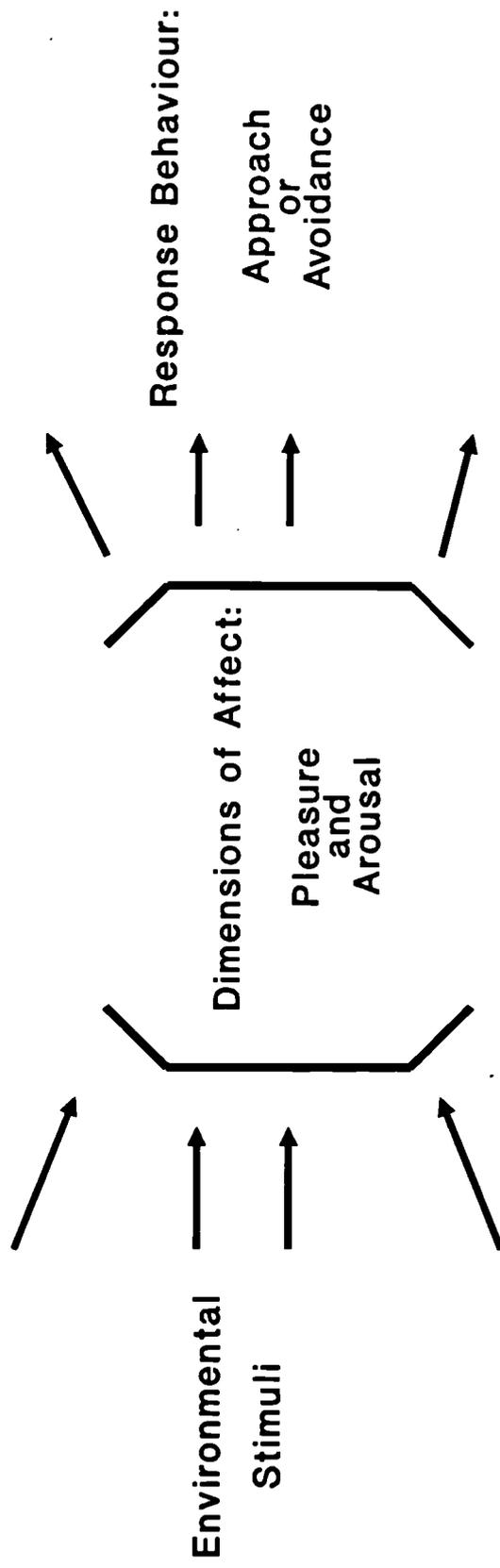
Russell (1980) and Russell and Pratt (1980) have proposed a comprehensive model of human-environment relations. They suggest that there are two basic bipolar emotional dimensions, pleasure and arousal, which modulate all human behavioural responses towards or within a physical environment. This behaviour can be classified as either approach or avoidance behaviour. Approach (avoidance) behaviour is considered to have four aspects: (1) the willingness or desire to stay in (get out), (2) exploration (remain inanimate), (3) communication with others (avoid communication), and (4) interact supportively and perform well (perform badly) in the environment. The model is schematically depicted in Figure 1-2.

Approach-avoidance behaviours are the result of emotional states individuals experience within the environment. According to Russell's (1980) circumplex model of affect<sup>1</sup>, all emotional states can be represented by some combination of two major dimensions - pleasure and arousal. Pleasure and arousal are hypothesized to interact in such a way that arousal amplifies approach behaviour in pleasant environments and avoidance behaviour in unpleasant environments (Donovan and Rossiter 1982).

---

<sup>1</sup> Russell's (1980), and Russell and Pratt's (1980) model of affect is a further development of the original Mehrabian and Russell (1974) model.

**Figure 1-2: The Revised Mehrabian and Russell (1974)  
Model of Environmental Responses**



Sources: Mehrabian and Russell (1974)  
Russell and Pratt (1980)

Applied to services marketing, Donovan and Rossiter (1982) suggest that cognitive influences such as price, location, variety and quality of merchandise can largely account for retail store selection. However, once within the store, emotional responses are the primary determinant of customer behaviour. The authors showed that simple affect or store-induced pleasure can be a very powerful predictor of approach-avoidance behaviour within a store.

The Russell model is strong in its treatment of mediating variables and response areas, but leaves the problem of an appropriate stimulus taxonomy largely untouched (Donovan and Rossiter 1982). From a managerial point of view, this model provides a good framework for testing the effects of specific stimuli in the service encounter. In other words, the service environment can be systematically manipulated, and via measuring the intervening variables, the behavioural consequences can be predicted and evaluated.

### **3.2 Perceived Control Theory**

White (1959) proposed that control is a principal human driving force. People are motivated to demonstrate their competence and superiority over their environment. Schutz (1966) suggests that social behaviours are driven by three interpersonal needs: inclusion, control and affection. And Proshansky et al (1970) suggest that environmental quality is determined by the extent of choice and control an environment can provide to its inhabitants. These researchers have shown with

their research that control is a crucial dimension in interpersonal and human-environment interactions.

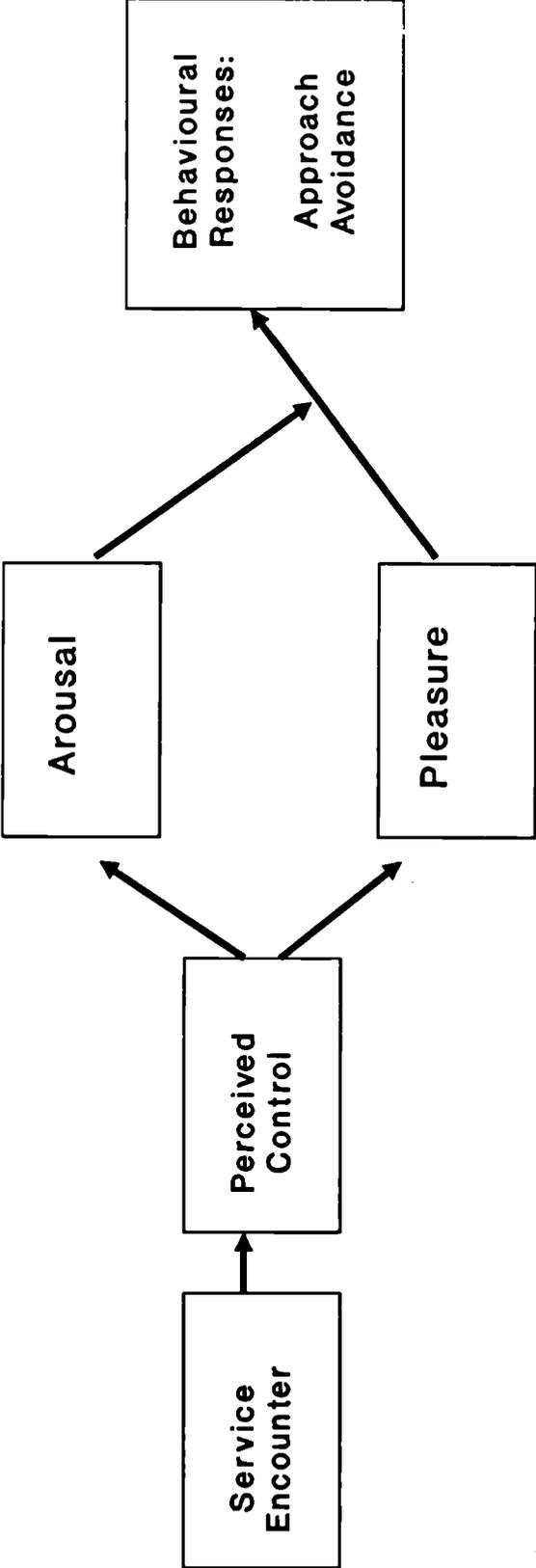
As the service encounter consists of these interactions, control has also been suggested as a crucial determinant of consumer behaviour with respect to services. The concept of control was first introduced by Bateson (1985b) to explain consumer behaviour in the purchasing and consumption of services. The control perspective suggests that in the service encounter, consumers' affect and behaviour are shaped by their level of perceived control. The basic argument of this model is that the higher the consumer's level of perceived control during the service encounter, the more positive will the consumption experience be perceived (Bateson 1989).

Bateson (1985b), and Bateson and Hui (1988) applied Averill's (1973) conceptualization of control in the context of the service encounter. Averill's (1973) threefold typology of control is one of the most extensively used in psychology. He has identified three different types of control from the existing literature: behavioural, cognitive and decisional control. Behavioural control is the actual or perceived availability of a response that may directly influence or modify the objective characteristics of an event. Cognitive control refers to the processing of potentially threatening information in such a way as to reduce the net long-term stress. This includes the perception of predictability (Miller 1981) as well as to the appraisal of the situation at hand and its implications. Decisional control is defined as actual or perceived ability to select or change outcomes or goals.

Rather than treat perceived control as a service characteristic in multiattribute type models, Bateson (1985b) and Hui (1988) regard it as a global indicator which summarizes the perceptual aspects of a service experience. They therefore consider perceived control as a crucial antecedent of any ensuing affective and behavioural responses to the service encounter. Their view is based on a revised Mehrabian and Russell (1974) model of affect. Bateson and Hui (1988) suggest that perceived control intervenes between the environment and human affective responses (Figure 1-3).

Empirical research has provided substantial support for the perceived control perspective. For example, consumers choose to participate more extensively in the production of a service (e.g. by choosing a self-service option) not because of monetary incentives but because they perceive a stronger sense of control as a result of their participation (Bateson 1985b). Furthermore, it has been shown that an induced sense of control can reduce the level of perceived crowding (Bateson and Hui 1988, Langer and Saegert 1977) and shape consumers' affect in the service encounter (Hui 1988). These studies show that perceived control can significantly influence consumers' decision making and behaviour in the service encounter.

Figure 1-3: The Perceived Control Perspective



Adapted from: Hui (1988)

### 3.3 Conclusions

The models currently employed in the context of consumer satisfaction (i.e. disconfirmation and multiattribute models) capture only cognitive dimensions of the service encounter. Any affective response during the consumption process is disregarded. These information processing perspectives therefore fail to adequately capture the experiential nature of services (Hui 1988), and "fantasies, feelings and fun" are ignored (Holbrook and Hirschman 1982, p. 132).

The environmental psychology and the perceived control perspectives on the other hand, try to capture affective responses that are intrinsic to the service experience and occur during the consumption process. The underlying conceptual framework has been provided by Mehrabian and Russell (1974), and further developed by Russell (1978, 1980), who propose affect as a mediating variable between stimuli and response behaviour. Russell's model has been applied to the service encounter and shown to explain consumer behaviour (e.g. Bateson and Hui 1988, Hui 1988, Donovan and Rossiter 1982).

In environmental psychology affect plays a central role as a mediating variable between stimuli and behaviour. In the context of consumer satisfaction, very little effort has been devoted to the modelling of affective responses during a consumption experience (Westbrook 1982, 1987), or to explaining the link between the cognitive processes and affect (Woodruff et al 1983). Nevertheless, the bringing together of these two different approaches seems particularly promising, as satisfaction

models, which seem to be good at capturing the cognitive dimensions of a consumption process, could be extended by including affect, which may be a powerful concept for capturing the way people feel while consuming a service. The question concerning satisfaction research is whether affect can and should be incorporated in the currently employed satisfaction models to better capture the experiential nature of services. This is one of the main themes of this thesis.

#### 4 Objectives and Thesis Outline

As service industries continue to grow in importance, consumers increasingly voice frustration and dissatisfaction with their service experiences (Koepp 1987). The level of consumer satisfaction depends directly and most immediately on the management and monitoring of individual service encounters (Parasuraman, Zeithaml and Berry 1985; Surprenant and Solomon 1987). For this, an understanding of the satisfaction process with services is needed. However, hardly any of the research into satisfaction has focused on the features typical of services. It is the aim of this thesis to contribute to the understanding of the satisfaction process with services.

This thesis focuses on the impact of two service specific features on consumer satisfaction. The classical satisfaction model, the disconfirmation-of-expectations model, is extended to include (1) heterogeneous performance expectations, and

(2) affect. The inclusion of affect is aimed at capturing the experiential nature of services.

A methodological issue concerning perceived performance measures is examined in the context of the classical disconfirmation model. This issue arose in the review of literature on satisfaction (Chapter II) and concerns conclusions that were drawn on the basis of inadequate perceived performance measures. This issue is relevant to satisfaction research in general and not only to services, and is of secondary importance in the thesis. It was included as it could be conveniently examined without jeopardizing the pursuit of the two central objectives of the thesis as outlined in the previous paragraph.

### Thesis Outline

In Chapter II, the mainstream satisfaction literature is reviewed. Then, the impact of some services specific features on the disconfirmation-of-expectations model, the classical conceptualization of the satisfaction process, are discussed. One outcome of this discussion is the fact that no research has looked at the impact of heterogeneous performance expectations on the disconfirmation process, which is to be examined in the thesis. Finally, a methodological issue is raised about the perceived performance measures generally employed, and a hypothesis is developed.

In Chapter III, the literature on performance heterogeneity in the area of consumer behaviour is reviewed. Two hypotheses are developed on the role of expected performance heterogeneity in the satisfaction process.

Affect seems to be the crucial variable for capturing the experiential nature of services in the recently developed perspectives of the service encounter. To gain a deeper insight into the concept of affect and its causes, relevant research and development in environmental, social and cognitive psychology is reviewed in Chapter IV.

Chapter V examines the relationship between consumer satisfaction and affect and reviews studies of the service encounter that used affect to explain consumer behaviour. As an outcome of this review, it is proposed that affect, or the way people feel during the service encounter, can be included in the standard satisfaction model. Four hypotheses are developed.

In Chapter VI, the reasons for the selected methods are outlined. The data analysis is presented in Chapter VII, and the conclusions and implications of the findings are discussed in the final chapter, Chapter VIII.

# **Chapter II**

## **Consumer Satisfaction Theory**

### **Applied to Services**

The purpose of this chapter is to obtain a clear understanding of the consumer satisfaction literature and its application to services. In the first section, the main stream satisfaction literature is reviewed, and it is concluded that the disconfirmation-of-expectations model is the best available conceptualization of consumer satisfaction. Section two discusses the impact of a number of widely recognized features of services on the satisfaction process. In the last section, a hypothesis concerning the role of perceived performance in the satisfaction process is put forward, and the two main research issues to be addressed in this thesis are outlined.

#### **1 Consumer Satisfaction Models**

All satisfaction models are based on some kind of a comparison process. In the first subsection, a review of disconfirmation models, which are the common conceptualization of satisfaction, is presented. These models use confirmation or disconfirmation as a mediating variable between a pre-consumption comparison standard, perceived performance and satisfaction. In the second subsection, equity

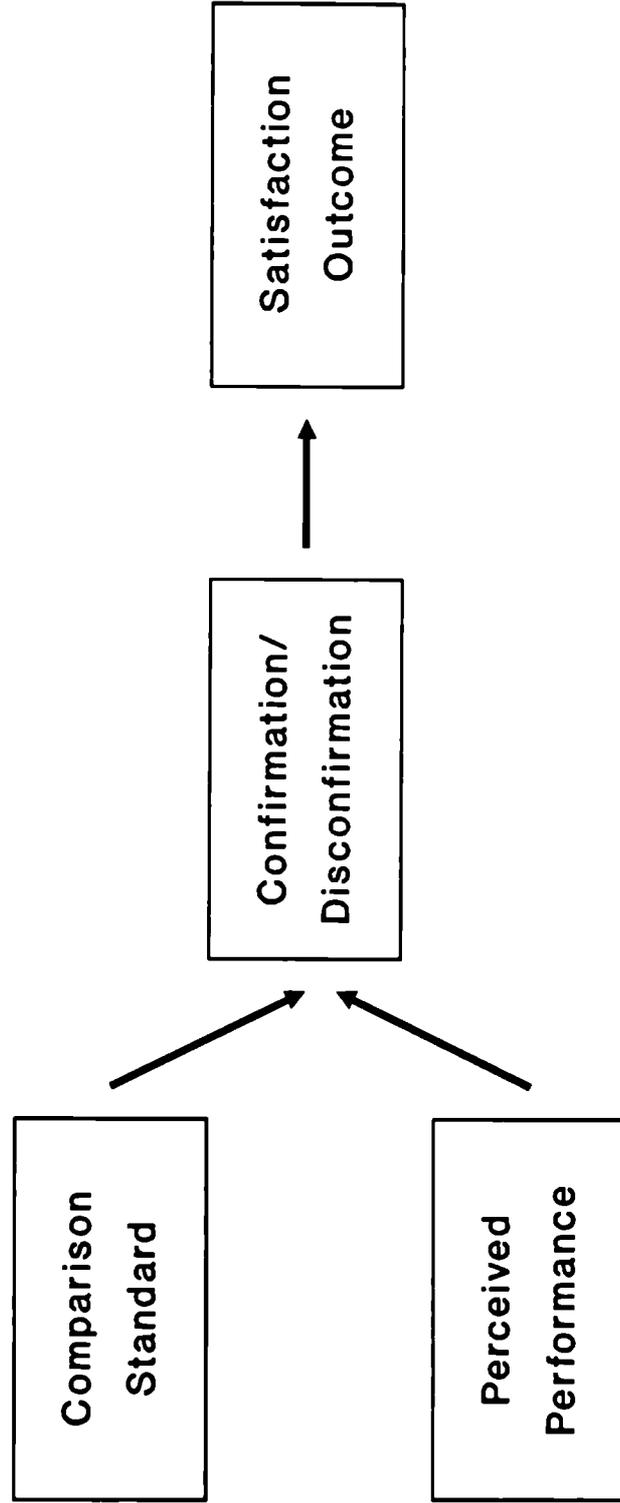
theory, an alternative conceptualization, is discussed. This theory is distinguished from other models because it does not incorporate any pre-experience standard or perceived performance. The net gains of consumer and seller are conceptualized as the driving force of satisfaction. The third subsection looks at empirical studies comparing alternative models. A summary is provided in the last subsection.

### **1.1 Disconfirmation Models and Their Components**

Numerous models have been proposed to examine the antecedents of satisfaction and provide theoretical explanations of consumer satisfaction. The vast majority of researchers have employed the confirmation/disconfirmation paradigm. In this paradigm consumers evaluate consumption experiences and make satisfaction decisions by comparing perceived performance with some pre-consumption standard (Figure 2-1). The level of satisfaction is related to the size and direction of the disconfirmation experienced. Satisfaction occurs when performance matches or is better than the standard. Dissatisfaction occurs when performance is less than the standard.

In each of the next four subsections, the basic components of disconfirmation models are discussed. These components are the pre-experience comparison standard, perceived performance, confirmation/disconfirmation and satisfaction. In the last subsection, their operationalization with multiattribute-type models is reviewed.

**Figure 2-1: Disconfirmation Models**



### **1.1.1 Comparison Standards**

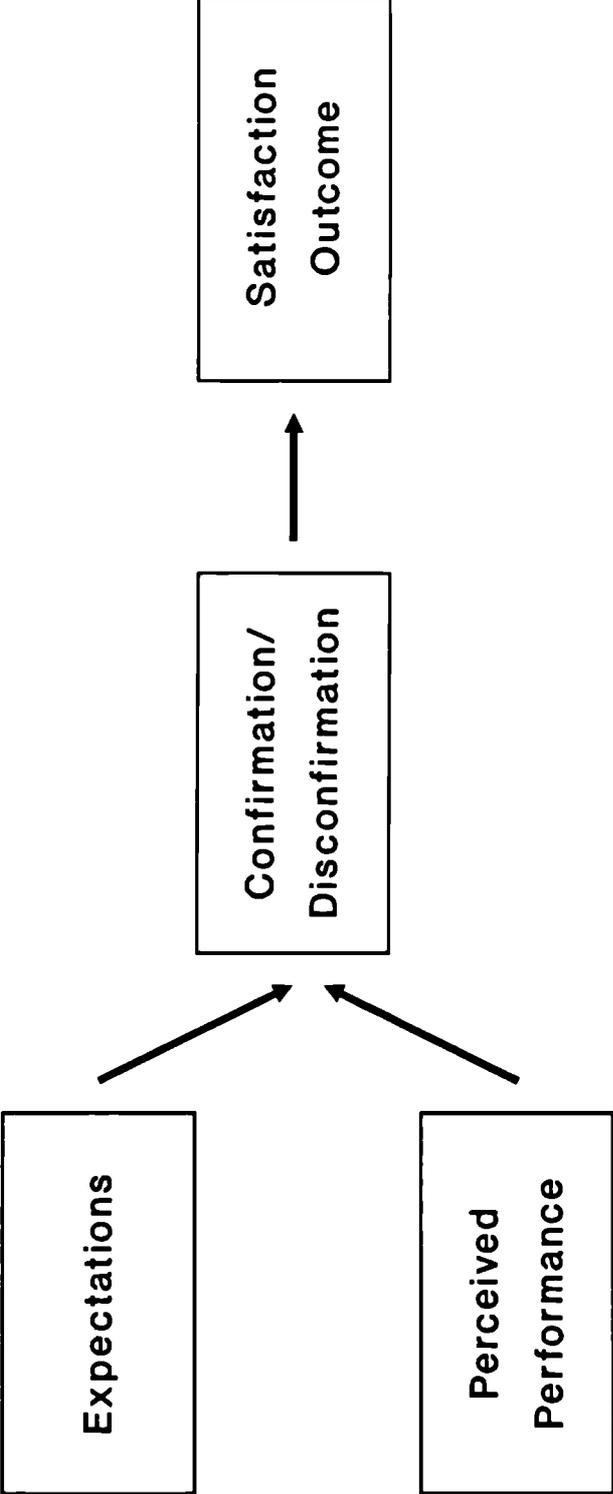
The most controversial component in disconfirmation models is the pre-experience comparison standard. A lot of theoretical debate and empirical research has revolved around the question: "What standard or standards do consumers use in the comparison process?" Numerous models have been developed and tested. The conceptual development has resulted in four broad classes of pre-experience standards. They are expectations (e.g. Bitner 1987, Day 1977, Liechty and Churchill 1979, Oliver 1980a, Woodruff et al 1983), ideal performance (e.g. Miller 1977, Sirgy 1984), needs and wants coined as value-precept (Westbrook and Reilly 1983, Swan and Trawick 1980) and experience-based standards (Cadotte et al 1987, LaTour and Peat 1979a and b, Woodruff et al 1983). These standards are presented in the next four subsections. Two further, less well corroborated, standards are reviewed in subsections five and six. They are the normative deficit standard, advanced by Swan (1983), who suggests comparisons with social norms, and empirically based multiple standards (e.g Sirgy 1984, Tse and Wilton 1988).

#### **1.1.1.1 Expectations**

The disconfirmation model with expectations as the comparison standard (Figure 2-2) has received the strongest empirical support and has been most widely applied. Theorists generally agree that satisfaction can be conceptualized as the consumer's response to the experienced discrepancy between pre-experience expectations and

performance (e.g. Bitner 1987, Churchill and Surprenant 1982, Day 1984, Oliver 1980a, Tse and Wilton 1988). Expected product performance represents the most likely performance (Miller 1977; Liechty and Churchill 1979) or predicted performance (Woodruff, Cadotte and Jenkins 1983). Some researchers refer to this comparison standard also as the focal brand performance standard. Expectations are a function of a consumer's frame of reference (Bitner 1987), which is determined by his/her experience with the product (Miller 1977), the observed experiences of other consumers, word-of-mouth (LaTour and Peat 1978) and advertising effects (Olson and Dover 1979).

**Figure 2-2: The Disconfirmation-of-Expectations Model**



### **1.1.1.2 Ideal Performance**

This standard was derived from ideal point models of consumer preference and choice (e.g. Holbrook 1984). Instead of comparing performance with expected or most likely performance as in the disconfirmation-of-expectations model, ideal performance theory suggests that consumers use as a comparison standard the optimal possible performance, the performance that best "can be," or the performance a consumer could "ideally hope" for (e.g. Sirgy 1984). Some empirical evidence for this standard has been furnished by Sirgy (1984) and Tse and Wilton (1988).

One limitation of this standard seems to be that consumers could rarely perceive any performance better than the standard, as hardly any firm can perform better than its customers could "ideally hope for."

### **1.1.1.3 Value-Precept Disparity**

The value-precept standard draws on the most basic concepts underlying marketing, which are those of human needs and wants. A human need is a state of perceived deprivation in a person, which takes the form of wants shaped by culture and individual personality. A motive is a need that is sufficiently pressing to direct a person to seek out satisfaction of this need, whereby the consumer is motivated

towards products and services that he/she believes will satisfy this drive (Kotler 1989, p. 5-6). So far, comparison standards have not included consumer needs.

Westbrook and Reilly (1983) consider it a major problem that the disconfirmation-of-expectations model does not provide sufficient differentiation between cognitive and evaluative notions. Expectations refer to beliefs about the most likely performance of a product or service. However, what is expected in a product may or may not correspond to what is wanted or desired. Product breakdown, improper function and unattractive appearance produce dissatisfaction regardless of whether they are expected. In practice, values and expectations often fall together because consumers choose purposefully to achieve their goals. When values and expectations have been separated experimentally, values rather than expectations determined satisfaction (Locke 1967). In other words, performance in relation to needs and wants rather than expectations appear to be the primary determinant of satisfaction (Westbrook and Reilly 1983).

For example, a consumer might be forced, due to time and location constraints, to dine at a restaurant that does not possess the wanted and valued attributes. Although the dining experience may be as expected, it is unlikely that this consumer will be fully satisfied because he/she probably compares this dining experience with one at a more highly valued restaurant (LaTour and Peat 1979a, Woodruff et al 1983).

Westbrook and Reilly (1983) drew on Locke's (1967) seminal analysis of job satisfaction. Locke asserts that satisfaction/dissatisfaction is a response triggered by a cognitive-evaluative process in which the perceptions of an object, action or condition are compared to one's values (or needs and wants). The smaller the discrepancy between perceptions of the object, action or condition and one's values, the more favourable is the evaluation and the greater the level of satisfaction. Conversely, the greater the value-precept disparity, the less favourable the evaluation, followed by goal frustration and dissatisfaction.

In the context of marketing, value-precept disparity has been operationalized as the extent to which a product provides the features and performance characteristics needed by a consumer (Westbrook and Reilly 1983). The researchers used a 7-interval semantic differential scale anchored with "provides far less than my needs" and "provides exactly what I need." Their model was not corroborated by the experiment. However, in the same experiment the empirically well supported disconfirmation-of-expectations model was not corroborated either. This might be an indication of problems with the method of investigation. All measurements were taken a long time after the actual consumption experience.

The difference between the ideal performance and the value-precept disparity models lies in the implicit consideration of needs and wants in the second model. Instead of using the optimal possible performance as a comparison standard, the value-precept disparity model employs the level of performance that is required to fulfil the consumer's needs and wants.

#### 1.1.1.4 Experience-Based-Standard

Woodruff, Cadotte and Jenkins (1981, 1983) proposed a comparison standard based on consumers' prior experiences. Furthermore, LaTour and Peat (1979a) suggested and supported empirically (1979b) that the comparison standard can be driven by prior experiences with a focal brand and with similar brands in a product category. Both papers indicate that the comparison standard can be influenced by perceived capabilities of brands other than the one purchased, and that the standard may be more than just expectations about the most likely or predicted performance of a focal brand.

For example, Woodruff et al (1983) suggest that breadth of experience may cause consumers to form standards that establish what focal brands should be able to achieve, and that each individual consumer may use some absolute comparison standard against which all experiences with a given product are compared. This standard may differ significantly from the expected or predicted performance of the focal brand. In general, Woodruff et al (1981) propose that a comparison standard can be based on (1) a brand unit (e.g. a particular McDonald's restaurant), (2) other units of the same brand (e.g. several other McDonald's restaurants), (3) similar brands (e.g. Burger King's and Wimpy's restaurants), (4) a range of brands competing for the same use situation (e.g. fish and chips restaurants) or even (5) a whole class of products competing for the same basic needs or wants (e.g. all restaurants).

Although Cadotte et al (1987) provided some empirical support for an experience-based standard, and although the conceptual framework seems promising, Woodruff et al (1983) and Cadotte et al (1987) failed to furnish a convincing operationalization of their standard. In their 1987 study, they measured beliefs about the "typical" performance, best brand performance, focal brand performance and focal brand expectation. However, they did not present nor explain how an experience-based standard could be developed from these measures.

#### 1.1.1.5 Normative Deficit

Normative deficit theory conceptualizes satisfaction as a comparison between perceived performance and perceptions of social norms (Swan 1983). For example, a consumer might be dissatisfied with a compact car if he or she feels that a person of his/her status should be driving a luxury car. Morris (1976) applied this theory to satisfaction with current housing. He hypothesized that satisfaction would be a function of the discrepancy between the perceived attributes of one's current house and the attributes that one's house should have, i.e. the attributes that are socially acceptable. For example, a family with older children should have a bedroom for each child. If their current house does not provide enough bedrooms, the family will be less satisfied.

So far, little empirical work has been done to validate normative deficit theory. Furthermore, Bitner (1987) suggests that it would have limited application because,

for the majority of products and services, social norms either do not exist or are only vaguely defined.

#### 1.1.1.6 Multiple Standards

Sirgy (1984) first introduced the concept of multiple comparison standards. Oliver (1985) as well as Wilton and Nicosia (1986) provided further conceptual support, and Tse and Wilton (1988) showed empirically that more than one comparison standard can be involved in satisfaction formation. Both expectations ( $r=.39$ ,  $p<.01$ ) and ideal standard ( $r=-.25$ ,  $p<.01$ ) were significantly correlated to satisfaction. Path coefficients showed that expectations exerted significant influence on disconfirmation and satisfaction, whereas ideal standards significantly affected perceived performance.

This conceptualization of satisfaction explicitly allows for interactions and/or temporal changes in key satisfaction determinants. Tse and Wilton (1988) suggest that post-experience comparison processes should not be modelled with a state variable, but as a continuous process involving different standards as the effects of the product consumption decay. However, they do not provide an adequate theoretical framework that determines which standards are to be included at each stage of the process. In consequence, operational models have not been developed. Much conceptual and empirical work will be necessary before models incorporating multiple standards can be considered alternatives to the models currently used.

### 1.1.1.7 Summary and Conclusions

From this review, the expected, value-precept and experience-based standards emerge as the most promising comparison standards. They have some conceptual basis, intuitively appealing theories and at least some empirical support.

Expectations have received much empirical support and have been the most frequently used standards in both applied and theoretical research. However, there are a number of situations in which expectations do not seem to be appropriate standards. LaTour and Peat (1979a) presented an example in which a consumer may be forced to purchase a brand whose attributes are not evaluated favourably (e.g. because the preferred brand is not available). Although the consumer's expectations of the purchased brand may be confirmed by the consumption experience, the consumer will still be dissatisfied with the product because of his/her unfavourable evaluation of its attributes. The disconfirmation-of-expectations model, however, predicts a satisfying consumption experience for this case.

Furthermore, it has been criticized that the disconfirmation-of-expectations model holds that previous experiences within the product category merely influence the accuracy of expectations of the focal brand performance, and no influence is predicted on the satisfaction outcomes. However, a number of theorists predict otherwise. For example, experiences with brands superior to the focal brand are suggested to raise the comparison standard used in the disconfirmation process (LaTour and Peat 1979a and b, Woodruff et al 1983).

The first of these limitations of the expectations standard seems to be addressed by value-precept theory, which explicitly proposes values, deriving from needs and wants, as standards. The second limitation is addressed by the experience-based standard, which holds that the comparison standard is shaped by the breadth of experience a consumer has. However, the experience-based standard disregards needs and wants in the same way as the expectations standard, and the value-precept standard does not reflect the reality of available product performances on the market, which would set the limits for the performances a consumer believes a focal brand should provide (Cadotte et al 1987). Here the theorists have the implicit assumption that consumers would be satisfied with a realistically achievable performance, even when needs and wants are not fully met.

In 1987, Cadotte, Woodruff and Jenkins developed a synthesis of the experience-based and the value-precept standards, and proposed what they called "experience-based norms." They suggest that consumers rely on a standard with two characteristics: (1) It reflects the desired performance in meeting needs/wants (as with the value-precept standard), and (2) it is constrained by the performance consumers believe possible, based on their breadth of consumption experiences (comparable to the experience-based standard).

From a conceptual point-of-view, this extended experience-based norm seems promising, however, much further conceptual and measurement research will be needed before it can be used. In their paper, Cadotte et al (1987) did not indicate

how an experience-based norm could be developed from the constructs measured in their study.

The development of alternative standards seems to indicate that researchers and practitioners have to be more careful when applying expectations as a comparison standard in satisfaction models. According to the review in this section, this appears to be particularly true in situations where consumers cannot choose the preferred brand which would best meet their needs and wants, or where they do not know what performance they can realistically expect in this product category. In other words, expectations seem to be a good pre-experience comparison standard when consumers have free choice and sufficient pre-purchase experience or information, so that they can choose purposefully and match their needs and wants with expectations and know about the performance they can realistically expect.

### 1.1.2 Perceived Performance

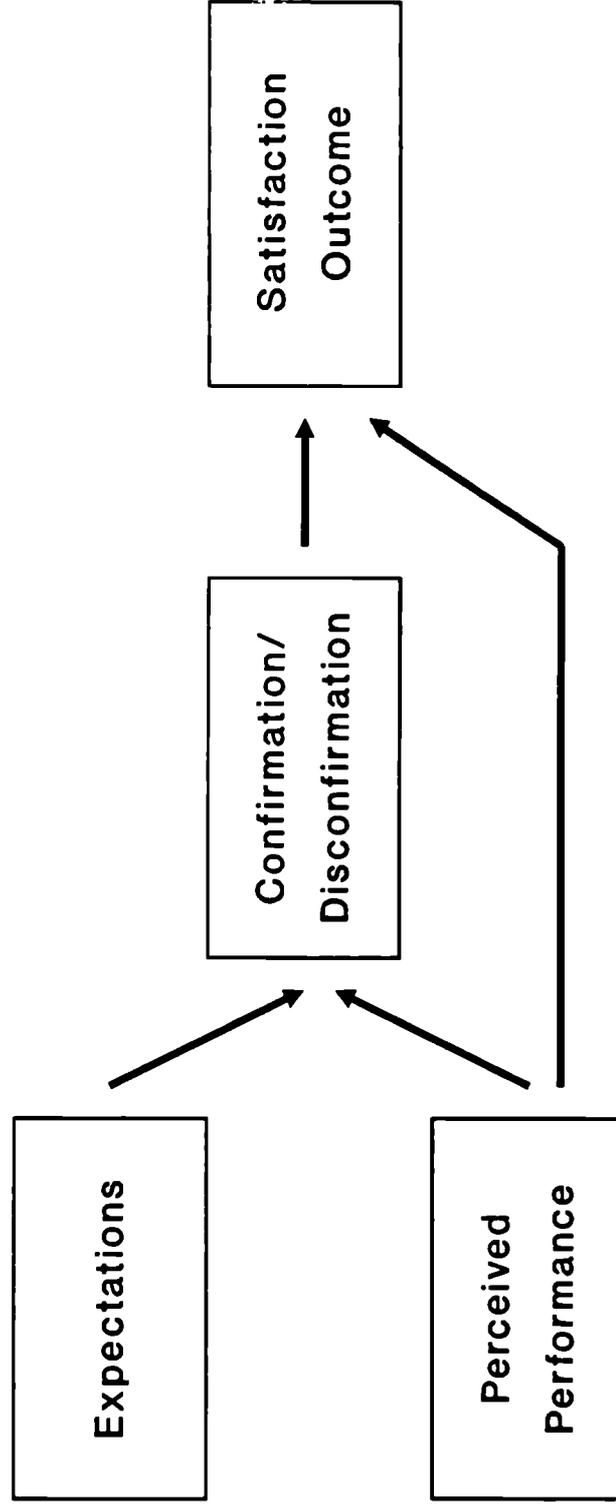
The previous subsection discussed comparison standards as the first basic component of disconfirmation models. This subsection presents perceived performance as the second model component.

Perceived performance is customers' subjective perception of actual product performance during the consumption experience. Perceived rather than objective performance has been primarily employed as a driver of the disconfirmation process (e.g. Bitner 1987, Cadotte et al 1987, Churchill and Surprenant 1982, Oliver 1980a).

Recent empirical findings seem to confirm the notion that, apart from driving the intervening disconfirmation variable, perceived performance directly influences satisfaction (Figure 2-3). It has been empirically shown that direct causal links from perceived performance to satisfaction increase significantly the proportion of explained variance in satisfaction. In fact, in some experiments perceived performance was a better predictor of satisfaction than the disconfirmation-of-expectations variable (e.g. Churchill and Surprenant 1982, Tse and Wilton 1988). Churchill and Surprenant conducted an experiment with video-disk-players. Perceived performance on its own accounted for 88% of the variance in satisfaction. Similarly, perceived performance explained 65% of the variance in Tse and Wilton's study. In both studies, the disconfirmation-of-expectations variable accounted for only a relatively small part of the variances in satisfaction, i.e. 50% in Tse and Wilton's study and no significant effect was

reported in Churchill and Surprenant's experiment using video disk players. These results seem to cast some doubt on the appropriateness of confirmation/disconfirmation as a mediating variable and/or expectations as the comparison standard.

**Figure 2-3: Perceived Performance as a Direct Determinant of Satisfaction?**



Sources: e.g. Churchill and Surprenant (1982)

Tse and Wilton (1988)

A possible explanation for these data might be a measurement problem, whereby the perceived performance and satisfaction measures applied captured completely or partially the same construct. In Churchill and Surprenant's study, some between-construct correlations (between perceived performance and satisfaction measures) were higher than within-construct correlations (between different satisfaction measures). This indicated a lack of discriminant validity between measures, and therefore, three out of five satisfaction measures had to be dropped (Churchill and Surprenant 1982). Other researchers encountered similar problems. Tse and Wilton (1988) obtained a correlation of .81 between satisfaction and perceived performance measures, which was the second highest correlation between any two measures in their study.

Cadotte, Woodruff and Jenkins (1987) indicated concern about the use of anchor words in performance measures that imply evaluation. Performance scales commonly used include anchor words such as good/bad, fast/slow and friendly/unfriendly. For instance, widely accepted standard measures employed in Churchill and Surprenant's (1982) paper used the following 7-point semantic differential scales anchored with "good sound - poor sound", "not very good - excellent", "very inferior - very superior" and "terrible - excellent." These scales clearly show strong evaluative components, and therefore probably measure components of the satisfaction construct.

Cadotte et al (1987) suggest that scales measuring true beliefs about the objective level of performance instead of evaluations ought to be used. Measures that cue

the respondent to an objective level of an attribute performance can be developed. For instance, the speed of a service can be measured in time units rather than fast/slow, and menu variety can be measured by the number of items on the menu instead of few/many. This approach is difficult when attributes do not have physical correlates. What objective measures can be used for service employee friendliness? On this issue, further measurement research is needed.

Independent of the line of argument of inappropriate measures, the notion of a direct causal link between performance and satisfaction can be rejected on logical grounds. A performance of X has no meaning unless we can position it in our value system. Performance can only have meaning in the context of evaluation, and evaluation implicitly requires some comparison standard. If, in the context of consumer satisfaction, the disconfirmation-of-expectation variable does not sufficiently explain the variance in satisfaction, a more suitable comparison standard ought to be used instead of introducing a direct causal link from perceived performance to satisfaction.

In summary, the generally employed perceived performance measures have strong evaluative components, which leads to the conclusion that these measures capture a part of the satisfaction construct. The high correlations between perceived performance and satisfaction, encountered in a number of studies are the result of inappropriate measures rather than an indication of a causal relationship between the two constructs. This proposition is tested in this thesis and the formal hypothesis is laid out in the conclusion section of this chapter.

### 1.1.3 Confirmation/Disconfirmation

The previous subsections discussed two components of disconfirmation models, namely pre-experience comparison standards and perceived performance. This subsection presents confirmation/disconfirmation as the third component.

Confirmation/disconfirmation occupies a central position as a crucial intervening variable in most satisfaction models. Disconfirmation arises from discrepancies between pre-consumption comparison standards and perceived performance.

#### 1.1.3.1 Theories About the Disconfirmation Process

Four common theories assist understanding of the disconfirmation process and enable predictions of its outcomes (Anderson 1973; LaTour and Peat 1978; Oliver 1977, 1980b). These theories are the contrast, consistency, assimilation-contrast and negativity theories.

Contrast theory predicts that if perceived performance is greater than or equal expectations, confirmation or positive disconfirmation<sup>1)</sup> occurs and this leads to satisfaction. If perceived performance is less than the expected performance,

---

<sup>1</sup> The terms "positive disconfirmation" and "negative disconfirmation" are frequently used phrases in the area of consumer satisfaction. The words "positive" and "negative" define the direction of disconfirmation. Positive disconfirmation means that the perceived performance was better than the comparison standard, and negative disconfirmation means that the perceived performance was worse than the standard.

negative disconfirmation occurs and this, in turn, causes dissatisfaction (Churchill and Surprenant 1982).

Consistency theory predicts that when disconfirmation occurs, consumers revise their perception of product performance to more closely match their pre-experience expectations (Anderson 1973, Olson and Dover 1979). These expectations-oriented responses might be triggered when individuals are either reluctant to acknowledge discrepancies from previously held positions (Oliver and DeSarbo 1988), or when the performance is ambiguous and therefore open to alternative interpretations (Hoch and Deighton 1989, Oliver 1980a). As a result, the consistency theory predicts that an increase in expectations also raises the level of satisfaction (Bitner 1987).

Assimilation-contrast theory (Sherif and Hovland 1961) reconciles the contradictory predictions of the consistency and contrast theories. The assimilation-contrast theory proposes that the magnitude of the perceived discrepancy determines, which of the two theories best explains the satisfaction outcomes of a particular consumption experience. According to this theory, if the discrepancy between expectation and perceived performance is small, consistency effects are predicted. Consumers will alter their perception of performance to match their pre-experience expectations. However, when the experienced disconfirmation exceeds a certain magnitude, contrast effects are predicted.

In other words, the assimilation-contrast theory qualifies the circumstances under which either consistency or contrast effects are predicted. Bitner (1987) suggests that there might be variables other than the magnitude of disconfirmation that influence what effects would be observed. For example, when performance is difficult to judge, as in the consumption of services, expectations might dominate and consistency effects would be observed. Other variables might include level of involvement or strength of expectations.

A fourth theory, negativity theory, predicts dissatisfaction when any disconfirmation, positive or negative, occurs (Anderson 1973, Oliver 1976). A study of the consequences of unconfirmed expectations was conducted by Carlsmith and Aronson (1963). They manipulated expectations about the taste (sweetness or bitterness) of solutions and then asked individuals to taste bitter and sweet solutions. The researchers assumed that bitter was an unpleasant taste and sweet was a pleasant taste. The results showed that any disconfirmation resulted in a significantly more bitter or less sweet judgement than the respective confirmation conditions. Carlsmith and Aronson concluded from these results that any disconfirmed expectancy results in a hedonically negative state of people.

Negativity theory has been empirically supported only when expectation were strongly held as in Carlsmith and Aronson's study (1963), or when subjects had high ego involvement, were committed to the outcome and were interested in the task (Oliver 1976). These results suggest that the negativity theory explains satisfaction decisions for only a rather limited set of consumption experiences.

### 1.1.3.2 Subtractive Versus Subjective Disconfirmation

Disconfirmation has been defined operationally in two different ways: as subtractive and as subjective disconfirmation. The subtractive disconfirmation approach assumes that the result of post-experience comparison can be expressed as a function of the algebraic difference between perceived performance and a comparison standard (e.g. LaTour and Peat 1979a; Parasuraman, Zeithaml and Berry 1985, 1988; Tse and Wilton 1988). On the other hand, subjective disconfirmation has been operationalized as the subjective evaluation of the difference between perceived performance and a comparison standard (e.g. Churchill and Surprenant 1982, Westbrook and Reilly 1983). In the course of this literature review three reasons were raised why subjective rather than subtractive disconfirmation should be used in modelling consumer satisfaction.

First, Oliver (1980b) suggests that subjective disconfirmation represents a distinct intervening cognitive state that results from the comparison process and precedes a satisfaction decision. Subjective disconfirmation accounts for a set of psychological processes that mediate performance perceptions, comparison standards and level of disconfirmation experienced. Examples of these processes are those that result in the contrast and consistency effects discussed in the previous section (section 1.1.3.1). Subtractive disconfirmation cannot account for these psychological processes.

Second, empirical comparisons of the two operationalizations indicate that subjective disconfirmation explains satisfaction outcomes far better than subtractive disconfirmation. In Tse and Wilton's (1992) study subjective disconfirmation explained 50% of the variance in satisfaction, whereas subtractive disconfirmation explained only 16.8%. Similar results were obtained in a study by Oliver and Bearden (1985). Of course, this superiority of subjective disconfirmation may simply be a reflection of the theoretical explanation provided in the previous paragraph.

Finally, a fundamental problem of any measure that builds on difference scores is lack of reliability. This conclusion has been virtually sanctified in the psychology literature to the point where difference scores are rarely used at all (e.g. Cronbach and Furby 1970). There are three factors contributing to the low reliability of difference scores. (1) The difference score unreliability increases as the unreliability of either score on which it is computed increases. (2) The difference score unreliability increases as the correlation between the two measures on which it is based increases. (3) As the two components in satisfaction research are measured at different points in time, difference score unreliability capitalizes on any effect of regression-to-the-mean over time. For a detailed explanation, refer to Prakash (1984) and Prakash and Lounsbury (1984). These researchers show empirically, using the disconfirmation-of-expectations model that difference scores have reliability far lower than subjective disconfirmation measures. As a result, the researchers recommend the use of subjective disconfirmation scales such as a five-point scale anchored with "much worse than expected" to "much better than expected".

#### 1.1.4 Satisfaction

The previous three subsections presented the following components of disconfirmation models: pre-experience comparison standard, perceived performance and confirmation/disconfirmation. This subsection discusses satisfaction as the final component.

It is generally accepted that satisfaction is determined by the magnitude and direction of the disconfirmation experienced (e.g. Bitner 1987, Swan and Trawick 1980, Tse and Wilton 1988). The satisfaction construct itself is defined as an evaluative response concerning the perceived outcome of a particular consumption experience (Westbrook and Oliver 1981).

Investigators of consumer satisfaction have frequently drawn on models from the job satisfaction theory. Herzberg's (1975) two-factor motivator-hygiene theory has been repeatedly proposed as being applicable to consumer satisfaction (Maddox 1981, Thomas 1988). This theory holds that satisfaction and dissatisfaction are two different constructs, caused by two different kinds of attributes of the consumption experience, namely hygiene and motivator attributes.

The theory suggested that negative disconfirmation along the hygiene attributes leads to dissatisfaction, and that positive disconfirmation of hygiene attribute performances will not cause satisfaction. A positive disconfirmation of motivator attributes is proposed to cause satisfaction outcomes. For example, if the durability

of a dress was a hygiene attribute and its design a motivator attribute, then durability would drive dissatisfaction, and design would drive satisfaction. This would mean that an improvement in design would not reduce dissatisfaction, and an improvement in durability would not increase satisfaction, which should be only driven by design improvements. Although the two-factor notion is an intriguing perspective with at least some intuitive appeal, empirical tests did not support it. Maddox (1976, 1981) focused on this issue in his dissertation and refuted the two-factor notion on the basis of two empirical studies with four different kinds of products.

In contrast to the two-factor perspective is the view that satisfaction and dissatisfaction are opposites on a single, bipolar continuum. This view has generally been accepted in the area of consumer satisfaction (e.g. Bitner 1987, Westbrook and Oliver 1991) and was further supported by Maddox's empirical work (1976, 1981). The one-factor perspective was also adopted in this thesis.

### 1.1.5 Operationalization of the Components in Multiattribute-Type Models

Four basis components of disconfirmation models were presented in the previous subsections. The operation of these components with multiattribute-type models is presented in this final subsection on disconfirmation models.

Multiattribute models have been used extensively to simulate evaluations of products and services in the decision making process (Bateson 1989, p.83-89; Sarel 1978) and to analyze the driving dimensions of overall quality perceptions (e.g Carman 1990; Parasuraman, Zeithaml and Berry 1985, 1988). Satisfaction researchers transferred the concept of multiattribute models to the evaluation of individual consumption experiences. The application of these models in the context of consumer satisfaction has received much empirical support (e.g. Churchill and Surprenant 1982, Day 1977, LaTour and Peat 1978, Woodruff et al 1981, 1983) and is the prevailing conceptualization of the four components of satisfaction models, i.e. the comparison standard (usually expectations), perceived performance, disconfirmation and satisfaction. In these models, a global satisfaction concept is operationalized as a combination of levels of satisfaction from various dimensions or attributes of a particular product. Commonly, a Fishbein-type linear compensatory model is used, which can be expressed as follows:

$$S_{jk} = \sum w_{ik} AS_{ijk},$$

where:  $i$  = attribute,  $j$  = consumption experience, and  $k$  = consumer such that:

$S_{jk}$  = consumer  $k$ 's overall satisfaction with the consumption experience  $j$ ,

$w_{ik}$  = the importance weight to given attribute  $i$  by consumer  $k$ , and

$AS_{ijk}$  = consumer  $k$ 's satisfaction with attribute  $i$  at the consumption  $j$ .

As marketers can manage their customers' expectations and perceived performance directly, and therefore disconfirmation and satisfaction indirectly, an analytical tool is needed to provide guidelines for management. In order to obtain an understanding of satisfaction outcomes with specific products, and of the impact of any actions, it is necessary to pinpoint the causes of satisfaction outcomes. For this reason, the levels of satisfaction with each salient attribute of a product is of particular interest.

This approach improves marketers' ability to analyze and manipulate their consumers' satisfaction. For example, low levels of perceived performance in attribute A, and too high expectations in attribute B can both cause disconfirmation. By identifying these causes managerial actions can be taken that increase the performance perception in A and lower the expectations in B. Furthermore, if one can estimate the multiattribute function well enough, one should be able to predict how a consumer's disconfirmation and satisfaction will change for given alterations in one or more product attributes (e.g. Shocker and Srinivasan 1979, Wilkie and Pessemier 1973).

## 1.2 Equity Theory

Disconfirmation models, which depict the main stream in satisfaction theory, have been discussed in section 1.1. An alternative to these models, equity theory, holds that consumers have a notion of fairness (e.g. Swan and Mercer 1982) and expect a form of distributive justice in exchange situations such as sales transactions (Cook and Messick 1983). Oliver and Swan (1989) propose that the fairness dimension reflects inputs and outputs of both customers and sellers. Messick and Sentis (1979) suggested the "weak proportional equity formula," which requires that the inputs and outcomes of both parties approximate one another. Similarly, Swan and Mercer (1982) propose that equity involves primarily the relationships between the marketer's and the consumer's net gain, where the two net gains have to be in balance.

Applied to consumer satisfaction, equity theory predicts that consumers are more satisfied with equitable exchanges than with inequitable ones. Empirical research supports this notion. Swan and Oliver (1985) found that equity and disconfirmation of expectations each contributed significantly and independently to satisfaction. Equity was measured on a five-point scale from negative inequity (the sales person came out ahead) to positive inequity (I came out ahead), and with equity at the midpoint (we both benefited equally).

Oliver and Swan (1989) also found strong empirical support for equity theory using data from a survey of new car buyers. The authors operationalized equity in the

form of buyer and seller inputs and outputs. Seven measures for each of the input and outcome variables were obtained. Four measures were for specific inputs and outcomes determined through pretesting. Buyer outcomes were: a good explanation about the car, a good deal (price), personal attention and an understanding of personal needs. Seller outcomes were: a high sales commission, getting ahead on the job, a satisfied customer and buyer referrals. The remaining three were general statements like: "The salesperson put a lot into dealing with me." Each was measured on a 7-point Likert scale. Oliver and Swan (1989) concluded that equity significantly influences satisfaction, and that it is a mediating variable between input, outcomes and satisfaction.

Equity theory may be criticized in a number of ways. Equity perceptions appeared to be largely driven by buyer outcomes, suggesting a self-centred and asymmetric buyer perspective (Oliver and Swan 1989). Buyer outcomes, rather than equity become the determinant of satisfaction. Furthermore, whether satisfaction could be captured equally well by disconfirmation models (e.g. by expressing buyer outcomes in the form "the explanation of the car was: much better than expected" to "much worse than expected") remains to be tested. Finally, the operationalization of satisfaction as satisfaction with a sales-person in a bargaining situation (car-purchase) is not typical for consumption experiences (Bitner 1987, p. 18-19). Therefore Oliver and Swan's findings may not be generalizable across other consumption experiences.

In conclusion, conceptual work and empirical evidence support the consideration of equity theory in the study of consumer satisfaction with exchange experiences. However, further research is clearly needed. First, investigations are required to establish whether the equity approach actually adds something to the disconfirmation model. Second, work is required to generalize results to situations where the relationships between buyer and seller inputs and outcomes are less clear than in bargaining situations (Swan and Oliver 1985, Oliver and Swan 1989).

### **1.3 Empirical Comparisons of Alternative Satisfaction Models**

In the first two sections of this chapter, disconfirmation models with alternative comparison standards and equity theory based models were reviewed. This existence of alternative models of consumer satisfaction has induced a number of researchers to conduct studies that compare the explanatory powers of these models. These studies are reviewed in this section.

Tse and Wilton (1988) manipulated performance expectations of a portable CD-player using written evaluations by an "independent consumer testing laboratory." Performance was manipulated using two versions of the product: an advanced model (representing good performance) and an earlier model which was modified to guarantee poor performance. The experimenters measured the following pre-experience constructs: ideal, equitable and expected performance. Ideal and expected product performance were measured for both the overall product and for

26 individual attributes. After having experienced the good or bad performance condition, subjects then rated perceived performance, subjective disconfirmation-of-expectations and satisfaction.

An equity model, with equity operationalized as perceived performance minus equitable performance, was the worst performing of the three models. It explained 38% of the variance in satisfaction. Furthermore, the equitable performance standard was not even significantly correlated with satisfaction.

The classical disconfirmation-of-expectations model, with subjective disconfirmation as the mediating variable between expectations, perceived performance and satisfaction, explained 50% of the variance in satisfaction.

The ideal performance standard was operationalized through a five-point bipolar scale (has exactly the combination of attributes you would like to see in a miniature record player...). This model explained 56% of the variance in satisfaction. Surprisingly, it seemed to have performed slightly better than the classical disconfirmation-of-expectations model.

However, there seems to be a problem with the comparison between the expectations and ideal standard models. Subjective disconfirmation was used in the expectations model and subtractive disconfirmation was used in the ideal standard model. Now, the ideal standard was not manipulated in the experiment and so a high correlation between subtractive disconfirmation and satisfaction

( $r=.75$ ) reflects the linear relationship between perceived performance and disconfirmation ( $r=.89$ ). But this should not be unexpected. The performance measures included a strong evaluative component (e.g. a scale ranging from "very good" to "very poor performance")<sup>2</sup> and hence there was a high correlation between perceived performance and satisfaction ( $r=.88$ ). In summary, the explanatory power of the disconfirmation-of-expectations model was actually compared with the explanatory power of a model in which the perceived performance measures already captured a large part of the satisfaction construct.

In another recent empirical study, which was briefly reviewed in section 1.2, Oliver and Swan (1989) compared the equity and disconfirmation-of-expectations models. The researchers operationalized disconfirmation with a three item "better than expected/worse than expected" scale. Oliver and Swan used structural equation modelling with Lisrel (Joereskog and Soerbom 1984) for the data analysis. The size of causal path coefficients showed that equity (standardized loading .680) had a greater effect on satisfaction than disconfirmation (standardized loading .344). The causal path between disconfirmation and satisfaction was still highly significant. The survey was conducted in an interpersonal context (car buyers' satisfaction with sales personnel). Oliver and Swan (1989) suggest that fairness is a more important determinant of satisfaction in interpersonal exchange situations than disconfirmation. However, these findings have not been generalized to non-bargaining situations and to satisfaction with overall consumption experiences.

---

<sup>2</sup> Refer to section 1.1.2 for a discussion of the problems inherent in using perceived performance measures measuring evaluations rather than perceptions.

Oliver and DeSarbo (1988) obtained results in their study of satisfaction with investment decisions that strongly supported the disconfirmation-of-expectations model. Disconfirmation-of-expectations (operationalized as subtractive disconfirmation) accounted for 32.3% of the variance in satisfaction. The direct effects of performance and expectations explained 4.6% and 2.6%, respectively, of the total variance, and equity (which was operationalized as the ratio of customers' transaction profit and brokers' commission) accounted for merely 1.2%.

Also, Oliver and DeSarbo (1988) encountered differences across individuals. Most subjects in their study responded to the disconfirmation-of-expectations treatment. However, the disconfirmation respondents could be separated into two groups, those who fit the classical disconfirmation model (50% of all subjects) and those who use disconfirmation with performance and, to a lesser extent, expectations as a direct determinant of satisfaction (37.5% of all subjects). The remaining 12.5% of all subjects seemed to be equity oriented. Their beta equity beta coefficients were of similar size to their coefficients for the disconfirmation variable, suggesting that equity on its own would not have sufficiently explained the variance in satisfaction even for equity oriented subjects. Unfortunately, determinants (such as demographic variables) of these response tendencies could not be found.

In spite of the differences between subjects, it should be noted that disconfirmation was for all groups of subjects an important and highly significant determinant of satisfaction, although for about half the subjects other significant determinants were found.

In summary, disconfirmation-of-expectations was shown to be an important determinant of satisfaction in all three studies. It seems that this model has the largest external validity of all models included in these empirical comparisons. The evidence for the equity theory is varied. At the least, the evidence seems to indicate that the equity theory is only applicable to a limited set of consumption experiences, e.g. of personal interchanges as in bargaining situations. However, no researcher has tried to determine the circumstances under which equity is an important determinant of satisfaction.

#### 1.4 Summary

All satisfaction models are based on a comparison process of one sort or another. The majority of these models are based on a comparison between perceived performance and a pre-consumption experience standard. Of all satisfaction models, the disconfirmation-of-expectations model seems to be the best available conceptualization of the satisfaction process. It has been the most widely applied and has received a lot of empirical support. Theorists generally agree that satisfaction can be seen as the consumer's response to the evaluation of the difference between expectations and perceived performance. Furthermore, the simplicity of this model is intuitively appealing to both researchers and practitioners.

The existence of alternative standards seems to indicate that there are limitations to the disconfirmation-of-expectations model. In particular, the model disregards

needs and wants, and the influence of non-focal brand experiences on satisfaction. However, under the assumption that consumers have free choice and sufficient pre-purchase information, and that they choose purposefully to match their needs and wants with expectations, the disconfirmation-of-expectations model seems to be a good conceptualization of the satisfaction process.

The review discussed problems with perceived performance measures that capture evaluations instead of beliefs about the performance level. A formal hypothesis on this issue is developed in the last section of this chapter. Four theories on the disconfirmation process, contrast, consistency, contrast-assimilation and the negativity theories, were reviewed. It was proposed that disconfirmation should be operationalized with subjective rather than subtractive disconfirmation and a number of reasons were provided. Furthermore, the literature reviewed suggests that satisfaction is a unidimensional construct with satisfaction and dissatisfaction on opposite sides on a single, bipolar continuum. The two factor theory of satisfaction was rejected.

## **2 Services Specific Features in Disconfirmation Models**

Research conducted in the context of the disconfirmation paradigm has, in general, neither conceptually nor empirically differentiated between products and services. Research was conducted with products and services alike without addressing the idiosyncratic features of the service encounter and the possible impacts of these on the satisfaction process. However, features of services (compare Table 1-1) may have considerable impact on the satisfaction process and, therefore, open up many interesting research issues. One of them, the experimental nature of services, was already introduced in Chapter I. Four further issues are outlined in the following subsections. In particular, it is suggested that (1) a high degree of ambiguous and experience attributes (Bateson 1977, LaTour and Peat 1978, Mills and Moberg 1982) can cause large assimilation effects between expectations and perceived performances, (2 and 3) face-to-face service employee and consumer interactions make it possible to manage consumer expectations (Bateson 1989, Gelb and Smith 1988) and attribution processes (Bitner 1987, 1990), and (4) a high degree of service performance heterogeneity (Berry 1980, Krughoff 1981, Zeithaml 1981) might have impact on the level of disconfirmation perceived during a consumption experience.

## 2.1 Ambiguous Attributes Foster Consistency Effects

Parasuraman et al (1985) drew on the economics literature (Nelson 1974, Darby and Karni 1973) and proposed three different service attributes: (1) search attributes, which can be evaluated before purchase, (2) experience attributes, which can be evaluated during and after the service encounter, and (3) credence attributes, which cannot be evaluated confidently even after the service has been consumed. Most services contain only few search attributes (e.g. price level and location of a restaurant). A large number of the service attributes can only be evaluated after the consumption (e.g. the taste of a meal). Some attributes cannot be evaluated even after having consumed the service (e.g. the hygiene conditions in the kitchen).

Hoch and Deighton (1989) and Hoch and Ha (1986) criticized Nelson's (1974), and Darbi and Karni's (1973) classification and, in particular, the assumption that consumers know that neither search nor experience can help them evaluate credence attributes. According to their own research, this does not seem to be the case. Hoch and Deighton (1989), therefore, proposed ambiguous attributes as a further category next to search, experience and credence attributes. They define ambiguity in the sense that the evidence is consistent with more than one hypothesis. Ambiguous attributes are perceived and evaluated by the consumer (nothing is missing in contrast to credence attributes), but they can be perceived in more than one way. For example, a messy desk in a travel agency could be interpreted by the consumer with regard to the perception of the service quality as: not relevant

information, an exceptional occurrence, or a direct indicator of the service quality provided. If consumers fail to recognize the full range of possible interpretations, the evidence (the messy desk) will be seen as more diagnostic than it is (Hoch and Deighton 1989).

Furthermore, when the evidence is ambiguous, people tend to encode information as consistent with prior knowledge and expectations (e.g. Herr, Sherman and Fazio 1982). In other words, when evidence is ambiguous, consumers are likely to experience what they expect to experience (Hoch and Deighton 1989, LaTour and Peat 1978). When applied to the satisfaction process, this means that consistency effects<sup>3</sup> are likely to occur when ambiguous attributes are encountered.

Services are physically intangible (they cannot be touched) and mentally intangible (it is frequently difficult to envision what has been obtained after having received a service) (Bateson 1977, Levitt 1981), and researchers in the field have argued that it is difficult for consumers to evaluate the quality of a service (Parasuraman et al 1985). These characteristics lead consumers to rely upon cues associated with the service production process, the service worker and the service environment when evaluating service quality and making satisfaction decisions (Mills and Moberg 1982). These cues may be prone to subjective evaluation because of their ambiguity and, therefore, consistency effects may be important in service satisfaction processes.

---

<sup>3</sup> Consistency theory predicts that when disconfirmation occurs, consumers will revise their perception of the service performance to more closely match their pre-experience expectations (Olson and Dover 1979). For a review of the consistency and assimilation-contrast theories refer to Chapter II, 1.1.3.

In summary, it is suggested that services are prone to triggering consistency effects because of their high degree of ambiguous attributes. Consequently, expectations that are higher than the objective performance level can increase the assessment of that performance and, in turn, can increase satisfaction. It follows then that expectations that are lower than the objective performance level can decrease the level of perceived performance and satisfaction.

## **2.2 Expectations Can be Managed in the Service Encounter**

In contrast to products, it is possible to manage the expectations of individual consumers during the service consumption process (e.g. a bank employee's investment advice can correct risk and return expectations of clients). From indepth interviews with professional services providers, Gelb and Smith (1988) developed a service satisfaction model using disconfirmation-of-expectations and incorporating Parasuraman's et al (1985) conceptualization of services quality. One of their major findings is that buyer expectations, shaped by pre-purchase influences, can be verified or corrected during the service encounter, and these expectations then constitute the comparison standard in the disconfirmation process. This can have important consequences on satisfaction outcomes as unrealistically high expectations can be corrected and thereafter met by the service provider.

Expectations can also be raised during the service encounter. These real-time expectations can be created explicitly as a promise ("your food will be ready in five

minutes") or implicitly by a behaviour pattern that sets a tone (Bateson 1989). Such expectations are often created unwillingly, for example, when a waitress promises to "be right back." This statement can be viewed as a farewell salutation by the waitress and as a binding promise by the diner.

In summary, expectations management in the service encounter can prevent consumer dissatisfaction by lowering unrealistically high expectations and by avoiding their creation.

### **2.3 Attribution Processes Can be Managed in the Service Encounter**

Bitner (1987, 1990) applied attribution theory developed in psychology to the service encounter. In this context, attributions are consumer perceptions of the causes of the outcome of a service encounter. In a major review on attribution theory, Weiner (1985) concluded that people engage in spontaneous causal thinking, particularly in cases of unexpected and negative events. It is generally accepted that most causes can be classified along three dimensions: locus (who is responsible?), control (did the responsible party have control over the cause?) and stability (is the cause likely to reoccur?).

Bitner suggested that consumers diagnose the causes of negative disconfirmation and that the perceived nature of these causes determines the level of satisfaction. She designed a study to test this hypothesis in which control and stability

attributions were manipulated in the context of a service experience with a travel agency. Her hypothesis was confirmed. The results suggest that when consumers perceive the cause of a service failure to be within the control of the service provider (control attribution) and likely to reoccur (stability attribution), they will be more dissatisfied than when the opposite conditions hold (Bitner 1990).

A further important finding of Bitner's work is that, in contrast to product marketers, service providers can actively manage customer attributions. Bitner used in her experiment non-verbal cues (the physical appearance of the service encounter) and service employee behaviour (offer of an internal or external explanation for the service failure) to manipulate consumer attributions. The physical appearance of the service encounter was used to manipulate the stability attributions. When the service setting looked "disorganized" a service failure was perceived to be more likely to reoccur, and when it looked "organized" it was perceived as less likely to reoccur. Verbal explanations given by service employees were used to manipulate control attributions. When the service employee admitted that the service failure was her fault, subjects felt that the failure was within the control of the firm, and when the employee offered an external explanation, the subjects thought the failure was beyond the control of the firm.

As these manipulations proved highly successful, Bitner (1987) concluded that the interactive nature of service encounters makes it possible for service firms to manage the attribution processes of their customers. This, in turn, can be used to reduce dissatisfaction when service failures happen (Bitner 1987, 1990).

## 2.4 Service Performances are Heterogeneous

In contrast to products, services show a high degree of unit-to-unit variability because of their real-time production and consumption (Berry 1980, Zeithaml 1981), and the human factor involved in service production (Krughoff 1981). Real-time production and consumption make quality control difficult (e.g. bad lots cannot be rejected, DeSouza 1989) and introduce time and place constraints on the service encounter (Bateson 1989, p. 8), which typically result in heterogeneous performances. Furthermore, service quality relies heavily on the technical and social capabilities of service employees, and the performance of service workers can vary considerably from one employee to another within the same service organization. Even for the same individual, the performance can fluctuate significantly over time (Krughoff 1981).

Solomon and Surprenant (1985) applied role theory to the service encounter and suggest that the congruity between employee and consumer role behaviour is an important determinant of the service production process. As the role behaviour varies according to expectations and capabilities (Solomon and Surprenant 1985), one can infer that not only employee but also consumer behaviour contributes to the heterogeneity of services.

So far, satisfaction researchers have implicitly assumed point-expectations. However, if consumers are familiar with the high variance in service performances, one might argue that they expect some kind of band within which the performance usually

occurs. That is, consumers may expect performance level distributions with the parameters of the distributions varying between service classes, service providers and even service employees of the same service organization. Extensive research has been conducted to explore the impact of variance and uncertainty on consumer choice behaviour (e.g. Diamond 1988, Meyer 1981), but not on satisfaction outcomes. This is in spite of the reasonable assumption that it might have considerable impact on disconfirmation whether a consumer holds a point-expectation (e.g. he expects waiting time for a meal of 15 minutes in a particular restaurant) or an expectation with a wide variance (e.g. a waiting time between 5 and 25 minutes). What are the disconfirmation and satisfaction outcomes in our example, if the consumer has to wait for 10, 20 or 40 minutes? Research that examines the impact on the satisfaction process of consumers holding performance level distributions is clearly needed. As this seemed to be an important and interesting issue in services marketing, it was selected as one of the focal points of the thesis.

## **2.5 Summary**

Idiosyncratic features of services should not be disregarded by satisfaction researchers. An understanding of those features can give further insight into satisfaction processes and can provide guidelines for the management of service organizations. For example, this review showed that (1) ambiguous attributes foster consistency effects, whereby increases in expectations result in increased

satisfaction outcomes, (2) expectations can be managed during the service encounter and dissatisfaction outcomes can be avoided, and (3) consumer attributions can be directed and dissatisfaction outcomes reduced. No research has been conducted into performance heterogeneity in the context of consumer satisfaction. This issue is to be addressed in the thesis.

### 3 Conclusions, Further Research and Development of Hypothesis I

The disconfirmation-of-expectations model seems to be the best available conceptualization of the satisfaction process. It has received extensive empirical support and has been generalized across a multitude of products and services. This thesis aims at further extending the model by the inclusion of two services specific features namely heterogeneous performance expectations and the affective state of the consumer during the consumption process. A secondary objective of this thesis is to clarify the role of perceived performance in the satisfaction process.

#### The Role of Perceived Performance in the Satisfaction Process - Hypothesis I

It has been repeatedly suggested and empirically supported that there is a direct causal link between perceived performance and satisfaction (refer to section 1.1.2 of this chapter). In fact, in some experiments, perceived performance was a better predictor of satisfaction than disconfirmation. Close examination of the measures employed in these studies revealed that perceived performance measures were used that capture strong evaluative components and may well be the underlying cause of the observed correlations between perceived performance and satisfaction. Furthermore, it was suggested that these high correlations would not be encountered if true beliefs about the objective level of performance were measured instead of evaluations. As a test of this proposition could be conveniently conducted without

obscuring the two main objectives of the thesis, the following hypothesis was put forward:

**Hypothesis I:**

**Direct causal links between perceived performance and satisfaction do not exist; high positive correlations between the two constructs are merely an indication of inappropriate perceived performance measures.**

**Heterogeneous Performance Expectations in Satisfaction Models - Further Research**

The review in this chapter found that disconfirmation models always implicitly assume point-expectations. There is, however, reason to believe that consumers are familiar with the high heterogeneity in service performances and therefore expect some kind of performance level distribution. As this seemed to be an important and interesting issue in services marketing, it was selected as one focus of this thesis and was examined for the first time in the context of consumer satisfaction. In Chapter III, the effects of performance heterogeneity on consumer behaviour and operations management are reviewed, and two hypotheses about the impact of heterogeneous performance expectations on the satisfaction process are developed.

## The Role of Affect in the Satisfaction Process - Further Research

In Chapter I it was concluded that disconfirmation and multiattribute models capture only the cognitive dimensions of the service encounter. These models disregard its experiential nature. The environmental psychology and perceived control perspectives, on the other hand, try to capture affective responses that are intrinsic to the service experience. In both theories, affect plays a central role as a mediating variable between stimuli and behaviour.

In the context of consumer satisfaction, very little effort has been devoted to modelling affective responses during a consumption experience (Westbrook 1982, 1987), or to explaining the link between cognitive processes and affect (Woodruff et al 1983). Nevertheless, explicit consideration of affective responses and the link between cognition and affect seem particularly promising. The classical satisfaction model, which appears to be good at capturing the cognitive dimensions of a consumption process, might be extended by including affect to capture the way people feel while consuming a service. The question concerning satisfaction research is whether affect can and should be incorporated in the currently employed satisfaction models to better capture the experiential nature of services.

Psychological research into the concept of affect and its causes is reviewed in Chapter IV, and four hypotheses on the inclusion of affect in the disconfirmation-of-expectations model are developed in Chapter V.

## Chapter III

# The Impact of Heterogeneous Service Performances on Consumer Satisfaction

In Chapter II, the potential impacts of four typical features of services on satisfaction were discussed. One of these features, namely heterogeneous performances of services, has been examined neither conceptually nor empirically by satisfaction researchers. However, there is reason to believe that consumers are familiar with the high degree of heterogeneity in service performances and, therefore, also hold heterogeneous performance expectations. This chapter tries to examine the impact of heterogeneous performance expectations on disconfirmation models.

In the first section, some operations management issues are addressed. In operations management the heterogeneity issue has been examined from an operations efficiency perspective with direct implications for consumer choice. Section two reviews the literature on performance heterogeneity in the context of the consumer choice process. In particular, it reviews work on perceived risk and multiattribute modelling. In section three, the decision science literature on decision making under uncertainty is reviewed and a number of conclusions are drawn. An important conclusion for this thesis is that neither the work in

consumer behaviour nor in decision science has addressed the post-decision stage which could help to develop theories on the impact of heterogeneous performance expectations on satisfaction. In the last section, two hypotheses about the impact of heterogeneous performance expectations on consumer satisfaction are developed on the basis of the disconfirmation-of-expectations model.

## **1 Why is Heterogeneity Important - An Operations Management Perspective**

Operations management and marketing theorists have dealt with the heterogeneity issue from the perspective of operations efficiency. They have concluded that workflow and task uncertainty introduced by customers decrease potential operations efficiency (Slocum and Sims 1980) by limiting the scope for industrialization of services via application of systems and machines, and by limiting opportunities for specialization and division of labour (Levitt 1972, 1976).

Several methods for reducing the workflow and task uncertainty inherent in service encounters have been proposed. Chase (1978, 1981) suggests that customer contact time should be minimized to reduce customer induced uncertainty. Mills and Moberg (1982) propose client selection and socialization, as well as routinization of the service encounter as means of reducing task uncertainty. These methods of uncertainty reduction result in simplified service encounters, fewer customer/service employee interactions and increased efficiency.

It has therefore been concluded that there is a negative correlation between possible operations efficiency and the degree of customization offered (Chase 1981; Mills and Moberg 1982). Service providers have to trade-off between the two (Surprenant and Solomon 1987).

Mills and Moberg (1982) showed the extreme points of this trade-off by developing two models: a full-service model and a restricted service model. The full-service model is characterized through a minimally restricted and routinized service encounter. Customers have a maximum number of choice and interaction possibilities. Service employees' expertise and decision making discretion are at very high levels. The control of the service employees is collegial and primarily related to the possession of skills rather than to the actual production process. Operations efficiency is limited by the high level of unpredictability (customer and employee behaviour) in the service encounter.

In the restricted service model the encounter is highly routinized and sealed off as much as possible from client-created uncertainties. Customer behaviour is standardized and service choice is very limited. Rules and regulations for employees and process control are enforced by management. Employees have very little discretion concerning decisions and behaviour. High levels of operations efficiency can be achieved as uncertainties are minimized.

Service organizations have to trade-off the costs and benefits of full-service against those of restricted-service as suggested by Mills and Moberg (1982). This requires

an understanding of the impact of standardization on the consumer choice and satisfaction processes. Standardization involves lower degrees of freedom in behaviour for both service employees and consumers in a service production process (Shostack 1987), which in turn leads to a lower performance heterogeneity (Levitt 1976). The impact of performance heterogeneity on consumer choice behaviour has been examined extensively (e.g. Diamond 1988, Meyer 1981, Sarel 1978). This research is reviewed in the next section.

## 2 Performance Heterogeneity in the Choice Process - A Consumer Behaviour Perspective

In the consumer behaviour literature, the service purchasing process is frequently divided into two complex stages: (1) the prepurchase choice, and (2) the consumption and postpurchase evaluation stage (Bateson 1989, p. 80-83; Fisk 1981). The prepurchase stage is basically a decision making process, whereby consumers evaluate available alternatives. The outcome of the prepurchase stage is a decision to buy the best option available in the eyes of the individual. The second stage encompasses the interactions between customer and service firm (this includes production, acquisition and consumption of services) and the service evaluation process. Its outcome is a satisfaction evaluation by the consumer.

Perceived risk plays an important role in theories of pre-purchase choice. Consumers perceive risk when they make purchase decisions and employ strategies to reduce this risk. The first subsection of this section reviews the impact of performance heterogeneity on perceived risk. Generally, multiattribute-type models have been used to describe consumer decision making processes. To increase the normative scope of perceived risk theory, researchers have incorporated risk at the attribute level (i.e. uncertain attribute performances) rather than at the overall product/service level. The second subsection focuses on consumer evaluate of alternatives with uncertain attribute levels, i.e. of heterogeneous performances.

Two main conclusions are presented in the third subsection. First, according to both the perceived risk and multiattribute literature, performance heterogeneity can have significant impact upon consumers' decision making. Second, neither of the two bodies of literature has addressed the post-purchase stage or provided a framework that would allow to develop theories on the impact of heterogeneous performance expectations on consumer satisfaction.

## **2.1 The Impact of Performance Heterogeneity on Perceived Risk**

Consumers have to make decisions regarding what services to buy, whom to buy from and when to buy. As the outcomes of such decisions are often uncertain, consumers face some degree of risk in making them. The concept of perceived risk has become well established in the marketing literature as one of the key determinants of consumer behaviour. Perceived risk is defined as the uncertainty that consumers face, when they cannot foresee the consequences of a purchase decision. Working from this definition, Bauer (1960) identified two dimensions of perceived risk: (1) consequence - the degree of importance and/or the degree of danger inherent in the outcome of any consumer decision, and (2) uncertainty - the subjective probability of an outcomes.

Other researchers have conceptualized perceived risk as multidimensional. Jacoby and Kaplan (1972) suggest that the major types of risk consumers perceive include performance, social, financial, physical and psychological risks. Furthermore,

Roselius (1971) suggests that time and hazard are also components of perceived risk.

Recently Diamond (1988) conducted a study to test the relative impact of Bauer's (1960) two dimensions of perceived risk which were operationalized as (1) the magnitude of possible negative consequences and (2) the probability of their occurrence. Diamond built on Kahneman and Tversky's (1979) prospect theory. This theory posits that people combine both probability and consequence perceptions to judge uncertain outcomes. Judgements are described by a value function of the consequences multiplied by a weighting function of the probabilities. Diamond conducted an experiment with imaginary car purchases whereby two variables were manipulated: the magnitude of a potential negative consequence (the amount of a possible repair bill), and the probability of that consequence (the repair being necessary).

The results of this study show that in low-probability/high consequence situations, the subjects were affected more by the magnitudes of the consequences than by their probabilities, which means that individuals may uniformly disregard very low probability outcomes (Diamond 1988). In the high-probability/low consequence situations, the subjects were affected more by probabilities rather than consequences. For example, subjects preferred a car with a 30% chance of a \$350 repair bill to a car with a 60% chance of a \$175 bill (the expected bill is \$105 in both cases). Diamond's research has provided further insight into the relative importance of Bauer's two dimensions of perceived risk. In particular, he draws

attention to the inappropriate, although often implicitly held assumption that perceived risk is the product of uncertainty and consequence.

Other determinants of perceived risk have been investigated. For instance, it has been shown that perception of risk varies between people (high-risk versus low-risk perceivers; Currim and Sarin 1983), the purchasing situation (e.g. non-store retailers versus store retailers; Korgaonkar 1982), the culture (e.g. risk is a less important determinant of consumer behaviour in Mexico than in the United States; Hoover, Green and Saegert 1978) and the product category (products can be classified on the basis of type and intensity of perceived risk; Cunningham 1967).

The perceived risk literature shows that consumers are risk averse and, therefore, strive to reduce the level of perceived risk in the decision-making process. The higher the perceived risk, the greater they strive to (1) increase the predictability of the various purchase decision outcomes, (2) limit the magnitude of the possible negative outcomes, and (3) decrease the probability of occurrence of those outcomes. For example, the magnitude of potential consequences can be reduced by buying smaller quantities or lowering aspiration levels. Some of the strategies applied by consumers reduce uncertainty are increased information seeking, brand loyalty, selection of well-known brand names, reliance on store image and seeking reassurance (Schiffman and Kanuk 1987, p. 214-20).

### Services as High Risk Products

Services have often been considered as being perceived by consumers as more risky than products (Guseman 1981, Zeithaml 1981). There are a number of reasons for this view.<sup>1)</sup> First, services show a higher unit-to-unit variability than products due to difficulties in quality control (DeSouza 1989), customer involvement in the production process (Zeithaml 1981) and the variance of the technical and social capabilities of service employees (Krughoff 1981)<sup>2)</sup>. Second, it is very difficult to evaluate service quality before purchase because many aspects of quality can be evaluated only after consumption (experience attributes) or may not be amenable to evaluation at all (credence attributes) (Parasuraman et al 1985, Zeithaml 1981)<sup>3)</sup>. Finally, consumer participation in the service production process implies that the cost of a service not only encompasses the price but may also include time as well as mental and psychological effort. A wrong decision may result in a loss of these inputs, and hence consumers perceive a higher level of risk (Hui 1988, p. 36-7).

---

<sup>1</sup> For a detailed review refer to Hui (1988), p. 36-7.

<sup>2</sup> Refer to Chapter II, Section 2.4 for a more detailed review of the causes of performance heterogeneity of services.

<sup>3</sup> For a review on search, experience, credence and ambiguous attributes refer to Chapter II, Section 2.1.

## 2.2 Performance Heterogeneity Included in Multiattribute Models

Multiattribute models have been used extensively to simulate evaluations of products and services in the decision making process (Bateson 1989, p. 83-9; Sarel 1978). According to these models, consumers evaluate a number of salient attributes of each alternative and weight them in relation to their importance. The final choice is made on the basis of a combination of these weighted attribute scores. A number of methods of combination have been proposed. The most commonly applied is the linear compensatory model, in which the weighted attribute scores are simply added and the alternative with the highest overall score is chosen.

To increase the normative scope of perceived risk theory, researchers have incorporated risk at the attribute level rather than looking at perceived risk of an overall product/service level. A review of these multiattribute models with risky attributes is provided in this section.

Performance heterogeneity in multiattribute models is basically concerned with (1) uncertain attribute performances and (2) the associated consequences for each potential outcome. Meyer (1981) conducted a laboratory experiment that examined the impact of uncertain attribute performances on attribute evaluation and choice behaviour. Sarel (1978) examined the impact of the associated consequences of potential outcomes on choice behaviour. Meyer's and Sarel's experiments are reviewed in the following paragraphs.

Meyer (1981) tested the hypothesis that an increase in expected variance of attribute performance results in a lower evaluation of the overall product or service. Descriptions of sets of hypothetical pizza restaurants in a number of imaginary cities were used as stimuli. Several "expert" ratings were provided about the overall quality of each restaurant. These ratings were on a 0 - 100 scale, 100 meaning best pizza restaurant in the US. In each imaginary city (that is, set of restaurant descriptions) one restaurant was singled out as the "target restaurant". The means and variances of the target restaurants' quality ratings were manipulated in a factorial design. The subjects were asked to indicate the probability of selecting the target restaurant in each city for a dinner out.

Analysis of variance revealed significant main effects for both the expected mean performance and its variance in each of three experimental subdesigns. This means that an increased variance in the expected performance had a significant negative impact on the probability of choosing the target restaurant. Meyer's hypothesis was therefore supported that an increase in the variance of expected performance results in a decreased perception of utility.

Furthermore, the expected mean performance and the variance of expected performance showed a significant interaction in two of the three subdesigns and a marginally significant interaction in the third ( $p < .10$ ). This interaction between mean and variance implies that the marginal effect of each is not constant, but varies as a function of the level of the other. For example, when the expected attribute performance is low, variability matters little (the utility remains low),

but when the expected mean performance is high, variability matters more. In other words, once one variable takes an unacceptable level (low mean performance or high variability) it is difficult to compensate for this value by improving the value of the other variable.

In summary, consumers prefer attributes with small variances in performance, and may therefore trade off between predictability (variance of performance) and expected mean performance. From a managerial point-of-view this means that ensuring a higher degree of predictability can offset a lower level of performance.

Sarel (1978) manipulated the magnitude of the consequences of potential outcomes while keeping the degree of uncertainty constant across the experimental conditions. He operationalized two levels of risk (magnitude of consequences) with a simulated choice of a restaurant for occasions of different importance. On one occasion, the subject would attend the restaurant with his/her boss and an important customer (high consequence situation). On the other occasion the subject would attend with his/her friends for a casual meal (low consequence situation).

The results of his work show that consumers evaluate alternatives with uncertain attribute performances lower in high consequence situations than in low consequence situations. This means, an alternative with high performance uncertainty becomes less attractive the higher the potential negative consequences of poor performance. Furthermore, Sarel showed that consumers do not base

their evaluations only on the mean performance of attributes but also on the extreme tails of the performance distribution (termed unacceptable and outstanding thresholds). Sarel's results indicate that consumers try to reduce risk by focusing on the unacceptable threshold in high consequence situations, while in low consequence situations they put more emphasis on the outstanding threshold (risk-taking behaviour).

Sarel's findings are in contrast to the generally applied multiattribute models. The implicit assumption in most of these models is that consumers base their decision strategies on some derivative of the expected mean performance, as it is the case in the classical Fishbein (1967) expectancy-value model. The Fishbein model assumes that an attitude towards an object or its relative attractiveness can be represented as a combination of the product of (1) a belief component for each attribute, which is intended to measure the probability that a brand offers that specific attribute, and (2) the importance an individual gives to that attribute. This model suggests that consumers should choose the alternative with the highest score irrespective of the shape of the probability distribution. The model assumes that decision makers are risk neutral (Coombs et al 1970). This assumption has been refuted by Sarel (1978).

## 2.3 Summary and Conclusions

When performance heterogeneity can have negative consequences for consumers (e.g. variance in the departure time of airplanes, the quality of food in a restaurant, the perceived attractiveness of haircuts), their perception of risk will increase and bear on their decision making processes (Sarel 1978). Perceived risk theorists have concluded that consumers are risk averse and strive to lower the level of perceived risk by reducing the magnitude of possible negative consequences and/or the probabilities of their occurrences.

Research using multiattribute models with uncertain attributes showed that consumers prefer attributes with small variance in performance over those with high variance. This suggests that consumers trade-off expected mean performance against expected variance in performance. Furthermore, a low variance in performance was shown to be more important for consumers when the magnitude of potential negative consequences was high.

In summary, according to both the perceived risk theory and multiattribute models, performance heterogeneity can have significant impact upon consumers' decision making. Unfortunately, neither of the two bodies of literature has addressed the post-purchase stage or provided a framework that would allow to develop theories on the impact of heterogeneous performance expectations on satisfaction.

### **3 Decision Science Literature on Variance of Expected Outcomes**

Decision making under risk has been extensively addressed in the area of decision science. First, a brief introduction to the main concepts of decision making under risk is given. The theories of the descriptive school seem to be in line with the perceived risk and multiattribute models discussed in the previous section. Similar to the work in consumer behaviour, the decision science literature also does not address the post-choice stage or provide a framework for developing theories on the impact of expected performance heterogeneity on satisfaction.

Second, the scarce literature on the perception of variance is reviewed. This literature shows that people actually can distinguish different levels of variability and can estimate confidence ranges. This supports further the proposition that people perceive heterogeneity in service performances. An other important finding is, from a methodological point-of-view, that variance expectations can be conveniently and effectively manipulated by presenting data sets to subjects. Finally, the findings of this section are summarised.

### 3.1 Decision Making Under Risk - Normative and Descriptive Models

Decision making under risk has been investigated by both theorists of the normative school, who prescribe rules for optimal choice, and of the descriptive school, who focus on how decisions are actually made. The main models of the normative school are reviewed in the next few paragraphs.

The Expected Value Model ignores all features of the distribution of outcomes other than the mean. The expected value (EV) is:  $EV = \sum P_i V_i$ , where  $P_i$  and  $V_i$  are, respectively, the probability of occurrence and the value (usually measured in monetary terms) of the  $i$ th consequence. This model assumes that decision makers are indifferent between alternatives with the same EV, and choose the alternative with the highest EV irrespective of the shape of the probability distribution of outcomes. This means for example, that a person would be indifferent between taking an action that leads to a sure £500 and an action that is characterized by a 50% chance of gaining £1,000 with a 50% chance of gaining nothing.

In the Expected Utility (EU) Model the objective values  $V_i$  are replaced by subjective utilities  $[U(V_i)]$  for all values. Decisions are based on maximization of expected utility:  $EU = \sum P_i U(V_i)$ . This model allows individuals to have different utilities for the potential consequences and hence different preferences among risky alternatives. For example, a risk-seeking person would prefer the option with a 50% chance of gaining £1,000, rather than taking the sure £500. A risk-

averse person on the other hand would prefer the sure £500. An individual with linear utility function is said to be risk-neutral (Hogarth 1987) and would behave in the same way as described in the expected value model.

The Subjective Expected Utility (SEU) Model is a further development of the EU model and substitutes for the objective probability  $P_i$  a perceived or subjective probability  $S(P_i)$ . In other words, this model allows for different probability perceptions. Decisions are based on maximization of SEU:  $SEU = \sum S(P_i) U(V_i)$ , where  $S(P_i)$  is the subjective probability corresponding to the objective probability of the  $i$ th consequence. Over the last three decades, the SEU model has received much attention in the literature. The model seems to provide a good fit to choice data (Sarel 1978). In contrast to the EV and EU models, which are strictly normative, the SEU model can also be used to describe decision making of under risk (descriptive school).

In contrasting to the models which postulate maximization of expected value or utility, several security models have been developed which emphasize minimizing or limiting risk. For a review refer to Day et al (1971).

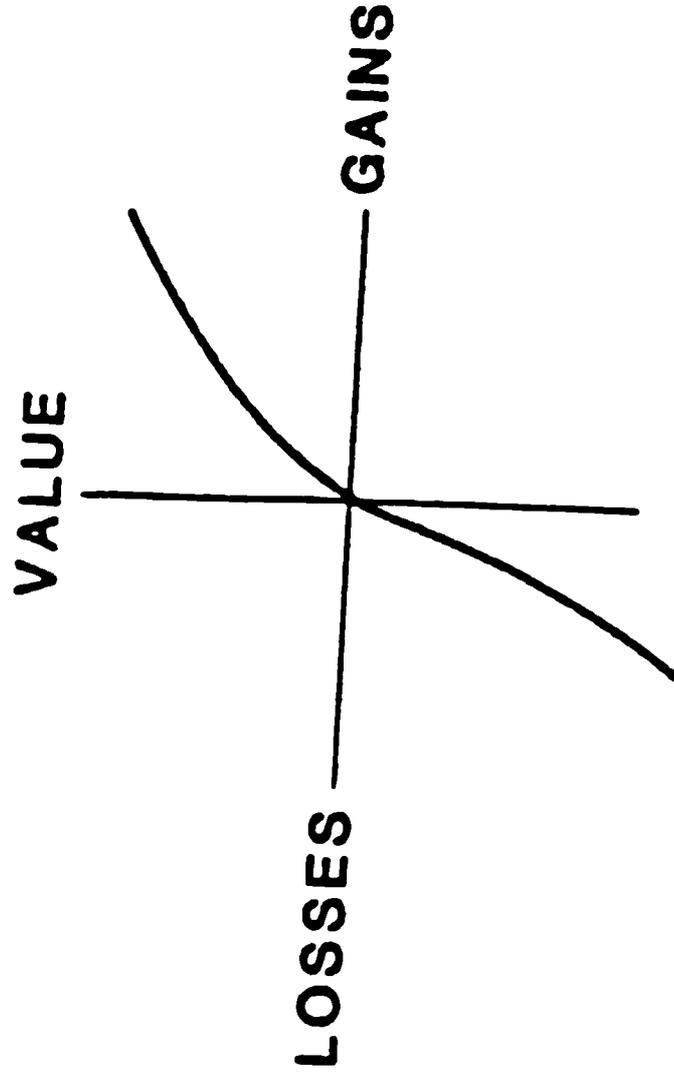
The descriptive school of decision making focuses on how decisions are made in the real world rather than trying to provide guidelines for decision making. The main descriptive models are Prospect Theory and the Ambiguity Model.

Kahneman and Tversky (1979) formulated Prospect Theory, which builds on the Expected Utility Model. Prospect Theory has two components, a value function that plays the role of the utility function in the expected utility model, and a weight function. The value function (Figure 3-1) has three characteristics. First, values are assigned to changes in wealth or welfare (gains or losses) rather than to final assets. That is, people are assumed to think in terms of gains or losses. The key psychological insight behind this assumption is that, since people have limited information-processing abilities, they are sensitive to differences. Thus, it is natural to think of outcomes in terms of deviations, usually increases or decreases from the status quo. The second characteristic of the value function is that its shape captures the notion that people are more sensitive to differences between outcomes the closer these are to the status quo. The third characteristic of the value function is that its slope over the range of losses is greater than its slope over the range of gains. This translates the notion that people experience losses and gains with different levels of intensity. Specifically, losses loom larger than gains.

According to Prospect Theory, decisions are based on maximization of expected increase in value (EIV):  $EIV = \sum P_i IV_i$ , where  $P_i$  and  $IV_i$  are the probability and the increase in value of the  $i$ th consequence, respectively.

Prospect Theory assumes that people know the precise (objective) probability distributions of outcomes as in the expected utility model. The cases where they do not are addressed by the Ambiguity Model.

Figure 3-1: A Typical Prospect Theory Value Function



In the Ambiguity Model (Einhorn and Hogarth 1985), people are assumed to assess ambiguous probabilities for alternative outcomes by first anchoring the probability of one outcome, and then mentally simulating the probabilities of other potential outcomes. The assessment of uncertainty in conditions of ambiguity therefore involves a compromise between "what do I know best" (the anchor) and "what could be the probabilities of other outcomes?" (the product of imagination and mental simulation).

The descriptive models discussed so far ignored the risk preferences of individuals. This issue was addressed by Coombs (1975), who proposed Portfolio Theory in which risk preferences play a central role. According to Coombs decisions are made on the basis of a compromise between maximization of expected value and optimization of risk. Portfolio Theory assumes that individuals have an ideal level of risk at each level of expected value, which allows for trade-offs between higher potential payoffs and uncertainty. The notion of an optimal risk level is intuitively appealing and might be useful as a complementary approach to theories proposed earlier (Sarel 1978).

### 3.2 The Perception of Variability

In the literature on perceived risk and multiattribute models it is often implicitly assumed that people perceive variability of product or service performances. However, none of this research has addressed the perception of variability directly. In the decision science literature a limited amount of work has been done that suggests that people perceive and compare the variability of distributions and can estimate confidence ranges. The first subsection reviews research on how people discriminate between the variability of different data sets. The second subsection looks at research on how people estimate the variability of a single data set.

#### Discrimination Between the Variability of Different Data Sets

Fike and Ferrell (1977) and Fike (1977) reviewed the literature on variability perception. They concluded that many individuals consistently detect the sample with the higher variability when comparing two data sets. This holds even when the data sets are presented in quite different forms, for example as points on a line or as verbal descriptions.

For example, Rich (1977) conducted a study in which data sets were presented as histograms, and subjects compared the variability of the data sets. Rich drew the following conclusions. In spite of making reasonably accurate estimations of variability, people tended to use heuristics that produced consistent mis-estimates: (1) Outliers tended to be ignored if they were sufficiently separated from the rest of the sample, (2) if the shape of the histogram tapered off on either end with

progressively increasing spacing between data points, the sample tended to be perceived as having a higher variability than if it were more uniform, and (3) if the shape and size of two histograms were very similar, the corresponding data sets tended to be perceived as having the same variability, independent of where the sample points were located.

### Perception of Variability of Single Data Sets

Some studies were conducted to investigate the perception of variability of single data sets. Fike and Ferrell (1977) distinguished between the perception of variability (1) when no sample data are provided, and (2) when sample data are provided.

When no sample data are provided, people can base their judgement about variability of outcomes only on logical derivation and/or their memory. Experiments show that people tend to underestimate confidence ranges (e.g. 90% confidence ranges) in these situations (e.g. when estimating the variability of the winners' final scores in golf tournaments based only on memory and logical derivation). The true values lie outside those ranges far more often than people estimate. This phenomena of hypercertainty (being more certain than the data allow) suggests that subjective probability density functions are much too narrow (Fike and Ferrell 1977).

In situations where sample data are provided, people estimate confidence ranges with reasonable accuracy. The estimates do not show the extreme hypercertainty

characteristic of estimates made without sample information (Fike 1977). The researcher suggests that when data are directly presented, the perception of variability is akin to basic perceptual processes and the heuristics that are used include organization and smoothing principles common to much perception (e.g. outliers tend to be ignored).

Fike and Ferrell (1977) concluded that when sample data are provided, the accuracy of estimation is always superior to the accuracy in situations where no data are available. This is because when no data are provided, people have to rely on their memory and logical thinking with large intersubject differences due to differing experiences, perspectives and logical thinking.

### **3.3 Summary and Conclusions**

The literature on decision making under uncertainty can be categorized into normative and descriptive. Several normative models, which can aid people in their decision making, have been proposed. Some commonly applied models, which are based on probability distributions, were discussed. Three descriptive models were reviewed. These models suggest that people assign subjective values to changes in wealth or welfare (gain or losses) rather than to final wealth or assets (Prospect Theory) and weight them with subjective probabilities (Ambiguity Model). Furthermore, Portfolio Theory suggests that people compromise between

maximizing their expected utility and optimizing their risks, which implies that each individual has an ideal level of risk for each expected value.

As variability is an important underlying dimension of risk, research on how people perceive variability has been carried out. It has been shown that people can estimate the variability and 90% confidence ranges of samples with reasonable accuracy when the raw data are provided. However, when no data are presented and people have to rely on their memory and/or logical thinking, they tend to underestimate variabilities and confidence ranges.

A number of conclusions with respect to heterogeneity in consumer choice processes can be drawn. First, expected variance in outcomes has been extensively researched in the decision science literature. The concepts reviewed (e.g. the prospect and ambiguity theories) support recent developments in marketing, such as theories about the evaluation of uncertain attribute performances (Meyer 1981) and magnitude of potential consequences (Diamond 1988, Sarel 1978).

Second, neither decision science nor consumer behaviour research has addressed the post-choice process which could help to develop theories on the impact of heterogeneous performance expectations on the satisfaction process.

Third, the research on variability perception shows that people actually perceive different levels of variability, although people tend to underestimate confidence ranges when no data sets are provided. This supports the implicit assumption in

much services marketing literature that consumers know about the high variability of service performances.

Finally, the relatively high accuracy in variability perception obtained with presentation of data sets suggests a convenient way for researchers to manipulate expectations about the variability of service performances.

#### **4 Summary and Hypotheses Development on the Impact of Performance**

##### **Heterogeneity on the Evaluation Stage - Hypotheses II & III**

The discussion in the first section of this chapter highlighted the fact that a predictable service encounter is favourable from an operations management perspective. Through industrialization of service production, minimization of customer contact time, client selection and socialization, as well as routinization of the service encounter, both workflow and task uncertainties can be minimized and operations efficiency maximized. For customers this generally means that a restricted and standardized (i.e. not customized) service is provided. Apart from direct consequences of a more efficient service production (i.e. lower price), it also has the following positive aspect. For consumers, perceived risk will be reduced through reducing the high degree of performance heterogeneity typically found in services.

The impact of performance heterogeneity on the consumer choice process and on the perceived risk inherent in any purchase decision has been well researched. This research has shown that perceived risk increases with the magnitude of potential negative consequences and the probability of their occurrence. Research in the context of multiattribute models has shown that higher uncertainty about the likely attribute performance results in discounted attribute evaluation. This discounting will be greater the greater the potential negative consequences. Therefore, consumers seem to trade-off expected performance heterogeneity against mean performance.

Research in the areas of consumer behaviour and decision making under uncertainty stands in sharp contrast to the study of post-purchase evaluation. Satisfaction researchers have always implicitly assumed point-expectations; performance heterogeneity has never been considered. This is surprising, considering the fact that services generally have a higher degree of unit-to-unit variability than products (e.g. Berry 1980, Krughoff 1981).

The review of variance perception has shown that people can estimate confidence ranges reasonably well. Therefore, it would seem logical that consumers hold some kind of confidence range in which they expect certain service performances to occur. In other words, it is suggested that consumers expect a mean performance with some confidence range and/or standard deviation, whereby the expected confidence range and/or standard deviation may vary between service classes, service providers and even service employees of the same service company. The impact of expected performance heterogeneity on satisfaction outcomes has not been addressed yet, although a clear understanding of its effects seems crucial for strategic decisions concerning the classical trade-off between standardization and customization, quality specification, quality control and investment in service uniformity within and across outlets.

Two simple hypotheses can be deduced within the context of the disconfirmation-of-expectations model. A test of these hypotheses should offer a first insight into the effects of expected performance heterogeneity on satisfaction processes. The

first hypothesis deals with processes leading to confirmation, and the second deals with processes leading to disconfirmation.

First, the disconfirmation-of-expectation model predicts that any performance that meets pre-consumption expectations leads to confirmation, and subsequently to satisfaction outcomes. Although never stated explicitly by satisfaction theorists, this well corroborated theory can be extended to variance expectations: assuming a unimodal distribution of performance expectations, consumers should not be "surprised" when they perceive a performance that falls within some confidence interval of their expected performance distribution. In other words, consumers should experience confirmation when the performance is close to their expected mean performance, independent of the expected variance. The following hypothesis can be put forward on the basis of this line of argument:

**Hypothesis II:**

**In situations where perceived performance and the expected mean performance are approximately equal, consumers experience confirmation independent of the expected variance in performance.**

In disconfirmation situations, the magnitude of disconfirmation is a function of the difference between expected performance and perceived performance. One can argue that the perceived probability of a particular difference occurring becomes larger, the larger the expected variability in performance. In other words,

consumers should be less surprised to encounter a performance that deviates from the expected mean, if they expect performances to vary a lot than if they do not expect performances to vary at all. If this relationship holds, the following hypothesis should be true:

**Hypothesis III:**

**In disconfirmation situations, consumers who hold mean performance expectations with some degree of variation around the mean, experience smaller magnitudes of disconfirmation than consumers who hold point expectations.**

A test of Hypotheses II and III was considered a first step in exploring the impact of expected performance heterogeneity on disconfirmation and satisfaction processes.

## Chapter IV

# The Concept of Affect: Research and Theoretical Developments in Psychology

In consumer behaviour theory, affect seems to be a crucial variable for capturing the experiential nature of services. In particular, the environmental psychology approach suggests that attributes of retail store environments, such as brightness, crowding, aisle width, colours and music, drive the affective state of consumers in the store (Donovan and Rossiter 1982). Control theory holds that perceived control is an important determinant of affect in the service encounter, and this affective state in turn determines the behaviour of consumers. If affect is to be included in satisfaction models to capture the experiential nature of services, we need to understand the concept of affect and its causes. The aim of the chapter is to provide this understanding.

Furthermore, this chapter lays the foundation for two important issues of the thesis. First, it provides the basis for selecting the currently best available conceptualization of affect to be used in the thesis. This is Russell's (1978, 1980) circumplex model of affect. Second, the review shows that cognitive processes of any degree of complexity can cause affective responses. This conclusion serves as a basis for the hypotheses development in Chapter V on the impact of disconfirmation-of-expectations on affect.

In this chapter the affect literature in psychology is reviewed. In the first section, the concept of affect is presented, which includes its definition, its delineation from cognition, and the different models of affect that have been proposed. The second section reviews the causes of affect. This section is of particular interest. In it, the recent challenge to the traditional view that cognitive processes are necessary causal antecedents of affect is examined. A summary of the chapter is provided in the last section.

## **1 The Concept of Affect**

Affect is a diffuse concept, hard to describe and even harder to differentiate and classify. This makes affect a difficult construct to investigate theoretically and empirically. As a consequence operationalizations and measures vary widely. Given this, a clear definition and terminology are needed, not because there might be right or wrong answers to them, but because unless they are settled, conceptual discussions will be impeded. The prevailing definitions in psychology are presented in the first subsection and a delineation of affect and cognition is outlined in the second subsection. In the third subsection, different conceptualizations of affect are discussed and compared. Finally, some conclusions are drawn.

## 1.1 Definition of Affect

In the psychology literature, affect is typically treated as a synonym for feelings and emotions. For example, Simon (1982) defines affect as a generic term that includes evaluations, moods (other theorists refer to those two as feelings) and emotions.

Feelings include subjective reactions that are essentially either pleasant or unpleasant. The feelings most frequently studied by social psychologists are affective evaluations: that is, simple positive and negative reactions (attraction, liking and so forth) to stimuli such as people, objects or experiences (Fiske and Taylor 1984). Evaluation refers to cognitive association of cognitive labels attributing positive or negative valence to objects or events (Simon 1982). Evaluations can be distinguished from moods which do not have a specific target and do not interrupt attention. One can have an evaluative reaction toward another person, but one typically does not have a mood reaction toward another person. Like evaluation, mood is primarily considered as simply positive or negative (Fiske and Taylor 1984). Moods refer to affect that provides a context for ongoing thought processes without noticeably interrupting them and whereby only low levels of arousal are experienced (Simon 1982). Typical examples for moods are sadness and happiness.

Emotion refers to affect that interrupts and redirects attention, usually accompanied by arousal of the autonomic nervous and endocrine systems. These interrupting emotions may originate in sensory stimuli (e.g. detection of a rapidly moving object),

in stimuli coming from the autonomic nervous system (hunger) or by evocation of affect-tinged items in long-term memory (recall of an embarrassing experience) (Simon 1982).

## **1.2 Delineation Between Affect and Cognition**

Affective and cognitive processes are difficult to separate because they are highly interrelated, at least on the prevailing assumption that the affective response to anything depends upon the way in which it is first perceived and cognized. Nevertheless, several theorists have developed conceptual differentiations between the two concepts.

Kuhl (1986) distinguished between cognition and affect in the following way. He assumes that cognitive and emotional processes relate to the world in different ways. The term cognition refers to those processes that mediate the acquisition and representation of knowledge about the world, i.e. processes that have a representative relation to the world of objects and facts. Affective processes evaluate the personal significance of those objects and facts.

Fiske and Taylor (1984) give some further differentiation: affect includes evaluation, preference, differentiated emotions and defensive attribution, while reactions representing cognition include attention, inference and memory. Affect and cognition differ in quality. For example, one can be wrong about a recognition

judgment but not about an affective judgment. One can also argue that recognition judgments are more complex than affective types of judgment. In general, Fiske and Taylor (1984) suggest that affective judgments are by nature subjective, simple and direct.

Russell et al (1981) distinguished between mainly affective components of meaning (the semantic representation of emotion) and perceptual/cognitive components (e.g. quiet versus noisy, clean versus dirty). Affect reflects a person's feelings about (including attitude toward) an object; cognition reflects a person's implicit semantic theory about (modes of categorizing) the object (Tzeng 1975).

In other words, the delineations between affect and cognition presented refer to cognition as processes related to thinking and knowledge about the outside world and its interaction with the self. Affect, on the other hand, refers to internal states that may be caused by sensations, perceptions and thoughts.

### **1.3 Conceptualizations of Affect**

Various theorists have tried to examine the content of affect, the dimensions that underlie affect and the types of affect that might be distinguished. Different approaches have been used. For example, typologies of affect were derived (1) from facial expression research (Izard 1977, Tomkins 1980, Ekman et al 1983), (2) from language-based research (Mehrabian and Russell 1974, Russell 1980, Russell et al 1981) and (3) from logical theory derivation and testing based on a psychoevolutionary perspective (Plutchik 1980). The derived conceptualizations of affect are discussed and evaluated in the following subsections.

#### **1.3.1 Classifications Based on Facial Expression Research**

Facial expression researchers build on the notion that the face provides immediate and specific information on the internal affective state of humans. In addition to its neural and experiential aspects, each fundamental affective state has a characteristic facial expression (Izard 1977). These expressions are instinctive in the sense that they occur reflexively as part of the emotional process.

Tomkins (1980), Ekman (1982) and Izard (1977) developed taxonomies of affect on the basis of facial expressions which classify the variety of affective states into small sets of fundamental or primary affects. These three taxonomies are in substantial agreement about the basic categories of affective experiences (for a

comparison, see Ekman, Friesen and Ellsworth 1982). Table 4-1 lists the fundamental affects identified by Izard (1977).

Scales that measure facial expressions and movements have been developed. One recently developed measure is the Facial Action Coding System (FACS) of Ekman and Friesen (1978). The FACS differentiates between facial muscle actions as observed by human operators. FACS includes most of the subtle differences in appearance that results from muscle action. FACS encompasses 33 Action Units, most of which involve a single muscle. The system has been shown to possess good internal and external validity (Ekman 1982). However, learning the system takes approximately five weeks on a part-time basis, which severely limits its application on a large scale. So far, only few researchers (less than a hundred) have acquired the skill to use the FACS.

To provide a tool for convenient measurement, Izard (1977) developed a subject self-report verbal scale, the so-called Differential Emotions Scale (DES). The DES instructions ask subjects to rate the extent to which each adjective of the scale describes the way subjects feels at the present time. Single five-point intensity scales are employed. The scale consists of thirty adjectives, three for each of Izard's (1977) ten fundamental emotions as listed in Table 4-1. The self-report DES is consistent with external observations of subject's emotional facial expressions (Ekman, Friesen and Ancoli 1980).

A drawback of facial expression based models is that they conceptualise affect as an unstructured collection of discrete emotions. In other words, our intuitions about the similarities and differences among emotions are NOT captured by these models (Smith and Ellsworth 1985). The conceptualisation of affect discussed in the next section tries to model these similarities and differences in a two dimensional space.

**Table 4-1: Izard's (1977) Taxonomy of Affective Experiences**

Basic Affect	Nature of Subjective Experience
Interest	Engaged, attentive, caught-up, curiously fascinated; when intense, a feeling of excitement and animation
Joy	Sense of confidence and significance; feeling loved and lovable; a good relationship to the object of joy
Anger	Hostility, desire to attack the source of anger, physical power, impulsiveness
Disgust	Feeling of revulsion; impulse to escape from or remove the object of disgust from proximity to oneself
Contempt	Superiority to other people, groups, or things; hostility (mild); prejudice; coldness and distance
Distress	Sadness; discouragement, downheartedness; loneliness and isolation; feeling a miserable sense of loss
Fear	Apprehension to terror, depending on intensity; sense of imminent danger; feeling unsafe; slowed thought; tension
Shame	Suddenly heightened self-consciousness, self-awareness; feeling of incompetence, indignity, defeat; in mild form, shyness
Guilt	Gnawing feelings of being in the wrong, "not right" with others or the self
Surprise	Fleeting sense of interruption of ongoing thought; brief uncertainty, amazement and startle

### 1.3.2 Russell's Circumplex Model of Affect

Mehrabian and Russell's (1974) work builds directly on Osgood's (1969) repeated identification of three dimensions of meaning: evaluation, activity and potency. The Mehrabian and Russell (1974) framework stresses three analogous dimensions, which they renamed to better reflect their interpretation of emotion: pleasure, arousal and dominance - which together constitute the so-called PAD paradigm. Mehrabian and Russell suggested that any specific emotions may be regarded as positions on the three PAD dimensions. These researchers also propose that the three basic bipolar dimensions mediate human behavioural responses towards environments and social situations. The three intervening variables (dimensions) link all stimuli of environments and social situations to a variety of human behaviours such as physical approach, task performance, explorations and interpersonal relationships. These behaviours can be generalized along an approach-avoidance continuum.

Mehrabian and Russell (1974) used more than a hundred affective adjectives along which subjects had to rate numerous environments. The three affective dimensions were determined through statistical data-reduction methods (principal component analysis, factor analysis and cluster analysis).

More recent research, however, has failed to demonstrate dominance as a dimension of affect similar to pleasure and arousal (e.g. Russell 1979, 1980; Russell and Pratt 1980; Russell et al 1981). Russell has concluded that only two dimensions, namely

pleasure and arousal, are sufficient to conceptualize affect. Dominance only accounted for a very small proportion of the variance in affect. The original Mehrabian and Russell (1974) model also hypothesised that the three dimensions were orthogonal. However, in Russell's (1980) study, dominance was highly correlated with pleasure ( $r=.65$ ) and marginally correlated with arousal ( $r=.25$ ). Similar results were encountered in other studies (Donovan and Rossiter 1982, Russell et al 1981). In addition, multidimensional scaling studies of affect indicated that apart from dominance, there were a number of other dimensions beyond pleasure and arousal contained in the meaning of verbal affect terms (Russell 1978, Russell and Pratt 1980). These dimensions included locus of causation, importance of the emotion and locus of control. Whereas pleasure and arousal each accounted for substantial proportions of the variance in affect, each dimension beyond these two accounted only for a trivial proportion.

Russell and his associates have contended from these results that there exist two categories of emotional terms that people employ to describe their affective states. The first category includes those terms that primarily describe an individual's internal state. Russell and his colleagues suggest that the two dimensions of pleasure/displeasure and degree of arousal denote the internal emotional state per se. All other dimensions belong to the second category and are suggested to be clearly interpretable as cognitive/perceptual rather than affective in nature. Examples of these dimensions are locus of causality, importance and dominance, which are general information processing concepts. They denote beliefs about antecedents and consequences of an individual's internal affective state. Accordingly,

dominance as a cognitive/perceptual construct is hypothesized by Russell to be an antecedent of pleasure and arousal rather than an independent affective dimension.

As a result of their stream of research Russell and his colleagues (Russell 1978, 1980, 1983; Russell and Pratt 1980; Russell et al 1981) have proposed a new circumplex model of affect which is a modification of the original Mehrabian and Russell (1974) model. The evidence for their circumplex model suggests that affect can be represented by a two-dimensional space with three specific properties. First, the two dimensions defining affect (pleasure and arousal) are independent of each other. In the Russell and Pratt (1980) study, two principal components (eigenvalues 3.56 and 3.04) accounted for 82.4% of the variance, the next component accounted only for a small increment in variance (eigenvalue 0.36). The first two components, clearly interpretable as pleasure and arousal, also accounted for about equal proportions of variance (44% and 38% respectively).

Second, the two dimensions are bipolar. As depicted in Figure 4-1, the eight unipolar scales (for measures refer to Tables 6-10 and 6-11 in Chapter VI) fall on four bipolar scales. Even though the scales were constructed separately and with no overlapping items, each pair of scales was empirically shown to measure opposite ends of the same continuum: arousing is opposite of sleepy, exciting is opposite of gloomy, pleasure is opposite of displeasure and distressing is opposite of relaxing.

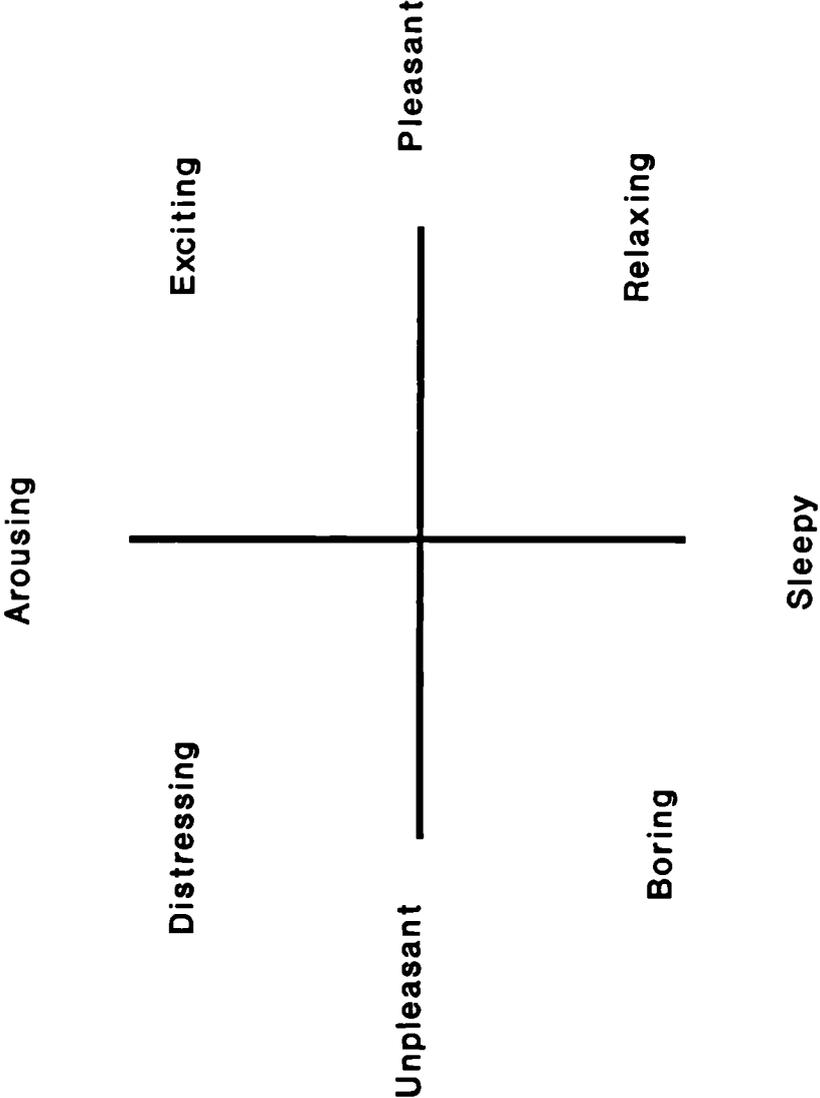
Third, the various affective states do not cluster about these axes (do not form a simple structure), but are placed meaningfully throughout the space. Figure 4-2

is an empirical plot of the eight scales located in a two-dimensional space defined by their rotated principal components (Russell and Pratt 1980). For example, festive or exciting places are both pleasant and arousing. Peaceful and comfortable places are also pleasant but unarousing.

The Russell (1978, 1980) circumplex model of affect should prove valuable to environmental and social psychologists in a number of ways (Russell and Pratt 1980, Russell et al 1981). (1) It can be used in studies which are aimed at understanding those objectively specifiable properties of environments and social situations that produce or influence affective reactions (determining the antecedents of affect). (2) The model can be used to explain behavioural reactions to environments or situations. For example, Mehrabian and Russell (1974) and Mehrabian (1980) propose that affective reactions to environments influence diverse behaviours like social interaction, drug consumption and work performance (determining the consequences of affect). (3) The circumplex model of affect can also be used as means of environmental and situational assessment (determining affective attributes of environments and situations).

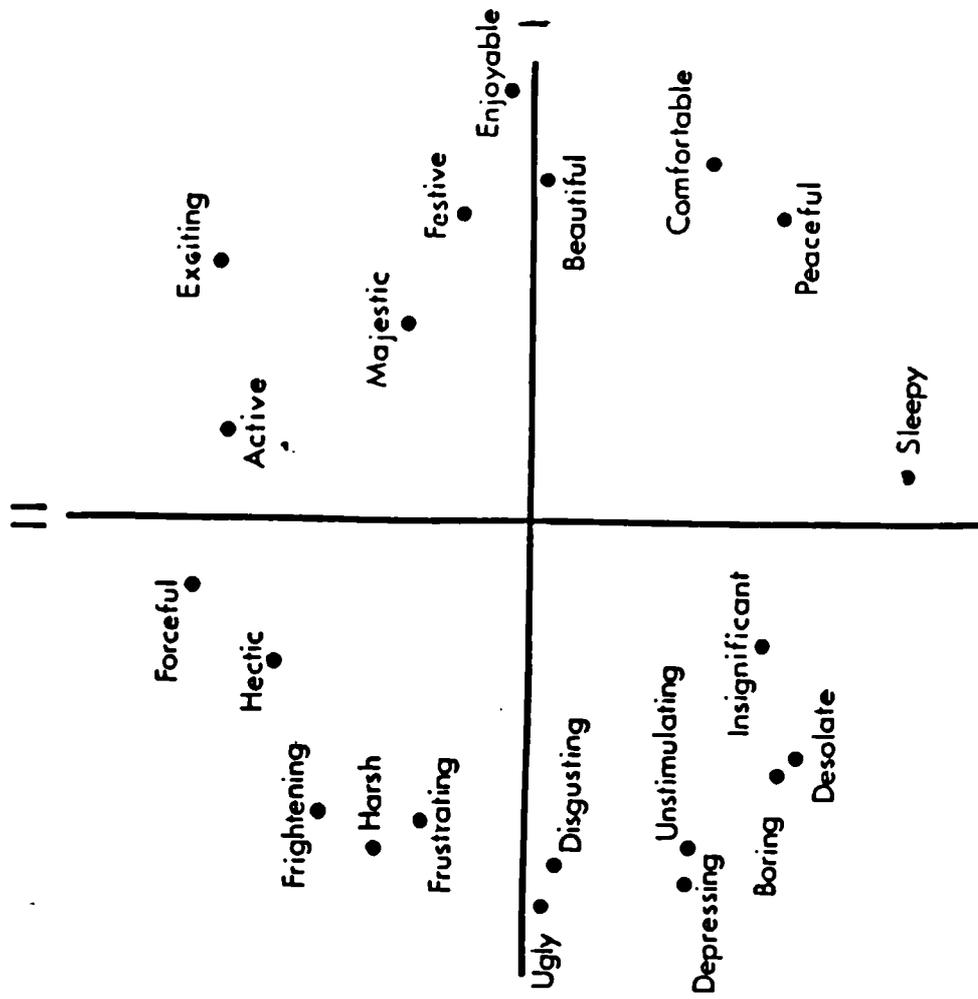
In summary, Russell's circumplex model of affect (1978, 1980) offers a means of replacing the commonly applied but broad and vaguely defined categories of affect with a conceptualization of more precise specification. This provides researchers with a well founded base for hypothesis development and testing. Further research will hopefully advance the understanding of the antecedents (e.g. stimuli and cognitive processes) and consequences (cognitive processes and behaviour) of affect.

**Figure 4-1: Russell's Circumplex Model of Affect**



Adapted from Russell and Pratt (1980)

Figure 4-2: An Empirical Plot of the Affective Quality of Different Environments



Taken from Russell, Ward and Pratt (1981) p. 277

### 1.3.3 Plutchik's Eight Basic Emotions

Plutchik (1980) derived a classification by adopting a psychoevolutionary theoretical stance in which the primary emotions were seen as performing key adaptive functions aiding species survival at all stages of evolution. In his opinion, the struggle for survival requires the performance of eight adaptive functions. A basic emotional and behavioural component corresponds to each adaptive function (for an overview refer to Table 4-2).

Table 4-2: Plutchik's Eight Basic Emotions

Adaptive Function	Basic Emotion	Behaviour
Incorporation	Acceptance	Affiliating
Rejection	Disgust	Repulsing
Protection	Fear	Escaping
Destruction	Anger	Attacking
Reproduction	Joy	Cooperating
Reintegration	Sadness	Crying for help
Orientation	Surprise	Stopping
Exploration	Expectancy	Exploring

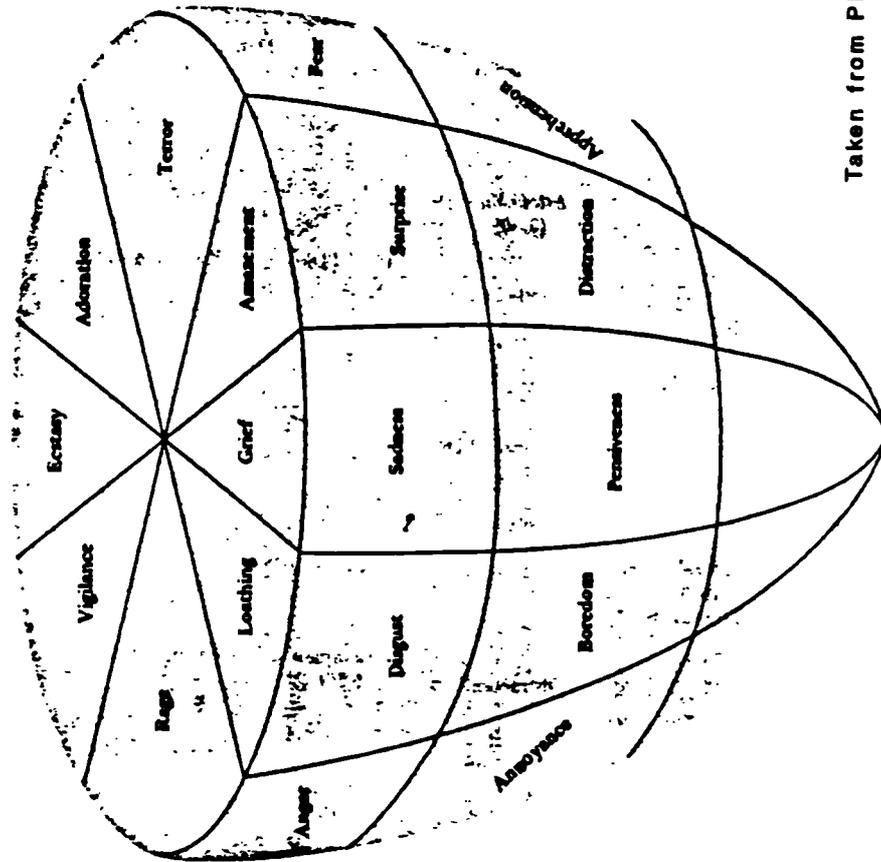
In a further structural elaboration, these eight primary components may be regarded as pairs of polar opposites (acceptance and disgust, fear and anger, joy and sadness, surprise and expectancy). These polar opposites may be arranged in a circular pattern or a circumplex (acceptance - fear - surprise - sadness - disgust - anger - expectancy - joy) within a conic shape in which height conveys intensity: boredom vs. disgust, fear vs. terror and anger vs. rage. More complex emotions can be derived by mixing these primary emotions in varying degrees - much as one might mix primary colours to obtain all possible hues. For example, joy blended with acceptance gives love or friendliness; acceptance mixed with fear yields submission; disgust combined with anger can produce hate, hostility or scorn (Plutchik 1980, p. 162).

Figure 4-3 presents this structural model of affect. It depicts the eight basic emotions somewhat like the section of half an orange, with the emotions of maximum arousal at the top. The level of arousal ranges from a maximum at the top to a state of deep sleep at the bottom. The shape of the model implies that the affective states become less distinguishable at lower levels of arousal.

Plutchik (1980) conducted extensive research using a set of verbal rating scales which he called the Emotion Profile Index (EPI). Application of data reduction techniques gave results that generally conformed to Plutchik's model of eight discrete emotions.

In summary, Plutchik proposes that affect can be defined by eight discrete emotions, their combinations plus one arousal dimension. The drawback of Izard's model of affect applies also to Plutchik's concept. It does not capture the similarities and differences among the emotions. A model with dimensions along which the different discrete emotions vary has not been provided by Plutchik.

Figure 4-3: Plutchik's Structural Model of Affect



Taken from Plutchik (1980) p. 157

### 1.3.4 Comparison of the Concepts

A comparison of the three different taxonomies of affect, derived from different conceptual approaches (facial expression research, language based approaches and psychoevolutionary theoretical argument) is made in Table 4-3.

As shown in the table, there is some agreement on the feelings and emotions encompassed by the affect concept, although the three conceptualizations are based on different approaches. For example, pleasure or joy is a dimension in all three conceptualizations. A difference between these models is that the Russell model explicitly encompasses arousal as a bipolar dimension. Izard's model has a number of unipolar dimensions with different degrees of arousal. Arousal has a more explicit role in Plutchik's model than in Izard's, but Plutchik also did not establish a clearly identified dimension of arousal.

The second major difference between the three models is that Russell's model can be used to describe similarities and differences between internal states of people on two orthogonal dimensions. The other two models only define discrete emotions, which do not capture our intuitions about the similarities and differences among emotions (Smith and Ellsworth 1985).

The third difference between the concepts is that the Russell model separates cognition from affect. Russell defines affect as an internal state of humans comprising only pleasure and arousal. All other dimensions of affect that have been

proposed are interpreted by Russell as being cognitive in nature and not belonging to the concept of affect. In contrast, Izard's and Plutchik's models implicitly include cognitive processes in their models. For example, both models include anger as a discrete emotion which implicitly requires the cognitive construct of attributions of causes.

Table 4-3: Comparison of Different Taxonomies of Affect

<u>Taxonomies:</u>	Izard (1977)	Plutchik (1980)	Russell (1978, 1980)
<u>Dimensions:</u>	10 discrete emotions	8 discrete emotions + 1 arousal dimension	2 bipolar dimensions
<u>Measurement:</u>	FACS, observation	EPI, verbal rating scale	verbal rating scale
<u>Theory Derivation:</u>	facial expressions	psychoevolutionary adaptation	language based
<u>Primary Affective States:</u>			
Pleasure/Joy	X	X	X <sup>1)</sup>
Interest	X		
Exciting			X
Expectancy		X	
Surprise	X	X	
Arousal		X	X
Anger	X	X	
Distress	X		X
Fear	X	X	
Sadness		X	
Disgust	X	X	
Unpleasant			X
Shame	X		
Guilt	X		
Gloomy			X
Contempt	X		
Sleepy			X
Relaxation			X
Acceptance		X	

<sup>1</sup> Eight unipolar affective states are presented in this column. According to Russell they can be conceptualized by two orthogonal bipolar dimensions of pleasure-displeasure and arousal-sleepy.

An empirical comparison of Plutchik's EPI and the original PAD paradigm was conducted by Holbrook (1986) in the context of product/service consumption experiences. Written consumption experiences were rated by subjects on Plutchik's EPI and Mehrabian and Russell's PAD scales. Both measures showed good levels of convergence validity. The Cronbach alpha for the EPI scales was .88 and .93 for the PAD scales. A canonical correlation analysis showed that the two PAD dimensions (pleasure and arousal) alone explained 53.6% of the variance in Plutchik's EPI, whereas four dimensions of the EPI model were necessary to achieve a comparable level of explained variance (53.7%) in the PAD model. Holbrook came to the conclusion that the PAD paradigm contains more information than Plutchik's EPI even though its scales contain fewer items. Also, Holbrook encountered strong intercorrelations between Plutchik's measures. This indicates that fewer than Plutchik's eight basic emotions might be sufficient to adequately represent affective states.

This empirical comparison was conducted with the PAD paradigm (Mehrabian and Russell 1974) and not with the later modified version, the Russell circumplex model of affect (1978, 1980). Similar results might be expected using only the two dimensions, pleasure and arousal, without including dominance as a third dimension. This is because dominance has accounted for only a small part of the explained variance in many studies (e.g. Holbrook 1986, Russell 1978, 1980).

Similar to Plutchik's (1980) measures, problems of discriminant validity within Izard's (1977) measures have been reported. Izard's measures of the basic affective

states, derived from facial expression research, have been shown to correlate highly (Westbrook 1987). These findings suggest some caution about the discriminant validity of the DES and EPI scales.

Empirical studies show that the Russell's measurement scales have the highest discriminant validity, and their predictive or explanatory power seems to be empirically well corroborated. In contrast, Izard's DES and Plutchik's EPI scales show high levels of correlation between the measures of their different dimensions. This contradicts their claimed orthogonality. Furthermore, Russell's model uses less dimensions and fewer scale items than the others, which makes its application more convenient.

In summary, Russell's (1978, 1980) circumplex model of affect is empirically the best corroborated model and has, in contrast to the other two models, no discriminant validity problems. This makes the Russell model the most promising framework for conceptualizing internal states of people.

#### **1.4 Conclusions**

A number of studies have provided strong support for Russell's (1978, 1980) two dimensional circumplex model of affect. Its predictive and explanatory power as well as its convergent and discriminant validity seem to provide good internal and

external validity. Furthermore, the small number of items in its measures makes this conceptualization convenient for use in marketing research.

Russell's model of affect suggests a redefinition of affect to the one presented at the beginning of this chapter. All hitherto proposed definitions have not clearly distinguished between varying degrees of arousal. The only hint at a differentiation is the delineation between feelings and emotion, whereby feelings are implicitly suggested to have low degrees of arousal and emotions to have high degrees of arousal. However, the conceptualization proposed in the Russell model is more explicit, concise and empirically corroborated. Therefore, the Russell (1978, 1980) conceptualization provided the definition of affect to be used in this thesis.

## 2 Theories on the Causes of Affect

The previous section discussed definitions and conceptualizations of affect. This section reviews theories on its causes. In particular, it is shown that cognitive processes of any degree of complexity can cause affect. This conclusion serves as a basis for the hypotheses development in the next chapter on the impact of disconfirmation-of-expectations on affect.

To provide some understanding of affective experiences and their causes, a number of theories about the relationship between affect and cognition are reviewed. The traditional approach to affect was, and still may be, that of Schachter and Singer (1962), Mandler (1975, 1982) and Lazarus (1982, 1984). These researchers argue that cognitions are necessary antecedents of affect. Affect is viewed as a product of cognitive processes and humans are viewed as primarily cognitive beings. However, since the early 1980's many studies have been conducted to improve our understanding of humans as cognitive-affective beings. Recent research in cognitive psychology indicates that affective processes might operate independently of cognition (Zajonc 1980) or precede cognition (Zajonc and Markus 1982).

This section does not try to provide a reconciliation of these views. Instead, it argues that a large part of the prevailing controversy can be attributed to different meanings given to the terms affect and cognition. On the one hand, some researchers exclude innate preferences from the definition of affect (e.g. Mandler 1982) and on the other hand, some theorists define cognitive processes so broadly

that they include all kinds of stimuli processing, even the unconscious perception of stimuli (e.g. Lazarus 1982). This section will show that a minimum of sensation, perception or information processing is necessary for affective response. Furthermore, it will show that the types of cognition that cause affect can vary from exceedingly simple to highly complex processes.

This section is structured so as to first review the independent and then cognition-affect sequence theories. This order also corresponds to the complexity of the cognitive processes involved. The independent theory deals with affective responses to exceedingly simple and sometimes even unconscious cognitive processes (e.g. the perception of physical or sensory stimuli). In contrast, the cognition-affect sequence theories deal with very complex cognitive processes, such as those explained by schema and appraisal theories. Schema theory looks at the match between a stimulus and its internal representation, and appraisal theories deal with the perceived meaning of stimuli beyond their physical and sensory aspects, such as their causes, consequences and evaluations.

Although these alternative theories about the relationship between cognition and affect are presented and discussed sequentially in the following three sections, they are not mutually exclusive and may often operate together in various combinations (Hoffman 1986). Hoffman pointed out that different cognitive processes may utilize different aspects of the individual's sensory, perceptual and cognitive apparatus and may even operate at different levels of consciousness. At the same time, Hoffman also suggests that affect caused by the more elaborate cognitive processes like

appraisal or schema processing may be more powerful and enduring than affect elicited through more primitive processes like mere exposure or unconditioned direct affective responses.

## **2.1 The Independent Theory**

Supporters of the independent theory argue that affective processes operate independently of cognitive processes. That means that under some circumstances, affective and cognitive processes might proceed in parallel without influencing each other significantly (Zajonc 1980). Zajonc went on to propose that affect in fact can precede cognitive processes. Moreover, affective processes may occur faster and at a more basic level than cognitive processes.

### **2.1.1 Mere Exposure Research**

A main pillar of the independent theory are the findings from mere exposure studies. A mere exposure effect is a person's growing to like an initially unobjectionable stimulus the more frequently it is encountered (Zajonc 1980). Zajonc (1980) and Moreland and Zajonc (1982) established this effect empirically. In their studies people were exposed to nonsense words, Chinese characters or tone sequences. The results consistently supported the mere exposure effect - people like frequently encountered stimuli better than less frequently encountered stimuli.

The mere exposure effect was shown to be independent of subjects' abilities to recognize a stimulus as familiar. The recognition of frequently and infrequently encountered stimuli approximated only chance guessing in Zajonc's (1980) study. In other words, people like frequently encountered stimuli better, even when they do not remember that they have been exposed to them before. The same results were found in a study employing dichotic listening tasks, in which subjects were presented with tone sequences in one ear while their attention was focused on a literary passage heard through the other ear. This experimental design virtually eliminated any possibility of subjects recognizing the tone sequences (Moreland and Zajonc 1982).

### **2.1.2 Unconditioned Direct Affective Responses**

Affective responses can be elicited by physical-sensory stimuli. All that is needed to elicit affect is the registration of a stimulus and a minimum of perceptual organization. These responses include unconditioned affective responses (learning is not necessary in contrast to conditioned responses) that appear to be involuntary and sometimes unconscious (Hoffman 1986). Some researchers also refer to these affective responses as innate preferences (Mandler 1982). This stimulus-affect connection may operate in early infancy, before complex cognitive appraisal becomes possible. These findings may be biologically based as explained by Gibson's (1979) notion that an organism's perceptual system has evolved in order to register information pertinent to survival.

### 2.1.3 Conditioned Affective Responses

It seems likely that almost any stimulus can be conditioned to produce affective responses. Among the important stimuli that often become conditioned elicitors of affect are the sounds of words and the tone of voices. Words can become conditioned stimuli for eliciting affect by being paired with an unconditioned stimulus for that affect, such as an electric shock. Another example is fear as an affective response to the sound of the word "cancer" (Hoffman 1986). Although conditioned affective responses do not fit comfortably with the independent theory of affect the underlying cognitive processes are still considerably less complex than those considered in cognition-affect sequence theories.

## 2.2 Cognition-Affect Sequence Theories

A wide range of social psychology theories and research implicitly assume that thought precedes affect. For example, most stereotyping or schema research is based on the assumption of a cognition-affect sequence. For instance, people apply a cognitive generalization to others on the basis of category membership, which in turn determines affective evaluation (Bringham 1971). The same applies in attitude research: cognitive consistency theories, such as dissonance theory, claim that cognitive discrepancies precede the affective discomfort that leads to attitude change (Abelson et al 1968, Festinger 1957). Weiner's (1980a, 1982) attribution theory holds that perceived causes shape people's affective response to success and failure.

All of this work provides strong evidence for a cognition-affect sequence paradigm and supports the notion that cognitive processes are indeed a major cause of affect.

Cognition-affect sequence work can be divided into two broad areas: (1) affect caused by schema interruption, fit and recall, and (2) affect caused by cognitive appraisal, such as perceived consequences of an event, attribution of its causes and comparison of the outcomes of some behaviour or situation with a standard. A brief review of these two broad areas is provided in the subsections below.

### 2.2.1 Schema-Triggered Affect

A schema is a cognitive structure that represents organized knowledge about a given concept or type of stimulus. A schema contains both the attributes of the concept and the relationships among those attributes (Rumelhart and Ortony 1977, Taylor and Crocker 1981). The schema concept specifically maintains that knowledge is stored in an abstract form, not simply as a collection of all the original encounters (Fiske and Taylor 1984). For example, a "professor" schema is stored in memory as a general case abstracted from various specific professors. The schema then organizes incoming information that is thematically related to professors. In other words, people's schemata enable them to focus their stimuli processing.

The schema concept reflects a concern with top-down, conceptually-based or theory-driven cognitive processes, as opposed to bottom-up or data-driven ones (Rumelhart

and Ortony 1977). Schema theory focuses on the ways in which prior theories or concepts guide people in how they perceive, remember and make inferences from raw data.

Two areas of research have dealt with the issue of schema-triggered affect. First, several theories about affective responses to the interruption and the match between a stimulus and a neutral representation (schema with neutral affective charge) are reviewed. Second, research into affective charged schema is presented.

#### **2.2.1.1 Interruption of Perceptual Schemata**

Mandler (1975) suggests that environmental inputs generate arousal, which in turn is cognitively interpreted. Physiological arousal is produced by interruptions or unexpected events that alert the organism to cope with an environmental contingency. Arousal comes from two related kinds of interruptions. The first type is the disconfirmation of a perceptual schema. The second type of interruption comes from interferences in complex sequences of intended behaviour or goals.

Arousal then sets off cognitive interpretation. The interruption may be interpreted as positive or negative, which determines the quality of the feeling or emotion. For some types of judgments, primarily those involving perceptual schemata rather than goals, the degree of interruption determines the positivity of the response (Mandler 1982). Mandler suggests that the disconfirmation of perceptual schemata

can range from zero (total familiarity) to extreme (total discord). On the whole, Mandler suggests, confirmation is good, a little novelty is better, but total disconfirmation is unpleasant (cf. Berlyne 1970, Leventhal 1974). Mandler mainly analyses what he calls judgments of value, which pertain to one's taste in music, art, food, beauty and the like.

In summary, Mandler's model focuses primarily on interruptions of schemata as causes of affect. The interrupted schemata create arousal and cognitive interpretations. The nature of the interpretation determines the quality of affect, while the extent of the interruption and subsequent arousal determine its intensity.

#### 2.2.1.2 Social Schema and Affect

In the area of social psychology, schema processing has been proposed as a cause of affective response. Examples are the complexity of schema at a given point of time and the organization of the different attributes of a particular schema over time.

Linville (1982a, 1982b) proposes that the greater the complexity of a schema at a given point in time, the more moderate the affect it elicits. For example, the person "A" may know a dozen dimensions that determine the quality of a given football team, whereas the person "B" can think of only two. A's football knowledge is more complex than B's, which means that A's evaluation of any random team's

attributes will be more mixed. As B is integrating fewer attributes, any piece of information is likely to have considerable input on his judgment and make it more extreme. People differ considerably in the complexity of their perception of given domains of knowledge and in the complexity of other people (Fiske and Taylor 1984). In summary, Linville proposes that affect tends to be more extreme when the complexity of a schema is low, and more moderate when the complexity is high.

Tesser (1978) researched the development of schema complexity over time. He proposes that affective responses to the attributes of a particular schema become polarized over time. The result is a reduced perceived complexity of this schema. This means, over time thought processes lead to a tighter organization of the attributes of a given schema and polarize affect. For example, the longer one considers the chances of a football team for the league championship, the more one will force all of the team's attributes to fall into place as consistently pro or consistently con. Over time, the schema-caused affective responses become more extreme as the attributes of the schema become more organized.

### **2.2.1.3 Affect-Charged Schemata**

Fiske's (1981, 1982) work emphasises affect as a reaction triggered through the fit of new information into prior knowledge or schemata. In this theory, some schemata include affective attributes, called affect-laden or affect-charged schemata. Fiske has suggested that the assimilation of a stimulus to an affectively charged

schema may result in the schema's affect being conferred on the stimulus, and that individuals respond to the stimulus with the schema's affect. For example, people perceive the affect triggered by a stereotype to the extent an encountered person fits to that particular stereotypic mould. Thus, deciding that someone is a good example of a shallow playboy triggers the affective responses that are stored in the playboy schema (Fiske and Taylor 1984). If there is a poor fit to a stereotype, the affective reaction moderates. In summary, schema-triggered affect is an efficient affective processing device: one can say, "I know that type of person, and I know how I feel about him or her."

### **2.2.2 Appraisal-Triggered Affect**

Whether or not individuals organize a particular stimulus into a schema, they may assess its consequences, infer its causes and compare it with some standard. Affective responses are hypothesized to be outcomes of these processes. A brief overview of the underlying theories is provided in the sections below.

### **2.2.2.1 Assessment of Consequences**

One's affective response to a stimulus can be caused by the assessment of its consequences. This assessment includes consideration of whether the consequences will occur immediately or in the future, whether they will be minor or major and whether they will have short-term or long-term impact (Hoffman 1986). For example, an adult may respond with intense anxiety to a barely perceptible body ache because he or she thinks it might be a sign of cancer. The same body ache may elicit little affect in a young child whose only concern is the current degree of physical discomfort.

### **2.2.2.2 Attribution of Causes**

Weiner's (1980a, 1982) attribution theory focuses on obtained outcomes and their perceived causes. His theory holds that causal attributions to these outcomes shape affective responses. People are thought to make these attributions because they search for understanding and seek to discover why an event has occurred.

Weiner's work provides a systematic set of rules about the ways that causal attributions cause affect (cf. Roseman 1979). This set of rules is based on the following dimensions: (1) locus (internal or external), (2) stability (stable or unstable), and (3) control (controllable or uncontrollable). Locus and

controllability determine the specific nature of the affective reaction to outcomes, whereas stability determines the intensity of affective responses.

For example, an individual feels gratitude towards another person who chooses to help him/her. The perceived cause is external (the other person) and controllable (the other person chose to help). If the other person had been forced to help (external locus but not controllable), the individual helped would not feel grateful. Similarly, Hoffman (1982) found that people are more likely to respond sympathetically to another's distress, if its cause is perceived as beyond the other person's control. However, if the cause is perceived as being within the person's control, the sympathy may be diminished.

### **2.2.2.3 Comparison with Moral Standards**

When one compares a behaviour to some standard (generally to one's own or other people's moral standards) an affective response may be elicited. The affective response depends on whether the behaviour fits the standard, exceeds it, misses it or violates it (Hoffman 1986). The performance standards used may be absolute or relative and they may or may not be internalized. An internalized standard can be viewed as a subcategory of the self-system that is charged with affects relevant to self-evaluation (e.g. pride, shame, guilt).

For example, considerable research has suggested that people tend to experience guilt when their actions violate a moral standard. Furthermore, when people resist the temptation to act in a way that violates the standard, they often experience a positive feeling of self-approval or pride (Hoffman 1983).

## 2.3 Conclusions

In this section, research on the causal relationship between cognition and affect was reviewed. The theoretical dispute over whether cognitive processes are necessary antecedents of any affective response (e.g. Zajonc 1980 versus Lazarus 1982) was altered to a discussion of affective responses to cognitive processes with varying degrees of complexity. At one extreme of the complexity continuum are exceedingly simple and sometimes even unconscious cognitive processes, like the perception of physical or sensory features of stimuli. At the other extreme we find highly complex cognitive processes, such as schema processing and appraisals. The review showed that affect can be caused by cognitive processes of any degree of complexity.

Furthermore, affect caused by the more complex cognitive processes may be more powerful and enduring than affect elicited through the more primitive processes. Finally, the reviewed cognitive causal antecedents of affect are not mutually exclusive and may often operate together in various combinations, they may utilize

different aspects of the individual's sensory, perceptual and cognitive apparatus, and may operate at different levels of consciousness.

### **3 Summary**

In this chapter conceptualizations of affect were reviewed. As a result, Russell's (1980) circumplex model of affect was chosen to serve as the conceptual framework for affect in this thesis. This model conceptualizes affect as having pleasure/displeasure and aroused/sleepy as two orthogonal dimensions. According to this model, all other affective states can be defined along these two dimensions.

The review on the causes of affect showed that affective responses can be caused by cognitive processes of any degree of complexity. The spectrum ranges from exceedingly simple cognitive processes, like the unconscious perception of stimuli, to highly complex processes, such as the attribution of causes. More elaborate cognitive processes seem to have a more powerful and enduring impact on affect.

# Chapter V

## The Concept of Affect

### Applied to Satisfaction Theory with Services

This chapter focuses on the inclusion of affect in satisfaction models applied to services. In particular, it examines at a conceptual level whether affect, or the way people feel while consuming a service, can be included in the disconfirmation-of-expectations model. The role of affect in the consumption process could theoretically also apply to products. However, in Chapter I it was proposed that: (1) the experiential nature of services makes affect especially important for services, and (2) the service firm actually can manage the affective state of its customers, whereas the goods firm has little control over the consumption experience of its customers.

An introduction to the role of affect in consumer behaviour is provided in the first section. The second section reviews the literature on affect in the context of the disconfirmation-of-expectations model. Review of the affect literature in psychology suggests that Russell's (1980) model of affect is the best available conceptualization of affect. The third section examines whether Russell's model is appropriate for capturing the affective quality of service encounters. In the fourth section, four hypotheses on the inclusion of affect, as conceptualized by Russell, in disconfirmation-of-expectation models are developed. A summary is provided in the last section.

## 1 An Introduction to the Role of Affect in Consumer Behaviour

In the area of consumer behaviour, the vast majority of theorists focus their research on humans as cognitive beings. The outcomes are cognition based models including multiattribute models, attribution theory and comparison based satisfaction models. To the extent that affect is considered at all, it is viewed only as an intervening variable that helps to explain additional variance in consumer behaviour (Holbrook and O'Soughnessy 1984). Affect in all its complexity, richness and variety is replaced by some very narrow conceptualization. For example, multiattribute models only deal with one aspect of affect, namely affective judgment (degree of liking) or value component as in Fishbein-type multiattribute models (Fishbein and Ajzen 1975). This conceptualization of affect ignores the remainder of the affective spectrum (Holbrook 1986) and is too narrow (Peterson, Hoyer and Wilson 1986).

The review of the affect literature in cognitive, social and environmental psychology in Chapter IV showed that, since the early 1980's, many studies have been conducted to enable a better understanding of humans as cognitive-affective beings. Despite a slowly growing number of conceptual and empirical papers on consumer affect, very few of the results have been applied to research in consumer satisfaction. Instead, most of the attention on affect has been focused on its role in consumer prepurchase processes, such as preference formation (Zajonc 1986), perception of affective advertising (Batra 1986, Batra and Stayman 1990) and affect-laden information processing (Cacioppo, Losch, Tassinary and Petty 1986). Affect in

postpurchase processes has been neglected in both goods and services marketing (Westbrook 1987, Westbrook and Oliver 1991).

This neglect is surprising for a number of reasons. Firstly, the postpurchase period provides ample opportunities for a variety of affective responses which may be of considerable significance for consumers (Holbrook and Hirschman 1982). Secondly, affective variables might contribute substantially to the explanation and prediction of postpurchase behaviours such as satisfaction, word-of-mouth and repeat buying (Westbrook 1987). And finally, affect might be a useful tool to manage consumer satisfaction more effectively than is currently practised (Wirtz and Bateson 1990).

## **2 Affect as an Important Variable in Satisfaction Models**

Apart from a few researchers, satisfaction theorists have mostly disregarded affect during the consumption process. At most, they occasionally define the construct of satisfaction itself as encompassing some vaguely defined affective component. In this section, it is first examined whether confirmation/disconfirmation can be considered as a causal antecedent of the affective state of people while consuming a service, and second whether this affective state can drive post-consumption satisfaction.

## 2.1 Confirmation/Disconfirmation as a Direct Determinant of Affect

No research has addressed the relationship between confirmation/disconfirmation and the affective state of people during the consumption experience. The standard confirmation/disconfirmation model is essentially a model of cognitive processes and treats affective outcomes, if they are considered at all, as outcomes of the disconfirmation process occurring after consumption, or as rather vaguely defined parts of the satisfaction construct itself (Westbrook 1987, Woodruff et al 1983).

This is not surprising. The disconfirmation model is mainly applied to goods, where firms have virtually no control over the consumption process. As was outlined in Chapter I, this is different for services firms which can design and manage the service experience and, therefore, the affective state of people during consumption becomes important.

The psychology literature reviewed in Chapter III showed that cognitive processes of all degrees of complexity can cause affective responses. In particular, disconfirmation with schemata and moral standards have been shown to elicit affect. As disconfirmation-of-expectations is a process similar to disconfirmation of schemata, there is no reason to believe that it cannot cause affect. Furthermore, services are chosen and expected to deliver bundles of benefits to consumers. Disconfirmation of expectations usually means that service performance falls short of (exceeds) what a consumer expected when making a purchase decision with negative (positive) implications for the service experience. Therefore, that a service

performance exceeding expectations can cause positive affect, and a shortfall in performance can cause negative affect seems plausible. This view is intuitively appealing, however, it has not been tested yet.

In addition, the psychology literature on affect suggests that more complex cognitive processes have a stronger and more enduring impact on affect than simple cognitive processes. The disconfirmation-of-expectations process would belong to the complex cognitive processes reviewed in Chapter IV, and may therefore be a potentially powerful determinant of the affective state of consumers in service encounters.

## **2.2 Affect as Direct Causal Antecedent of Satisfaction**

In contrast to disconfirmation, whose relationship with affect has not been researched yet, few studies have addressed the relationship between affect and consumer satisfaction. Westbrook (1987) and Westbrook and Oliver (1991) conducted empirical studies to examine the impact of affect on consumer satisfaction. These two papers are reviewed in the following paragraphs. Furthermore, a number of researchers have employed affect as a mediating variable between stimuli in the service encounter and response behaviours (Bateson and Hui 1988, Donovan and Rossiter 1982, Hui 1988). In accordance with the Mehrabian and Russell (1974) model these researchers used approach-avoidance as the outcome variable, and not satisfaction. Their studies are reviewed from a consumer satisfaction point-of-view.

Westbrook's (1987) field study showed empirically that there is a direct causal link between affect and consumer satisfaction. In his study of cable-TV and automobiles, positive and negative affect influenced satisfaction outcomes significantly. The affective variables explained almost as much of the variance in consumer satisfaction (61% for satisfaction with cars and 48% with cable-TV services) as did the two cognitive variables (expectations and disconfirmation-of-expectations) together (72% for cars and 55% for cable-TV). As a result, Westbrook proposes that satisfaction is not only driven by the cognitive comparison process as commonly assumed, but also by consumption related affective experiences.

There are three major limitations of Westbrook's (1987) conceptualization of affect. First, he adopted the notion that affect arises as a function of the individual's cognitive appraisal or his/her evaluation of the meaning, causes, consequences and personal implications of a particular stimulus. This conceptual approach excludes many possible causes of affect, for example, mere exposure effects and innate preferences.

Second, Westbrook (1987) adopted Izard's (1977) measures of affect in a way that might have excluded relevant aspects of affective states. Westbrook only included primary affects that involve the attribution of causal agency for any consumption related outcomes to the product/service or its seller. They were interest, joy, anger, disgust, contempt and surprise. By doing this, Westbrook disregarded the other four basic affects determined by Izard (1977), namely distress, fear, shame and guilt. From a services marketing perspective, one can hardly consider these excluded

feelings as irrelevant. For example, a consumer who is not familiar with the appropriate behaviour or the service script of an upmarket French restaurant may feel distress and shame if he/she perceives himself/herself as behaving "stupidly" in the presence of his/her company, other customers and service employees. These two kinds of affect (distress and shame) would probably influence the consumer's level of perceived pleasure and arousal in the restaurant, although no causal attribution might be given to the service provider. In the service encounter, all interactions between the customer and the configuration of service employees, service setting, support material and other customers are part of the experience. Therefore, the whole spectrum of affective responses that might be caused by other individuals, objects and physiological stimuli should be included in a model that aims to capture the affective quality of service experiences.

Third, Westbrook applied Izard's (1977) conceptualization of affect. Izard's model, however, has no explicit intensity or arousal dimension like the Mehrabian and Russell (1974) and the Russell (1980) models. Izard's surprise scale seems to be too narrow to capture the concept of arousal.

Westbrook and Oliver (1991) also used Izard's (1977) conceptualization of affect to capture the emotional reactions of 125 car owners to their newly purchased automobiles. Five clusters of emotional responses were identified from the data. They were: (1) Happy/Content, (2) Pleasantly Surprised, (3) Unemotional, (4) Unpleasantly Surprised, and (5) Angry/Upset. An analysis of the relationship between the emotional response patterns and satisfaction clearly showed that the

level of satisfaction was the highest for the Happy/Content and Pleasantly Surprised customers, and progressively declined for the Unemotional, Unpleasantly Surprised and Angry/Upset groups. A high correlation between distinct emotional states and satisfaction furnishes further support for the hypothesized causal link between affect and satisfaction.

Several other researchers have applied the Mehrabian and Russell (1974) model to the service encounter. However, they have used approach/avoidance and not satisfaction as outcome variables. These researchers focused on the input variables and their impact on affect. Examples are Donovan and Rossiter (1982), who tested the impact of different service setting designs on affect, and Bateson and Hui (1988) and Hui (1988), who focused on perceived control, perceived choice and perceived crowding in their experimental designs. Although they focus primarily on the input side, all of these studies also show that affect can have a major impact on consumer behaviour in service settings.

As a result, Donovan and Rossiter (1982) suggest that cognitive influences such as price, location, variety and quality of merchandise (usually conceptualized through multiattribute models), may largely account for retail store selection. Within the store however, affective responses might be the primary determinants of consumer behaviour (Donovan and Rossiter 1982).

In summary, most researchers would agree that consumer behaviour results from the combination of two factors: affect and cognition. Westbrook (1987) and

Westbrook and Oliver (1991) provided empirical support for the hypothesis that a causal link exists between affect and satisfaction. Empirical studies conducted by Bateson and Hui (1988), Donovan and Rossiter (1982) and Hui (1988) show that consumer behaviour in service encounters is significantly determined not only by cognition but also by affect. The results of these five studies together provide strong support for the hypothesis that affect is an important causal antecedent of satisfaction with services.

### 3 Conceptualizations of Affect Applied to the Service Encounter

The psychology literature review suggested that Russell's (1980) conceptualization of affect, which is a further development of the Mehrabian and Russell (1974) model, has the best internal and external validity. Furthermore, its measures have the best discriminant validity, which poses a major problem for alternative models. The question now arises as to whether Russell's circumplex model of affect is appropriate in a services marketing context.

The original Mehrabian and Russell (1974) model was tested in 323 different environments (Russell et al 1981). Among those environments were a variety of service settings, e.g. a restaurant, a nightclub, a grocery store, a hairdresser's salon, a bank, an embassy waiting room, a computer centre, an airport, a museum, a racetrack, a city bus, a hospital and a child care centre. Most of the subjects were surrounded by other people, as in a shopping mall or at a hockey game. The events that occurred in or near those places that were assessed, ranged from a ballet, a rock concert to an antirape rally. It can therefore be concluded that Mehrabian and Russell's very wide conceptualization of environment includes service encounters with their interpersonal interactions. This should also apply to Russell's (1978, 1980) model, which was derived from the original Mehrabian and Russell (1974).

Furthermore, Russell and Pratt (1980) explicitly stated that their concept and scales can be used (1) to assess those objectively specifiable properties of environments and social situations which may influence affective reactions, (2) to predict and

explain behavioural reactions to specific environments and situations, (3) to assess the affective quality of specific environments and situations, and (4) to design and evaluate research by employing more precise specifications to affective variables which have so far only been vaguely defined and operationalized in environmental and social psychology. In this paper Russell and Pratt explicitly stated the appropriateness of their model in capturing environment-human interactions as well as social interactions. Since service encounters consist mainly of these interactions, it would seem that Russell's circumplex model is indeed a good conceptualization of affect in the context of service encounters.

A number of researchers have already applied Russell's (1978, 1980) circumplex model to capture the affective quality of service encounters. Examples are the studies of Bateson and Hui (1988), Donovan and Rossiter (1982) and Hui (1988)<sup>1</sup>. The cited researchers agree on the appropriateness of Russell's circumplex model of affect for modelling consumer behaviour in service encounters. Their view seems to be empirically well corroborated.

---

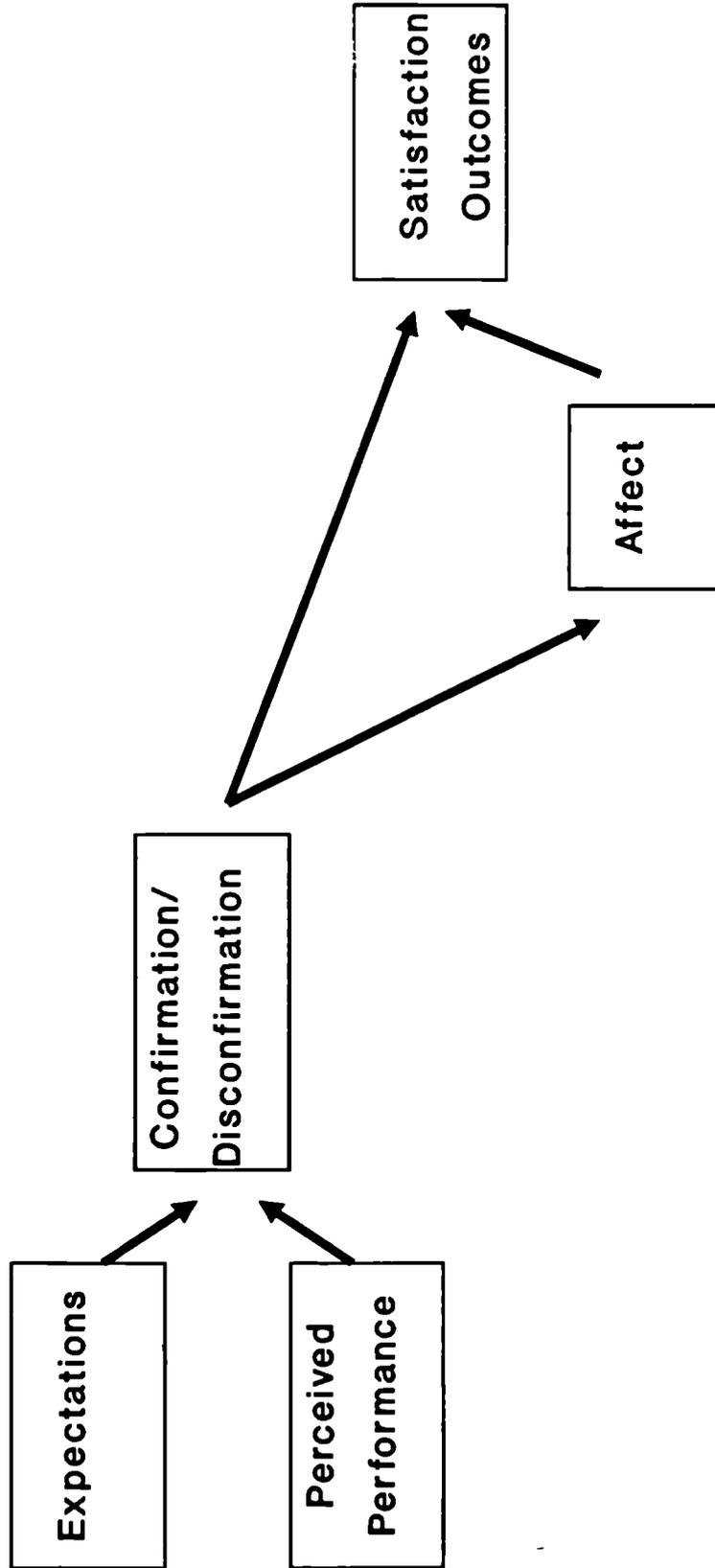
<sup>1</sup> A review of these studies was provided in Chapter I, Section 3.

#### **4 Inclusion of Affect in Disconfirmation Models - Hypotheses IV to VII**

Three conclusions were drawn in the previous sections. First, disconfirmation probably can drive the affective state of people during a consumption experience, however, this relationship still remains to be tested. Second, empirical evidence suggests that satisfaction is not only driven directly by the cognitive disconfirmation-of-expectations process but also directly by the affective state of people during the consumption experience (Westbrook 1987). These propositions are incorporated in the model shown in Figure 5-1. Third, it was concluded that Russell's (1980) model of affect is suitable for modelling the affective state of people while consuming a service.

So far, Russell's concept of affect has never been applied in the context of consumer satisfaction. Russell's two dimensions of affect (pleasure and arousal) offer a means of replacing commonly applied but broad and only vaguely defined categories of affect (e.g. positive and negative affect) with a concept of more precise specification which can be used for hypothesis development and testing (Russell 1980). In the following paragraphs, four hypotheses about the relationships between confirmation/disconfirmation, arousal, pleasure and satisfaction are developed.

**Figure 5-1: Incorporating Affect and Cognition  
in Satisfaction Models**



A first step in understanding the role of affect in the satisfaction process would be to confirm the proposed causal relationship between disconfirmation and "positive" affect or pleasure, as defined more precisely by Russell. This proposal that positive disconfirmation causes pleasure and negative disconfirmation displeasure has never been tested empirically. Therefore the following hypothesis is put forward.

**Hypothesis IV:**

**The degree of pleasure experienced in a service encounter is an increasing function of the perceived disconfirmation-of-expectations<sup>2</sup>).**

Mehrabian and Russell (1974) propose that the information rate of an environment or a situation directly drives the level of arousal of people. High information rates are suggested to cause high levels of arousal, and vice versa for low information rates. The researchers define information rate as the degree of novelty and complexity of the environment or situation. Novelty refers to the unexpected, the surprising, the new and the unfamiliar, whereas complexity refers to the number of elements and the extent of motion or change.

As noted in Chapter IV, Mandler (1977) suggests that arousal is produced by interruptions or unexpected events that alert the organism to cope with environmental contingencies. Specifically, arousal is suggested to be caused by

---

<sup>2</sup> Disconfirmation is defined as being negative for Expectations > Perceived Performance, and positive for Expectations < Perceived Performance.

disconfirmation of perceptual schemata or by interferences in sequences of intended behaviour or goals.

Applied to consumer satisfaction, both of these theories predict that disconfirmation-of-expectations produces higher levels of arousal than confirmation. Applying Mandler's (1977) theory, disconfirmation rather than confirmation can be considered as unexpected event, interruption of a service schema (or script as employed by Smith and Houston 1983) or perhaps even intended behaviour (customer participation in the production process) and goals (obtaining a specific bundle of benefits). According to this line of argument, disconfirmation causes arousal, whereas confirmation does not.

The information rate theory predicts the same outcomes as Mandler's (1977) theory. In particular, as disconfirmation-of-expectations experiences are more unexpected and surprising than confirmation experiences, the information rate theory would hold that disconfirmation situations have a higher information rate and cause higher levels of arousal than confirmation situations. Furthermore, the information rate theory proposes that an increased level of unexpectedness and surprise means an increased information rate which causes higher levels of arousal. On the basis of this discussion the following hypothesis is proposed.

**Hypothesis V:**

**The level of arousal experienced in a service encounter is an increasing function of the perceived magnitude of disconfirmation-of-expectations.**

Westbrook's (1987) and Westbrook and Oliver's (1991) studies clearly provided empirical support for the hypothesis that pleasure is a causal antecedent of satisfaction. In these two studies, Izard's (1977) conceptualization of affect was used. However, for conceptualizing pleasure in the service encounter, Mehrabian and Russell's (1974) model and its further development, the Russell (1980) model of affect, have generally been preferred. Unfortunately, neither of these two models has been used in the investigation of consumer satisfaction. It is hypothesized that Westbrook's (1987) and Westbrook and Oliver's (1991) results can be replicated with the Russell model of affect.

**Hypothesis VI:**

**Satisfaction is an increasing function of the pleasure experienced during the service consumption process.**

Arousal has generally been described as an "amplifier" of the impact of pleasure on behaviour (Mehrabian 1980). Mehrabian and Russell's (1974) empirical findings show that there is a significant pleasure-arousal interaction effect on approach-avoidance: in pleasant environments approach is a monotonic increasing function of arousal, whereas in unpleasant environments approach is a monotonic decreasing

function of arousal<sup>3</sup>). The same results were obtained in a study with retail settings (Donovan and Rossiter 1982).

In the context of consumer satisfaction, arousal is also commonly referred to as amplifier of the impact of pleasure on satisfaction (Oliver 1989, Westbrook 1987, Westbrook and Oliver 1991). The assumed relationship between arousal, pleasure and satisfaction is schematically shown in Figure 5-2. Again, this relationship has never been submitted to an empirical test. Therefore, Hypothesis VII is advanced.

**Hypothesis VII:**

**The strength of the impact of pleasure on satisfaction is an increasing function of the level of arousal experienced during the service consumption process.**

The model presented in Figure 5-3 summarizes Hypotheses IV, V, VI & VII.

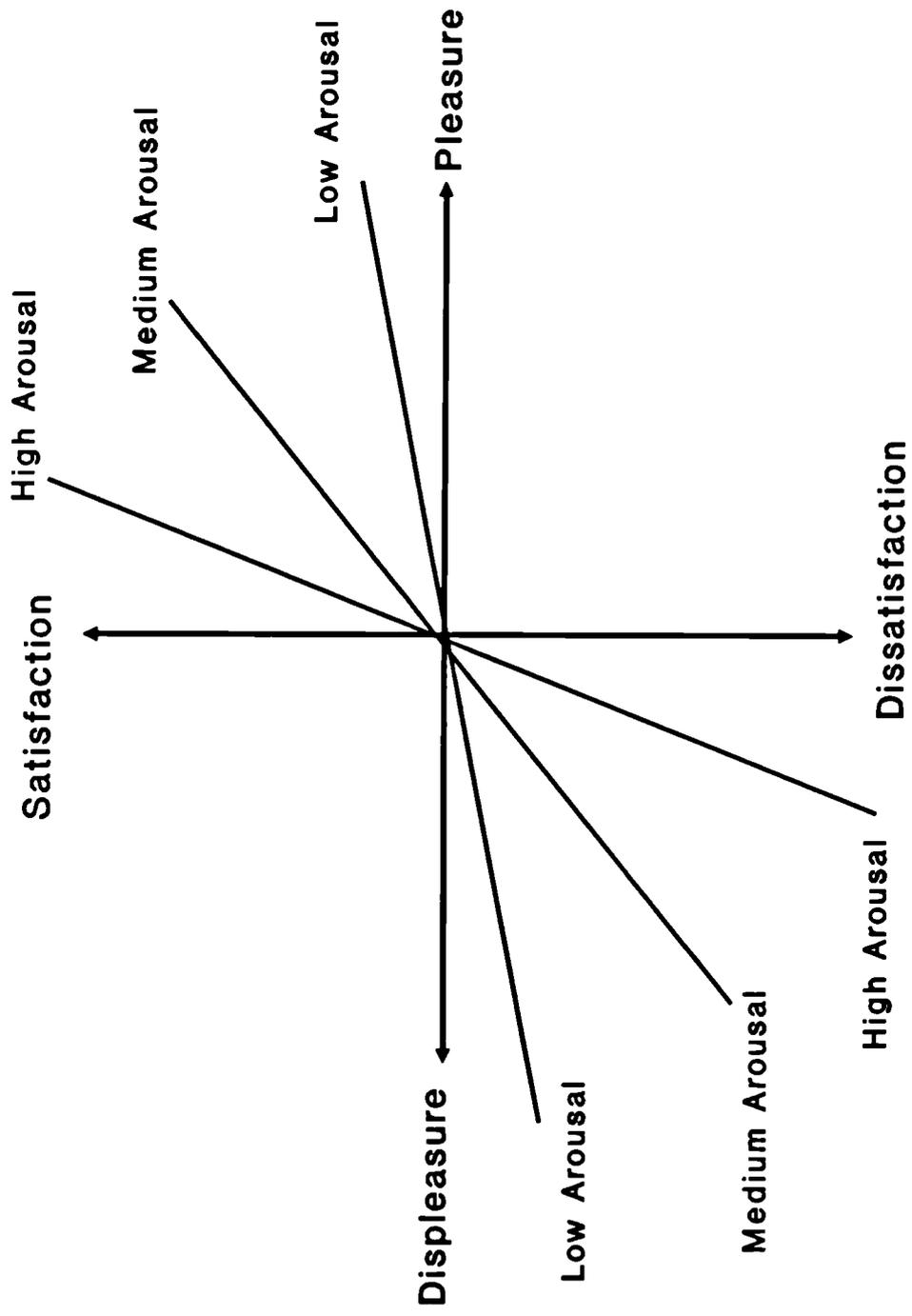
A direct causal link between disconfirmation and satisfaction is central to the classical disconfirmation-of-expectations model and has been supported empirically in a number of studies (e.g. Cadotte et al 1987, Oliver 1980a, Oliver and Bearden 1985, Oliver and DeSarbo 1988, Tse and Wilton 1988). This link was also

---

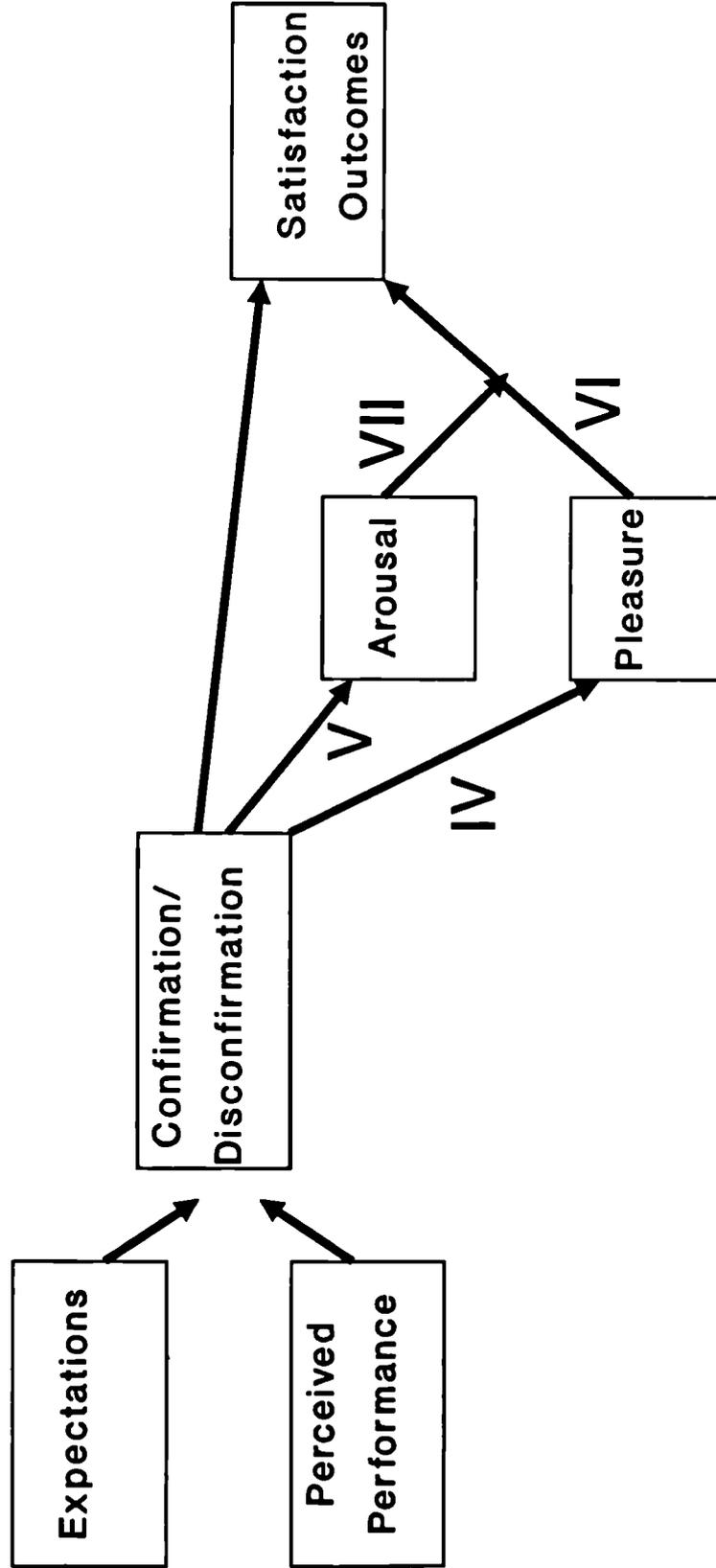
<sup>3</sup> Some caution needs to be raised for extremely low and extremely high levels of arousal, as the hypothesized impact on approach-avoidance can reach saturation and eventually reach a turning point (Mehrabian and Russell 1974, Mehrabian 1980). These situations were neglected, as it was not aimed to study extreme cases of affect.

confirmed in Westbrook's (1987) study on affect which showed that satisfaction was determined directly by both disconfirmation and affect. Therefore, a direct causal link between disconfirmation and satisfaction is also included in the model presented in Figure 5-3.

**Figure 5-2: The Hypothesized Relationship Between Pleasure, Arousal, and Satisfaction**



**Figure 5-3: Including Russell's (1980)  
Conceptualization of Affect in Satisfaction Models**



IV, V, VI and VII refer to the hypotheses developed in this chapter. All links without number are according to the classical disconfirmation-of-expectations model

## 5 Summary

Research reviewed in this chapter supports the use of Russell's circumplex model of affect in capturing the affective quality of service encounters.

In Chapter I it was suggested to include affect in satisfaction models to better capture the experiential nature of the service encounter. In this chapter, Hypotheses IV to VII were developed on the inclusion of affect (as conceptualized by Russell, 1980) in the classical disconfirmation-of-expectations model. Figures 5-2 and 5-3 summarize the hypotheses put forward.

# Chapter VI

## An Experimental Study: Methodological Issues and Procedures

The issues to be researched are (1) the impact of perceived performance on satisfaction, (2) the impact of heterogeneous performance expectations on disconfirmation, and (3) whether the disconfirmation-of-expectations model can be extended to include the affective state of people (as conceptualized by Russell 1980) while consuming a service. Specific hypotheses were developed in Chapters II, III and V respectively.

This chapter presents methodological considerations and the method used to test the hypotheses proposed. In the first section, four basic types of validity are discussed. In the second section the relative importance of these types of validity in relation to research objectives pursued is examined. In the third section, the three classical research designs and their strengths and weaknesses in relation to the four types of validity are reviewed. A true experimental design was chosen because of its superior internal validity which is crucial for testing causal relationships such as the ones examined in this thesis.

In the fourth section, the advantages and problems of laboratory and field settings are discussed. A laboratory setting was selected because it minimizes extraneous

sources of variation and thereby maximises internal validity. For the laboratory simulation a PC-based homebanking service was selected and expectations and performance on the system's response time were manipulated. There were three reasons for this choice. First, a machine-based service is highly controllable and replicable which is in contrast to people delivered services. Second, because a PC-based homebanking service had not been launched in the United Kingdom, subjects could be selected not to have previous experiences with this kind of service. This allowed the use of expectations as comparison standard and manipulate them plausibly. Third, response time is an important attribute of homebanking services for which value-free perceived performance measures can be used and which shows little ambiguity in contrast to many other service attributes. The importance of these features for the study is outlined in section four.

In the fifth section, three potential objectives that can be pursued when selecting subjects are discussed and the recruitment strategy for this study is presented. The details of the methodological procedures for the experiment conducted are presented in the sixth section. A summary is provided in the last section

## 1 Four Basic Types of Validity

A number of criteria for validity have been suggested to determine the extent to which research designs are useful in testing hypotheses. Cook and Campbell (1979) have identified four basic types of validity each associated with one of the four major questions which may concern any researcher. These questions are: (1) Is there a relationship between the variables of interest (statistical conclusion validity), (2) is that relationship plausibly causal from one operational variable to the other (internal validity), (3) what are the particular constructs involved (construct validity), and (4) how generalizable is this relationship across persons, settings and times? A short definition and discussion of each type of validity are given below.

Internal validity, without which any experiment is uninterpretable, addresses the question of whether an observed covariation can be considered as a causal relationship between two variables, or whether the absence of covariation implies the absence of causality. Therefore, in an ideal research design the covariation between two variables should be free from any influence of extraneous variables that might offer rival explanations for the observed variance between the variables of interest. The most commonly cited extraneous variables, also called threats to validity, are maturation, history, instrumentation, mortality and selection (Kidder and Judd 1986).

Maturation refers to the processes within the respondents operating as a function of time per se. That is, observed effects might not be due to the experimental treatment, but to respondents growing older, more experienced, hungrier and the like. History refers to rival hypotheses based on specific events other than the experimental treatment, occurring between pretest and post-test. Instrumentation is a threat when an effect might be due to a change in the measuring instrument or observer. For example, the observer might become more experienced in the course of an experiment and change the calibration of the measures applied. Mortality can be a rival hypothesis when an effect may be due to a differential loss of subjects from the comparison group. Selection refers to the threat that comparison groups differ and that any difference observed in the variables of interest could be explained by group differences rather than the experimental treatment. Campbell and Stanley (1963) and Cook and Campbell (1979) considered further potential threats to validity, particularly testing, statistical regression to the mean and interactions with selection (selection-maturation, selection-history and selection-instrumentation).

Construct validity refers to whether the operational variables or measures employed actually capture the theoretical constructs of interest. A necessary, but not sufficient condition for construct validity is that any employed measure should correspond perfectly to one, and only to one, theoretical construct at issue. All threats to construct validity have either to do with the measures failing to incorporate all of the dimensions of the target construct or with the measures containing dimensions irrelevant to this construct. Cook and Campbell (1979) call the first category of

threats "construct underrepresentation" and the second "surplus construct irrelevancies".

External validity considers whether an observed causal relationship can be generalized beyond the immediate research setting, population, measures and time. The extent to which a causal relationship can be generalized across conditions is in essence a test of statistical interactions. For example, only if we know that there is no interaction between treatment and age of respondents, can we say that the observed causal relationship holds across different age groups. Cook and Campbell (1979) suggest the following threats to external validity: (1) interaction of selection and treatment, (2) interaction of setting and treatment, and (3) interaction of history and treatment. Campbell and Stanley (1963) further suggest interactions of testing and treatment as potential threats to external validity particularly with experimental designs.

Statistical conclusion validity examines whether statistical inference about covariation between variables are justified. This encompasses the questions of whether the statistical techniques used and the sample size are appropriate, the appropriate magnitudes of the probabilities of Type I and II errors, and whether the obtained error variance has been inflated inadvertently by some systematic error. The common threats to statistical conclusion validity are: (1) low statistical power when sample sizes are too small, (2) violated assumptions of the statistical tests used in the analysis, (3) fishing and the error rate problem, and (4) the application of measures with low reliability.

## 2 The Relative Importance of the Types of Validity

Although all four types of validity are important in evaluating research, their relative importance is a function of the pursued research objective (Kidder and Judd 1986, p. 28). Calder, Phillips and Tybout (1981) identified two distinct types of research objectives in consumer research, which they termed effects application and theory application.

Effects application refers to obtaining findings that can be applied directly to a real-world situation of interest (Calder et al 1981, p. 198). An example of effects application would be a test of how a company's customers respond to alternative designs for a new product packaging. Obviously, effects application is the main objective of applied or problem-solving research (Myers, Greyser and Massy 1979). No meaningful prediction of the response of the company's customers can be derived from the results of a market study if the generalizability of the results to one's target population, shopping environment and shopping situation is low.

Accordingly, the key methodological concern for effects application is to ensure that all features of the real world application are represented in one's research in terms of the respondents employed, operationalizations of variables, research setting and research design. This minimization of artificiality (Hui 1988) has been termed a correspondence procedure by Calder et al (1981). Specifically, the selected sample should be representative of the real-world population. Convenience samples like students or the general public are obviously inappropriate for effects application.

The research setting should be as realistic as possible. This means that a field setting is preferable to a laboratory setting. Variables should be operationalized in the most naturalistic way possible (e.g. unobtrusive observation). Finally, even for the examination of causal relationships, although unlikely when effects application is the main interest, surveys or quasi experiments may be more appropriate than true experiments, because artificial settings and research procedures can often not be avoided in true experimental designs (Cook and Campbell 1979, Kidder and Judd 1986).

The primary goal of theory application is to "identify scientific theories that provide a general understanding of the real-world" (Calder et al 1981, p. 198). Theory application requires falsification test procedures that have the potential to refute a theory. Theories that survive rigorous efforts of falsification are accepted. It follows that the specific context and effects are not the focus of the research, rather their significance lies in the information they furnish about the adequacy of the theory (Calder et al 1981). It is not the aim to generalize any particular outcome observed, but to generalize the theoretical relationship between the variables employed.

Accordingly, the key methodological concern for theory application is minimizing the possibility that third variables have impact on any observed relationship between independent and dependent variables. In other words, internal, construct and statistical conclusion validity should be maximized. External validity is only of marginal importance, because most theories are context free and the refutation of

a theory in any context would anyway demonstrate the need to reformulate or abandon the theory.

In summary, when effects application is the pursued research objective, external validity should be the primary concern of the researcher. On the other hand, with the aim of theory application, the researcher has to focus on internal, construct and statistical conclusion validity. Here, external validity is not a primary concern for any single research study (Calder et al 1982, 1983).

#### Research Objectives of this Study

All of the objectives of this study are concerned with theory rather than effects application, and the hypothesized relationships are causal in nature. This means that the emphasis on the different types of validity should be in the following descending order of importance: internal, construct, statistical conclusion, and finally, external validity.

### **3 Research Designs**

In the previous two sections, the criteria for evaluating research and their relationships to the pursued objectives were discussed. In this section three classical research designs and their implications for the different kinds of validity are discussed. The most suitable design for this study is then determined.

#### **3.1 True Experiments**

True experimental designs are characterised by at least two treatment conditions (two-group designs) with random assignment of subjects to the experimental conditions. Because of the presence of at least two treatment conditions (either a treatment and no-treatment, or a treatment one and treatment two condition) the researcher can compare the outcomes and test the null-hypothesis that there is no difference between the outcomes of the two groups.

The true experiment, when conducted in a highly controllable environment such as a laboratory, allows for holding most extraneous variables constant and manipulating only the independent variables of interest. Even subject or organismic variables, such as religion, race, education and the like, can be controlled for by selecting only subjects with the desired criteria. However, there is potentially an indefinite number of subject variables (Calder et al 1983) that might impact on the outcome variables and offer rival explanations to the treatment as causal

antecedents of the observed effects. A random assignment of subjects controls for these subject variables by equalizing two or more groups before an experimental treatment begins.

With a carefully designed true experiment the internal validity is maximized by ruling out its primary threats: maturation, history, instrumentation, mortality and selection. Furthermore, true experiments allow for purpose-specific operationalizations which facilitate high construct validity. The minimization of systematic error also allows for high statistical conclusion validity.

On the other hand, true experiments have been extensively criticized for their low external validity. All extraneous variables not of interest to the researcher are held constant. They are, however, part of the real world situation. The artificial environment created by the experimental procedure and the frequent operationalization of distinct theoretical variables that cannot be observed in the real world cast further doubt on the external validity of such experiments.

In summary, a carefully designed true experiment can maximize internal, construct and statistical conclusion validity. This makes randomized experiments the method par excellence for examining causal relationships (Cook and Campbell 1979, Kidder and Judd 1986) or for theory application (Calder et al 1981). However, the generally low level of correspondence between experimental conditions and the real world results in low external validity (Calder et al 1981). For any effects

observed in an experiment, the generalizability beyond the immediate setting and sample is questionable.

### 3.2 Quasi Experiments

Quasi experiments are similar to true experiments. Quasi experiments, unlike true experiments, do not have randomly assigned treatment and comparison groups. This means that the comparison between different treatment and/or non-treatment conditions has always to be made with nonequivalent groups or with the same subjects prior to the treatment. This lack of control on what happens to whom (nonequivalent groups) introduces a number of threats to internal validity.

For example, history effects are a problem for interrupted time-series and regression-discontinuity designs. Selection by maturation interaction is a threat to internal validity in regression-discontinuity and pretest-posttest nonequivalent control group designs. Construct validity is usually lower than in true experiments, because tailored and multiple measures are more difficult to apply in quasi experiments. Statistical conclusion validity is also lower because the error variance is inflated by the effects of a potentially greater number of background factors that are not controlled for in quasi experimental designs.

Experiments appear to be the best methods for researching causal relationships because they have higher internal validity than quasi experiments (Kidder and Judd

1986). If, however, random assignments cannot be implemented, a quasi experiment is the second-best choice, and with careful design and the collection of many data points, most threats to internal validity can be avoided. Quasi experiments can have higher levels of external validity than true experiments, because quasi experiments can be designed to use real-world settings, processes and situations, and to be unintrusive. Furthermore, guinea-pig effects can be avoided because no random assignment is made.

In summary, quasi experiments offer a trade-off between the high internal, construct and statistical conclusion validity and the low external validity of true experiments. Quasi experiments show lower internal, but significantly higher external validity than true experiments. This makes possible the use of quasi experiments for theory application research when a true experimental design cannot be used. Quasi experiments are also suitable for effects application research, because they allow for a far better implementation of correspondence procedures than true experiments.

### 3.3 Surveys

Usually, surveys are carried out to assess the relative incidence, distribution, correlations and even magnitude of correlations of naturally occurring phenomena in a target population. With these objectives, the main concern is external validity, which can be maximized with carefully designed surveys (Kidder and Judd 1986). High levels of external validity require the study populations to be representative of the entire population to which it is to be generalized, and the accurate measurements of all variables or phenomena of concern.

Survey designs are subject to threats to internal validity, because any interpretation in terms of causality has to rely on correlational data only. In other words, survey designs are close to useless for researching causality between variables (Kidder and Judd 1986). An exception to this rule is that some survey designs allow for discrimination among competing causal hypotheses through the process of controlling for third variables which may induce spurious correlations. However, the success of this process depends on whether the right variables have been chosen in the first place. Ultimately successful interpretation of survey data depends on the strength of the theory that guides the research. In other words, for any causal interpretation, survey designs need well developed bodies of theory for determining potential causes before the data are collected.

In spite of the potential threats to internal validity, surveys are still used even for theory application research because many research problems cannot reasonably

be dealt with by true or quasi experiments. Questions like the impact of education on the vocational success of product managers, or the lifetime effects of being raised in poverty versus wealth, cannot reasonably be studied by research techniques other than surveys.

In summary, survey designs can achieve a maximum of external validity, which seems to be ideal for effects application research. On the other hand, survey designs generally cannot answer any questions of causality because of the many potential threats to their internal validity and are therefore the least suitable design for theory application research.

### **3.4 Selected Research Design**

The selected research design should be determined by the research objective. The objective of this study is theory testing or the identification of scientific theories that provide a general understanding of the real world (Calder et al 1981). It is the theoretical explanation or theoretical relationship between variables that is aimed to be generalized and not the particular effects obtained, which is important in problem-solving research (Calder et al 1981).

There is a consensus that the true experiment is more appropriate than other methods for investigation of causal relationships because it minimizes the possibility that a third variable causes any observed correlation between independent and

dependent variables. Experimental designs rank higher on internal validity than any other design (Calder et al 1981, 1982, 1983; Campbell and Stanly 1963; Cook and Campbell 1979). As the primary goal of the study is theory testing and the hypothesized relationships are causal in nature, an experimental design with random assignment of subjects to experimental conditions was chosen.

To provide a context for the choice of the research design for this study, Table 6-1 lists the major empirical studies on consumer satisfaction which use disconfirmation and/or affect. This overview shows that no true experimental design had been used to examine the role of affect in satisfaction processes. The choice of a true experiment for this study provides a rigorous test of the relationship between affect and satisfaction investigated by Westbrook (1987) and Westbrook and Oliver (1991) using survey methods (for a review of these surveys refer to Chapter V, Section 1).

Apart from Churchill and Surprenant's (1982) experiment, which tested the disconfirmation-of-expectations model, no other experiment has been conducted with a substantial number of subjects ( $n > 100$ ).

**Table 6-1: Major Empirical Studies of Consumer Satisfaction Using Disconfirmation-of-Expectations Models and/or Affect**

Study/Reference	Subject of Study	Research Design	Research Setting	Subjects	Product/Service
Cadotte, Woodruff and Jenkins (1987)	Comparing expected brand performance, best brand norm and product norm as standard for comparison	survey	field	87 patrons of fast food restaurants	fast food
Churchill and Surprenant (1982)	Empirical test of the disconfirmation-of-expectations model	experiment, manipulation of expectations and performance	laboratory	200 patrons of a shopping mall	video disk player and hybrid plant
Oliver (1980a)	Test of the disconfirmation-of-expectations model	survey	field	356 residents of a US town, and 248 students	federal flu vaccination program
Oliver and Bearden (1985)	Comparison of the role of subjective and subtractive disconfirmation in the satisfaction process	survey	field	353 members of a consumer panel	appetite suppressant
					cont...

Study/Reference	Subject of Study	Research Design	Research Setting	Subjects	Product/Service
Oliver and DeSarbo (1988)	Comparison of equity and disconfirmation models, individual differences	experiment; manipulation of attributions, expectations, performance, disconfirmation and equity	laboratory	40 students	stock market investment via a broker
Westbrook (1987)	Examination of the relationship between affect and satisfaction, complaint behaviour and word-of-mouth	survey	field	200 car owners, and 154 heads of households	automobiles, pay-TV services on cable
Westbrook and Reilly (1983)	Test of the value precept disparity standard	survey	field	72 students	automobiles
Westbrook and Oliver (1991)	Examination of the relationship between affect and satisfaction	survey	field	125 owners of recently purchased cars	automobiles
Tse and Wilton (1988)	Comparison of alternative comparison standards in disconfirmation models	experiment, manipulation of expectations and performance	laboratory	62 students	portable CD-player

## **4 Research Settings**

The selection of a research setting has major implications for the different types of validity. In the first section, the two basic types of research settings are presented and their implications for internal and external validity are discussed. In the second section, the chosen research setting is described and the reasons for simulating a PC-based homebanking service are outlined.

### **4.1 Trade-Offs between the Field and Laboratory Setting**

As in all other aspects of a study, the choice of the research setting depends on the objectives of the research. When effects application is the goal, correspondence procedures have to assure a maximum similarity between the research setting and the real world situation of interest. Usually, the real-world situation is heterogeneous in a number of background factors and this has to be reflected in the research setting. Ideally, a random sample of background factors would be employed but, if this is not feasible, the researcher may try to identify and include those factors that are most likely to affect the constructs of interest. However, regardless of the procedure used for treating setting heterogeneity, external validity is best achieved through a field setting (Calder et al 1981).

One major drawback of field settings is that they introduce many threats to the internal validity of the study and make the examination of causal relationships very

difficult. This is because the experimenters lose control of many background factors whose effects on the independent variables are often unknown.

Theory application requires the selection of an entirely different research setting. To test a theory, its constructs must be tied to a particular set of measures in specific circumstances. Here it is not important that these circumstances be representative of other settings. Rather, the circumstances are only important for operationalization of the theory. To test causality between variables, the research setting should be free of extraneous sources of variation, which increase the risk of Type II errors. The best procedure for minimizing the number of random irrelevancies is to employ a controlled laboratory setting. In contrast to field settings, laboratory settings facilitate the use of standardized procedures and treatment implementation, and allow the researcher to control rigorously the stimuli impinging upon subjects (Calder et al 1981). Laboratories also allow greater latitude in tailoring operationalizations to the constructs they are meant to represent. Furthermore, laboratory settings offer a greater potential for employing multiple measures.

In summary, laboratory settings usually offer high internal and low external validity. This makes them the first choice for theory application research. Field settings, on the other hand have, in general, high external and low internal validity. Their high external validity makes them ideally suited for effects application research.

## 4.2 A Laboratory Setting is Chosen and a Homebanking Service Simulated

Again, because the objective of this study is theory testing, a laboratory setting was selected as it offered a higher level of internal validity than a field setting. A PC-based homebanking service was used, and expectations and performance of its response time were manipulated. A PC-based homebanking service allows consumers to handle most of their bank transactions from their home or office. All that is required is a PC with a piece of software installed on its hard disk (the software is provided by the bank) and a modem that links the PC to the bank's mainframe via the customer's telephone. Typical transactions available are: obtaining account details (e.g. account balance, credit limit, interest accrued), bill payments, inter-account transfers (e.g. from chequing account to investment account), and ordering cheque books and account statements.

There were three reasons for selecting a PC-based homebanking service and manipulating response time expectations and performances. First, a machine-based service is highly controllable and replicable which is in contrast to people delivered services. Second, because a PC-based homebanking service had not been launched in the United Kingdom, subjects could be selected not to have previous experiences with this kind of service. This allowed the use of expectations as comparison standard and manipulate them plausibly. Third, response time is an important attribute of homebanking services for which value-free perceived performance measures can be used and which shows little ambiguity in contrast to many other

service attributes. The importance of these features for the study is outlined in the following subsections.

#### **4.2.1 Performance of a Homebanking Service is Controllable and Replicable**

Experimentation with a PC-based homebanking service allows maximum control over independent and extraneous variables. In particular, all attributes (with the exception of the independent variables) can be held constant. This is a unique feature of machine delivered services. In situations involving interactions between people (e.g. subject-experimenter or consumer-service employee interactions) it is far more difficult, if not impossible, to hold extraneous variables constant and to manipulate independent variables at the same time.

This difficulty is illustrated in the following example. Fisher et al (1976) conducted a field experiment to examine the impact of a short touch of hands on the customer's affective state and his/her evaluation of a library clerk. To ensure that all effects observed could be attributed to the manipulated touch-of-hands conditions, every other background factor had to be either controlled for or randomized. This meant that the library clerk could not add an eye contact or a smile to the touch of hands which would provide a rival explanation for the effects observed. To control for or randomize all such potentially relevant aspects of the behaviour of a person, here the library clerk, in an encounter with another person is exceedingly difficult.

Experimental investigation of a service encounter requires replication over many experimental sessions. In these circumstances problems arise from heterogeneity in performance between people and by the same person over time. A machine based service has the advantage of delivering exactly the same performance over many experimental sessions.

#### **4.2.2 Subjects Not Having Previous Experiences with Homebanking Could be Selected**

The discussion about alternative comparison standards highlighted the limitations of expectations as standard (Chapter II, Section 1.1.1). In particular, expectation based standards disregard consumers' needs and wants as well as their previous consumption experiences. However, there is reason to expect that, at least in some situations, needs and wants and previous consumption experiences influence the comparison standards used in disconfirmation processes (Cadotte et al 1987, Westbrook and Reilly 1983).

In Chapter II it was concluded that, for expectations to be an appropriate comparison standard, consumers have to be able to choose freely and purposefully so that they are able to match their expectations with needs and wants as limited by what they think can be realistically expected. In an experimental design with random assignment subjects obviously cannot match expectations and limited needs and wants. To be able to use expectations as a comparison standard and to permit

experimental manipulation of these expectations, subjects should not have had previous experience with the service. Then, the problem of congruence between expectations and needs/wants as limited by previous experiences can be overcome and expectations can be manipulated plausibly.

The subjects selected did not have previous experience with PC-based or television set-based homebanking services. At the time of data collection, PC-based services had not yet been launched, and television-set based services (e.g. HOBS by the Bank of Scotland) had not been actively marketed and had gained very few users in the United Kingdom.

#### **4.2.3 Time is an Important Attribute of Homebanking With Value-Free Perceived Performance Measures and Low Ambiguity**

There are three advantages for manipulating response time expectations and performance of a PC-based homebanking service.

First, response time is important and was therefore assumed to initiate confirmation and disconfirmation processes which were necessary for testing the hypotheses advanced. The amount of time involved in a service transaction has been found to be a crucial dimension along which consumers appear to make service purchase decisions (Bateson 1985b) and assess service quality (Parasuraman et al 1988).

In the context of the human-computer interface time, or more precisely, system response time, plays a critical role. Brown (1988) argues that "adequate system response time is critical to user performance, accuracy, productivity and satisfaction with the computer system" (p. 115). In the context of homebanking the system response time assumes even greater importance because telecommunication costs are directly related to time on line. Communication with the bank's mainframe is usually via the customer's telephone line and charges appear on the customer's telephone bill.

Second, the perception of response time performance can be measured with value-free measures (in time units). This is in contrast to many other service attributes for which value-free measures do not exist (e.g. for pleasantness of the service environment and friendliness of service employees). The use of value free perceived performance measures was crucial for testing Hypothesis I (advanced in Chapter II, Section 3) which proposed that there is no direct causal link between perceived performance and satisfaction. It was proposed that the high positive correlations observed in a number of studies stem from perceived performance measures with evaluative components that capture a part of the satisfaction construct (refer to Chapter II, Section 1.1.2).

Third, the response time of a PC-based homebanking service shows little ambiguity in contrast to many other service attributes. The significance of this feature is discussed in the following paragraphs.

Services are physically intangible (they cannot be touched) and mentally intangible (it is frequently difficult to envision what has been obtained after having consumed a service) (Bateson 1977, Levitt 1981). These characteristics lead consumers to rely upon cues associated with the service production process, the service worker and the service environment when evaluating service quality and making satisfaction decisions (Mills and Moberg 1982). It is suggested that these cues are prone to subjective evaluation arising from ambiguity, in the sense that the evidence is consistent with more than one hypothesis. That is, ambiguous attributes are perceived and evaluated by the consumer, but they can be perceived in more than one way. When evidence is ambiguous, people tend to encode information in a way consistent with prior knowledge and expectations (e.g. Herr, Sherman and Fazio 1982). Applying this to satisfaction processes with ambiguous attributes, consumers revise their perceptions of attribute performance to more closely match their pre-experience expectations (Anderson 1973, Oliver 1980a, Olson and Dover 1979). The outcome of this process is a consistency effect (see Chapter II, Section 1.1.3).

For testing the effect of expected variance in performance on disconfirmation (Hypotheses II & III) and for testing the impact of disconfirmation on affect (Hypotheses IV to VII), it was necessary that disconfirmation situations could be created. Therefore, it was important not to manipulate the performance of highly ambiguous attributes that might cause strong consistency effects with expectations and cause subjects to experience confirmation regardless the objective level of performance. Ambiguity would also make it difficult to manipulate expectations

on variability of performance because perception of variability implicitly requires perception of differences between different levels of performance.

Time performance had the potential to be manipulated in a way so that strong consistency effects could be avoided and that different levels of performance heterogeneity could be expected by subjects. Now, it was important to know how people perceive variability and time intervals. A review of literature on the perception of variability is presented in Chapter III.

Hawes (1979) presented a thorough review on time perception and time as an important variable in consumer behaviour. He came to the conclusion that apart from variations in objective time, a number of situational, demographic and psychological variables can influence the perceived length of a time interval (Hawes 1979). Examples of situational variables are: the extent of attention paid to the interval, the way in which the interval is occupied, the number of stimuli filling the interval. Examples of demographic and psychological variables are: sex, age, marital status, education, income, attitudes and personality characteristics.

The influence of situational variables on time perceptions can be controlled for in an experimental design. In the experiments for this thesis the environment and the time-interval filling activities were kept constant for all subjects. The large number of demographic and psychological variables and the large sample size did not allow control by sample selection. Instead, randomised subject allocation to the experimental conditions was used to control for the effects of these variables

on subjective time perception and hence on the dependent variables in the experimental design.

Furthermore, the experiment was designed to cue subjects to a more objective perception. In manipulating time performance expectations, the subjects were first informed about the time performance measured in seconds. They were then made to experience this time performance (for a detailed description of the experimental procedure refer to Section 6).

Finally, pretests assured that the response time manipulations could be clearly distinguished and that major consistency effects between expected and perceived response time could be avoided.

## 5 Subjects

In this section, three subject selection strategies appropriate to each of three general research objectives are discussed. Then, the strategy for subject selection and recruitment for this study is determined in accordance with its specific objectives.

### 5.1 Three Potential Objectives and Subject Selection

Research may be directed toward generalization of results to a specified population. Here, the key concern is achieving a maximum external validity by using samples that are statistically representative of the real-world of interest. Suitable sampling procedures include nonprobabilistic and probabilistic procedures (Kidder and Judd 1986).

A second objective might be study of the impact of specific subject characteristics on dependent variables. Here, the researcher would purposefully sample individuals who vary on those characteristics. He or she would not primarily be concerned with the representativeness of the sample. For determining those characteristics, a well developed underlying theoretical framework is necessary, because theoretically there is an indefinite number of dimensions that could impact upon the independent variables.

A third objective, theory application, only requires a sample that permits falsification of the theory. Although any sample in the theory's domain can potentially falsify the theory, homogeneous samples are preferred for two reasons (Calder et al 1979). First, homogeneous respondents permit more exact theoretical predictions than may be possible with heterogeneous samples. Homogeneous samples may also allow more specific and narrowly defined independent variables and measures. Both features make failure of the theory easier to detect. Second, homogeneous samples increase the statistical conclusion validity by minimizing the error variance introduced by irrelevant respondent characteristics. In other words, by selecting homogeneous samples along the characteristics believed to inflate error variance, these sources of error can be controlled and the likelihood of Type II error reduced.

## **5.2 Subject Selection and Recruitment**

The objective of this study is consumer satisfaction theory application. Therefore, a sample that allows a test of the theory is needed. A representative sample is not required, because statistical generalization of the findings to the population of potential homebanking users is not intended. As discussed earlier, an ideal sample would be homogeneous on dimensions that potentially inflate the error variances in outcome variables. However, there is no reliable theory available in the domain of satisfaction theory that would allow determination of these dimensions. Some experiment-specific characteristics, such as computer literacy or previous experience with on-line services, can only be guessed.

As a minimum requirement, the sample has to allow the falsification of the theory. In the context of this study, this means that the subjects have to be able to quickly understand and operate a homebanking system with which they had no previous experience. This requires PC and menu driven software experience as well as basic knowledge about current accounts and automatic teller machines (ATM). In consequence, subjects were selected who had a current account, an ATM card and PC experience. A primary source for subjects was the London Business School. Participants of executive, senior executive, part-time and full-time MBA courses, as well as London Business School staff (primarily library and secretarial staff) were asked to "participate in a study designed to pre-test a new PC-based homebanking service that is to be launched in the UK".<sup>1)</sup> White collar workers were considered more suitable for this experiment than were undergraduates or the general public. This is because white collar workers are more likely to have PC-experience and a good understanding of the concepts of homebanking. As remuneration, all participants were either offered £10 in cash or food and drinks, which they were told had been provided by the launching bank.

---

<sup>1</sup> For a sample leaflet in the recruitment of subjects refer to Appendix A.

## **6 Methodological Procedures**

The design of the homebanking system is outlined in the next subsection, and the experimental sessions are described in the second. The third and fourth subsections describe specific aspects of the experiment: reason, system description and instructions given to subjects, the manipulation of independent variables and measures employed.

### **6.1 Design and Production of the Homebanking Service System**

The following paragraphs give a short description of a television set based homebanking service called Home and Office Banking Service (HOBS) offered by the Bank of Scotland. This service was the first of its kind in the United Kingdom and served as the basis of the design of an up-to-date PC-based homebanking service used in the experiment.

HOBS uses a fairly simple keyboard, costing about £95, which connects a television set via a modem with a telephone line. At the press of a button, the modem dials up to the bank's mainframe and a welcome message appears on the television screen. After entering a user number and password, it is possible to view all account balances (including investment and money market accounts), scroll back through statements, transfer money between accounts, and effect the payment of bills. At

the end of each transaction, the computer sends the details to the screen and asks for confirmation.

HOBS uses a system called Viewdata, which displays large characters on the television screen. This means that rather little information (only a quarter of the 25 lines of 80 characters that PC users are accustomed to) is provided per screen. Communication with HOBS, which was designed in the early 1970s, is quite slow. HOBS has an irritating delay between entering of a figure or character and its appearance on the screen. After every data input, the user has to wait while the system accesses the Bank of Scotland's mainframe and sends a new screen of information down the telephone line (the complete Bank of Scotland graphic is sent unnecessarily each time). As one tester reported: "Sometimes, you could kick HOBS. It knows nothing of the jumps, short-cuts and elisions that make modern programmes quick and fun to use" (Wilkinson 1990). A new version tailored to work with personal computers is expected to be launched later in 1991.

One of the main benefits of HOBS is the possibility of active cash management. For example, a customer can keep most of his/her cash in an interest bearing money market account, leaving the current account balance near zero. The money market account allows withdrawals at any time without interest rate penalty, provided the withdrawal is at least £250. Whenever the current account balance becomes too low, it can be increased by transfer from the money market account. It is even possible to write substantial cheques against an empty account provided one remembers to transfer the money via HOBS soon afterwards. Deadlines for

payments can be exploited efficiently. HOBS can be instructed to pay a bill on a predetermined day. For example the outstanding balance in a credit card account can be paid on the day before it is due, without the cardholder having to remember that day, because the date can be stored in the system and the transaction executed automatically.

Another important benefit is that the user can avoid trips to the bank, since most standard transactions can be handled and most information obtained from the user's home or office.

The Bank of Scotland's television set based HOBS service served as the basis for the development of an up-to-date PC-based homebanking service. A programmer wrote a piece of software in Turbo-Pascal that was installed on a PC's harddisk to simulate a PC-based homebanking service. All transactions available on HOBS were programmed. Apart from the response time, which could be manipulated, the simulated system performed as any modern menu driven PC-software. For building up the screens Hewlett Packard's graphics drive was used to ensure a modern and professional appearance. Some screens of the transactions performed by the subjects are enclosed in Appendix E.

## **6.2 Experimental Sessions**

19 experimental sessions were conducted in the Hewlett Packard Teaching Centre (Room number E202) at the London Business School. In each session four to twelve subjects were exposed to the experimental conditions. An overview of the elements of the experiment and the chronological order of all actions is provided in Table 6-2. The manipulations and measures are described in detail in the remainder of this chapter.

Table 6-2: Chronological Overview of an Experimental Session

No	Action	Reference
1)	Welcome the subjects, Presentation about this study on homebanking	For the full text refer to Appendices B and C
2)	Subjects read instruction for the study	Page 1 of the questionnaire, Appendix D
3)	Subjects read information on the homebanking system (manipulation of expectations)	Page 2 of the questionnaire, Appendix D
4)	Subjects conduct transactions on the "Simulator" (reinforcing expectation manipulations and becoming familiar with the system)	Page 3 and 4 of the questionnaire, Appendix D
5)	Manipulation check of expectations	Page 5 of the questionnaire, Appendix D
6)	Subjects conduct transactions on the "real system" (manipulation of performance)	Page 6 and 7 of the questionnaire, Appendix D
7)	Manipulation check of performance	Page 8 of the questionnaire, Appendix D
8)	Measurement of disconfirmation, affect and satisfaction	Page 8 to 13 of the questionnaire, Appendix D
9)	Measurement of the personal background of all subjects	Page 14 and 15 of the questionnaire, Appendix D
10)	Subjects are thanked, receive their remuneration and leave the room	

### 6.3 Reason, System Description and Instructions Given to Subjects.

At the beginning of each experimental session the instructor gave a short presentation (for the entire text refer to Appendix B). In order to avoid experimental artifacts, the subjects were told that an American bank was planning to introduce a PC-based homebanking service to the British market. The applications offered and equipment needed were described briefly. Subjects were told that the objective of the experiment was to examine potential user reactions to the different features offered by the system and its performance.

A description of a potential customer, Mr. Smith, was presented on a slide (see Appendix C). The subjects were asked to imagine themselves as being Mr. Smith while testing the service. They were asked to judge the feelings of Mr. Smith, rather than their own, when answering the questionnaire.

A similar approach was adopted by Havlena and Holbrook (1986) in their study of consumption experiences. The authors argue that this approach has two advantages. First, it provides subjects with a projective task and thereby discourages social desirability effects, and second, it avoids problems involving individual differences in reactions to specific types of activities. This approach has also been adopted by Hui (1988), who encountered some discrepancies with subjects' consumption habits and his experiment design. For example, in a pre-test subjects had been asked to choose a human teller when a cash machine was available. A subject noted: "When the machine is on, I will never use the cashier. Therefore, the

situation does not apply to me." Hui avoided these problems by employing a hypothetical figure.

There are many potential sources of similar problems when simulating a homebanking service experience. For example, subjects might not own a PC, might be very satisfied with the current way of dealing with their bank or might value their personal contact with human tellers highly. To minimize the impact of heterogeneous consumption habits of individuals, all subjects were asked to role play the hypothetical Mr Smith (refer to Appendix C for a description of Mr Smith).

Subjects were provided with a summary description of the homebanking service<sup>2</sup>. This consisted of information on the transactions available and the response time of the system. Test results about system response times were also provided. The system description, the test results and a run on a simulator were used to manipulate the subjects' expectations about the service performance (for further details on manipulation related issues refer to section 6.4).

The subjects were asked to conduct a trial run on a homebanking service simulator. This was done to familiarize subjects with the system and to manipulate expectations about response time performance.

It was intended that subjects not have problems in conducting the experimental tasks and focus on assessing the service. Therefore, tasks to be conducted with

---

<sup>2</sup> Page 2 of the questionnaire presented in Appendix D.

the "Simulator" and with the "real system" were nearly identical. The instructions for these tasks are set out on pages 3,4 and 6,7 of the questionnaire presented in Appendix D.

#### **6.4 Manipulation of the Independent Variables**

In this study a 2 (expectations about the mean response time) X 2 (expectations about the variance of the response time) X 2 (actual levels of performance) factorial design was employed. The employed manipulations are described below.

##### Performance

The subjects tested the response time performance of the homebanking service with the following transactions: account balance enquiry and payment of a bill to the London Business School.

In his book on human-computer interfacing, Brown (1988) reviewed issues related to computer system response times. Although response time is potentially a major factor in system development costs and is assumed to be critical to user performance and satisfaction, he noted that response time guidelines have not been developed for use by system designers. Brown came to the conclusion that guidelines should be dependent on task-specific, system-specific and user-specific considerations. The

US Department of Defense (1981) has laid down the response time guidelines shown in Table 6-3 (1981). However these guidelines are qualified by the caveat that system response times are highly application driven. Systems like fire-control systems or command and control systems may require far more rapid response times than those proposed in the guidelines, whereas other systems may be designed to respond more slowly. Brown's review and the US Department of Defense guidelines show that response times ranging from immediate response to delay exceeding ten seconds seem to be acceptable for professional applications. It seems fair to assume that response time can be even longer for non-professional applications like home-shopping or homebanking.

Pretests with ten subjects were conducted to ascertain that the response time manipulations applied could be clearly distinguished and that disconfirmation-of-expectation processes could be experienced. Based on the literature review and the pretests, two levels of response time performance were used: 3 seconds (P1) and 9 seconds (P2).

**Table 6-3: System Response Time Guidelines from MIL-STD-1472C**

<b>Action</b>	<b>Definition</b>	<b>Max. Response Time in Seconds</b>
Key Print	Press key until appearance of character	0.2
XY EntEntry	Select field until visual verification	0.2
Function Selection	Selection until response	2.0
Simple Inquiry	Display of commonly used message	2.0
File Update	Update requires access to host file	10.0
Complex Inquiry	Seldom used calculations in graphic form	10.0

## Expectations

Expectations about the mean response time (E) and its variance (V) were manipulated using the written service descriptions and test reports presented on the following pages (refer also to page 2 of the questionnaire in Appendix D).

Two levels of point expectation [expected variance in performance equals to zero (V1)] were used: 1 second (E1V1) and 12 seconds response time (E2V1). These two levels of expectation were necessary in combination with the performance variation (P1 and P2) to test the classical disconfirmation-of-expectations model and Hypothesis I (Chapter II). A positive disconfirmation (E2V1,P1) and a negative disconfirmation (E1V1,P2) were needed to test Hypotheses IV and V (Chapter V).

Expectations E1V1 and E2V1 are point expectations. This means, that subjects were manipulated to expect the same performance level for every consumption experience with a variance in performance equal to zero (V1). To test Hypotheses II and III (Chapter III), however, expectations with the same mean performance but with a standard deviation greater than zero were required (V2). In the high variance condition (V2), the subjects were told that the performance could vary between immediate response (0 seconds) and 15 seconds.

Written descriptions of the service and a test report from the American bank were used to manipulate subjects' expectations about the time performance and its variance. To evoke point expectations (V1), the system description and the report

of the consumer association both gave identical and constant time performance descriptions. The stated variance in time performance was zero. For creating variance expectations (V2), the system description warned that deviations in time performance might occur, and the test report provided data that showed a possible time performance between zero and 15 seconds.

This manipulation was based on Fike's (1977) findings that people can estimate variability and 90% confidence ranges of samples reasonably well when the data are presented<sup>3</sup>.

---

<sup>3</sup> For a review of perception of variance when sample data are provided refer to Chapter III, Section 3.2.

1 second response time (E1) and low variance (V1):

In its advertisements, the American Bank claims that its system needs only one second to respond to each data input (= each ENTER strike), and that this response time always remains the same. The American Bank published the following test results on the response time of their system:

Table 1: Test results on the response time

Response time	% of trials falling in this category
0 - 2 seconds	100 %
3 - 5 seconds	-
6 - 10 seconds	-
11 - 15 seconds	-
16 - 20 seconds	-

"As you can see in Table 1, in all tests the system needed less than 2 seconds to respond."

---

1 second response time (E1) and high variance (V2):

In its advertisements, the American Bank claims that its system needs on average only one second to respond to each data input (= each ENTER strike), and that this response time can vary between immediate response and 15 seconds between each time the service is used. This variance is due to peak and low times. For example, the response time can vary between different times during the day and different days of the week. The American Bank published the following test results on the response time of their system:

Table 1: Test results on the response time

Response time	% of trials falling in this category
0 - 2 seconds	74 %
3 - 5 seconds	17 %
6 - 10 seconds	6 %
11 - 15 seconds	3 %
16 - 20 seconds	-

"As you can see in Table 1, in around three-quarters of all tests the system needed less than 2 seconds to respond. In the remaining quarter however, the system was slower, in some rare cases as slow as 15 seconds."

**12 seconds response time (E2) and low variance (V1):**

In its advertisements, the American Bank claims that its system needs **only twelve seconds** to respond to each data input (= each ENTER strike), and that this response time **always remains the same**. The American Bank published the following test results on the response time of their system:

Table 1: Test results on the response time

Response time	% of trials falling in this category
0 - 2 seconds	-
3 - 5 seconds	-
6 - 10 seconds	-
11 - 15 seconds	100 %
16 - 20 seconds	-

"As you can see in Table 1, in all tests the system needed between 11 and 15 seconds to respond."

---

**12 seconds response time (E2) and high variance (V2):**

In its advertisements, the American Bank claims that its system needs **on average only twelve seconds** to respond to each data input (= each ENTER strike), and that this response time **can vary between immediate response and 15 seconds** between each time the service is used. This variance is due to peak and low times. For example, the response time can vary between different times during the day and different days of the week. The American Bank published the following test results on the response time of their system:

Table 1: Test results on the response time

Response time	% of trials falling in this category
0 - 2 seconds	3 %
3 - 5 seconds	6 %
6 - 10 seconds	17 %
11 - 15 seconds	74 %
16 - 20 seconds	-

"As you can see in Table 1, in around three-quarters of all tests the system needed between 11 and 15 seconds to respond. In the remaining quarter however, the system was faster, in some rare cases as fast as 1 second."

A "Simulator" was used to further reinforce the expectations created by the written service descriptions. Subjects worked first on the Simulator and were told that they were not online connected to a mainframe. They were also told that the simulator works in exactly the same way as the real homebanking service. In other words, the response time of the simulated service was the performance described in the service description and test reports.

The following reasons for response times were given to the subjects. In the point-expectation condition (V1), the time elapsed before the system responded was explained in terms of telecommunications and mainframe processing times. In the high variance condition (V2), variations in utilization of the mainframe and telephone lines were given as reasons for varying response times.

## **6.5 Measures**

In this section measures frequently used in satisfaction research are discussed, and some studies that compare commonly applied measures are reviewed. The measures used in the experiment reported in the thesis are then discussed. The manipulation checks and measures employed are listed in Table 7-1. Wherever possible, the order of single items within a scale was randomized and 50% of the items were reversed.

For making the different satisfaction measures comparable, their order in the questionnaire was randomized. Six versions of the questionnaire, each with a different (randomized) order of the questions, were employed.

### **6.5.1 Measures of Expectation, Perceived Performance and Disconfirmation**

#### Measures of Expectations

Before conducting the experimental tasks with the "on-line" homebanking service, performance expectations were manipulated through a service description and a trial with a "Simulator". After using the simulator, subjects answered five questions about their expectations about the "real" or "on-line" connected homebanking service (refer to Table 6-4, or page 5 of the questionnaire in Appendix D). After that, subjects were exposed to the different experimental conditions, in which the response time performance of the system was varied.

Questions 3 and 4 locate the tails of the expected response time distribution. The difference between the minimum and maximum expected response time was used as a measure of expected variance.

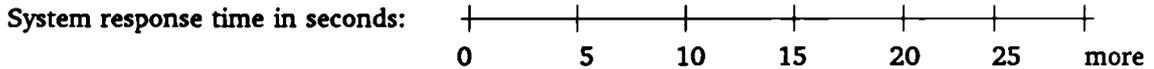


Table 6-5: Measures of Expected Variance of Performance

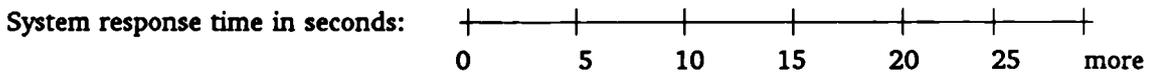
2. Please read the following statement carefully then circle the number which best describes your agreement or disagreement with the statement (e.g. 5=Strongly Agree)

	Strongly Agree	Agree nor Agree	Disagree	<u>Disagree</u>	Strongly <u>Disagree</u>
<p>After reading the <u>information page</u> and using the <u>Simulator</u> Mr Smith believes that the response time of the <u>real American Homebanking System</u> could vary by more than 5 seconds each different time he uses the system</p>	5	4	3	2	1

3. What would Mr Smith expect the maximum possible response time of the real American Homebanking System to be?



4. What would Mr Smith expect the minimum possible response time of the real American Homebanking System to be?



### Measures of Perceived Performance

Perceived performance and all subsequent constructs were measured after the subjects had carried out the experimental task with the "on-line" connected service.

In many empirical studies extremely high correlations have been found between perceived performance and satisfaction measures. In the literature review in Chapter II, Section 4.3 it was hypothesized that the reason for these results are inappropriate perceived performance measures. These measures include evaluative components and, in consequence, capture some component of satisfaction. In this study, it was therefore crucial to measure performance perceptions rather than evaluations, i.e. to measure perceived system response time in time units rather than with a slow/fast scale.

---

**Table 6-6: Perceived Performance Measures**

---

1. How long did the American Homebanking Service need approximately to respond to each "ENTER" strike. Please indicate on the following scale like this: (~~—X—~~)



3. In what range did the American Homebanking Service respond to each "ENTER" strike? (please tick)

<input type="checkbox"/>					
0-2	3-5	6-10	11-15	16-20	21 and more
seconds	seconds	seconds	seconds	seconds	seconds

## Measures of Confirmation/Disconfirmation

Perceived disconfirmation rather than subtractive disconfirmation (perceived performance scores minus expectation scores) was used because of the problems inherent in the subtractive measures (for an explanation refer to Chapter II, Section 1.1.3.2).

Disconfirmation was measured using attribute-specific and overall measures. This approach has been adopted by Churchill and Surprenant (1982), Bitner (1987) and many other researchers. In accordance with the generally applied linear compensatory multiattribute models, the responses to individual attribute measures were weighted by their relative importance and then summed to generate an overall construct score, which again measured overall disconfirmation or satisfaction (e.g. Bitner 1987, Churchill and Surprenant 1982)<sup>4</sup>).

In this study the disconfirmation and satisfaction associated with the following attributes was measured: response time, visual appeal, reliability, ease of use, range of transactions offered and security. These attributes were identified through telephone interviews of product managers of telephone or television-set based homebanking services, and through review of the multiattribute literature related to services (e.g. Bateson 1985, Parasuraman et al 1985, 1988). In contrast to most experiments on consumer satisfaction, only one attribute performance, namely

---

<sup>4</sup> Refer to Chapter II, Section 1.1.5 for a review on the application of multiattribute models in consumer satisfaction.

response time, was manipulated. All other attribute performance levels remained constant.

Scales commonly applied in consumer behaviour were used for this study. Attribute specific and overall confirmation/disconfirmation were measured with Oliver's (1980a) and Bitner's (1987) seven-point semantic differential scale ranging from "The response time of the home banking system was worse than expected" to "better than expected" (question 2). Churchill and Surprenant's (1982) seven-point semantic differential scale (question 4), with the attribute specific measure for the service response time ranged from "My expectations of the response time were too high: It took longer than expected" to "Too low: It took less time than I thought", was also used. The importance weights were measured using Churchill's (1991, p. 440-1) comparative rating scale (question 10).

---

**Table 6-7: Disconfirmation-of-Expectation Measures**

---

2. How was the actual performance of the following features of the American Homebanking Service in comparison to Mr Smith's expectations? (Please tick)

	worse than expected		just as expected		better than expected		
The visual appeal of the system was	___	___	___	___	___	___	___
The reliability of the service was	___	___	___	___	___	___	___
The ease of use was	___	___	___	___	___	___	___
The range of transactions possible was	___	___	___	___	___	___	___
The response time was	___	___	___	___	___	___	___
The security of the service was	___	___	___	___	___	___	___
Overall, the homebanking service was	___	___	___	___	___	___	___

cont...

---

4. What feelings would Mr Smith have about his previous expectations after using the American Homebanking Service? (please tick)

In comparison to the American Homebanking Service, his expectations were

too high:

Accurate:

Too low:

It was poorer than he thought

It was just as he expected

It was better than he thought

His expectation of the security of the system was \_\_\_\_\_

His expectation of the response time was \_\_\_\_\_

His expectation of the visual appeal of the system was \_\_\_\_\_

His expectation of the ease of use was \_\_\_\_\_

His expectation of the range of transactions possible was \_\_\_\_\_

His expectation of the reliability of the system was \_\_\_\_\_

Overall, his expectations about the homebanking service were \_\_\_\_\_

10. Now, please indicate how important each of the following features or characteristics of homebanking are to Mr Smith in general. Please divide 100 points between the following six features to show their importance to Mr Smith.

Range of transactions possible \_\_\_\_\_

Ease of use \_\_\_\_\_

Visual appeal of the system \_\_\_\_\_

Security of the system \_\_\_\_\_

Response time of the system \_\_\_\_\_

Reliability of the system \_\_\_\_\_

### 6.5.2 Satisfaction Measures

The selection of satisfaction measures proved to be a difficult task because a large number of very different measures has been used by satisfaction researchers. The measurement scales range from one to twelve items, and from predominantly cognitive (usually belief) to predominantly affective (usually like-dislike) measures. The research on satisfaction measures is reviewed in this section.

In most studies with findings that can be generalized directly to managerial marketing problems, simple one-item satisfaction measures have been used. This is also true for most theory application research that is not directly concerned with examining the construct or process of satisfaction but is concerned with some related issue. However, these generally applied one-item measures have been shown to be inferior to a number of more sophisticated multi-item scales. The most promising measures of both kinds are reviewed in the next section.

#### Review of Studies Comparing Satisfaction Measures

In the following section three studies comparing satisfaction measures are reviewed: Churchill and Surprenant (1982), Westbrook (1980), Westbrook and Oliver (1981). Westbrook (1980) tested the appropriateness of the delighted-terrible scale for the measurement of satisfaction. Westbrook and Oliver (1981) reviewed several types of scales used in consumer behaviour and other disciplines. Measures from other

disciplines adapted to consumer satisfaction included job, marital, patient and life satisfaction. Churchill and Surprenant (1982) employed five standard scales widely used by consumer behaviour researchers.

Westbrook (1980) examined the suitability of the Delighted-Terrible scale (D-T scale) for measuring consumer satisfaction. This scale was developed and extensively tested by Andrews and Withey (1976) in sociological research on the perception of life quality. It has been applied successfully in assessing evaluative responses to major life concerns such as one's job, health, family and neighbourhood.

Three other measures were tested by Westbrook. The first of these was a Percentage Scale, ranging from 100% - Completely Satisfied to 0% - Not at all Satisfied. The second was a Need Satisfaction - Dissatisfaction seven-point semantic differential scale with the question "To what extent does this car meet your needs at this time?" The scale ranged from "Extremely Well" to "Extremely Poorly". The third measure was a general, nondirective question on subjects' thoughts and feelings about the experience of owning and using the product, followed by clarifying probes, and a specific open-ended follow-up question about any particular aspects of the product liked or disliked. The responses were then coded into four satisfaction categories.

Westbrook conducted three different studies for testing these measures. In Study I, self-administered questionnaires were completed by 72 students about their satisfaction with automobiles, bank services and wristwatches. Study II consisted

of personal interviews with 151 female heads of households on their satisfaction with their washing machines, refrigerators and colour television sets. In Study III 47 students were questioned about their cars.

The D-T scale seemed to be superior to the other measures and showed good reliability, convergent and discriminant validity, and nomological validity. The evidence from the three separate studies indicates the superiority of the D-T scale for the measurement of consumer satisfaction. First, the skewness of satisfaction responses typical of the other satisfaction measures was significantly reduced by use of the D-T scale. This may be attributed to its second advantage, the superiority of the D-T scale in capturing the affective component of satisfaction. For example, of 25 washing machine owners in Study II who selected the uppermost scale position on the percentage scale "completely (100%) satisfied" only 14 chose the uppermost scale position of the D-T scale "delighted". The remaining 11 respondents describing their satisfaction as "pleased" or "mostly satisfied". Similar results were obtained for the other product/service categories. However, the ability of this scale to capture affective components is only assumed by the researchers rather than empirically tested.

Since the publication of Westbrook's (1980) paper the D-T scale has been tested in other consumer satisfaction studies and evaluated favourably (e.g. Bitner 1987, Westbrook 1987, Westbrook and Oliver 1981). Westbrook (1980), however, recommended that despite the predominance of simple, single-item satisfaction measures, future research should focus on the development of multi-item measures.

The primary aims of this research would be to reduce measurement error and improve the construct validity of satisfaction measures.

Westbrook and Oliver (1981) continued Westbrook's (1980) work by developing and testing multi-item satisfaction measures. They reviewed research in the areas job, marital, patient and life satisfaction and found widespread use of multi-item scales for satisfaction measurement. Westbrook and Oliver selected five measures to be tested empirically. The first was a three-item verbal scale (Verbal) for overall assessment of product satisfaction (1) the Delighted-Terrible seven-point rating scale, (2) a "Completely Satisfied-Not at all Satisfied" 11-point rating scale ranging from 100% to 0%, and (3) a behavioural tendency 11-point rating scale ranging from "Certain I'd do it again" to "No chance I'd do it again." These specific items were combined additively to obtain an overall satisfaction score. The scale was adapted from Andrews and Withey (1976).

The second measure made use of four distinct graphic rating scales: faces, thermometer, circles and ladder. According to Westbrook and Oliver, these particular items represent a desirable mixture of nonverbal content, and the faces and thermometer scales are chiefly affective, while the circles and ladder are more cognitive-evaluative in tone. They were combined additively to yield an overall graphic scale (hereafter referred to as Graphic) (Andrews and Withey 1976).

Third, a Likert summated scale in which 12 statements indicating varying sentiments of overall satisfaction with the product were presented to respondents for their agreement. Responses were made on a five-interval Likert-type scale.

Fourth, a set of seven semantic differential items, again dealing with various questions about overall satisfaction judgements. To reduce the cognitive strain on respondents only five intervals were used on the semantic differential instead of the conventional seven. These items were combined into a simple additive index (hereafter referred to as S-D).

Fifth, the differences between perceived attribute performances and the ideally desired levels were computed. This measure, called the Porter Scale, was used as an inferential satisfaction measure.

Self-administered questionnaires were completed by 175 students. Respondents were questioned about their experiences with currently owned automobiles and hand-held calculators. The Verbal, Likert and S-D measures emerged as the best measures. They showed high internal consistency, converged with other satisfaction measures, and succeeded in discriminating between unrelated constructs. At the same time, they also produced fairly symmetrical, dispersed distributions of individual responses. The few items (three) of the Verbal scale make it attractive from the standpoint of administrative efficiency. The Porter Scale had low internal consistency and did not converge with other satisfaction measures. Westbrook and Oliver concluded that the Porter Scale is not a good measure of satisfaction.

In summary, Westbrook and Oliver (1981) concluded that the Verbal scale, the S-D scale and the 12-item Likert-type scale were the most reliable measures for consumer satisfaction.

Churchill and Surprenant (1982) used five satisfaction scales in their experiment with video disc players and hybrid plants. The researchers used two attribute specific scales: (1) A seven-interval Likert-type scale measuring beliefs and (2) a seven-interval semantic differential scale with "like" and "dislike" at the extreme points. For both scales an overall measure was obtained by adding the scores of the single attribute evaluations. The researchers also employed three overall measures: (3) a seven-interval continuum scale, ranging from "completely satisfied" to "completely dissatisfied"; (4) a faces scale, and (5) a behavioural tendency scale asking for the subjects' purchase probability.

All measures but the two attribute specific measures had insufficient discriminant and convergent validity. Churchill and Surprenant concluded that the two multi-item attribute specific belief and affect measures are good measures of consumer satisfaction. This is in line with Westbrook and Oliver's (1981) recommendation of use of multi-item measures.

## Selected Measures

Three categories of measures were employed in this study: attribute specific, single-item overall and multi-item overall measures. The measures shown in Tables 6-8 to 6-10 have been selected in accordance with the results of the literature review on satisfaction measures above.

### Attribute Specific Satisfaction Measures:

The attribute specific satisfaction scales listed below were weighted with the same importance weights as described in the discussion on disconfirmation scales. The measures employed were Churchill and Surprenant's (1982) like-dislike (question 8) and belief scales (question 11). The belief scale was put together with a four-item perceived control scale and a twelve-item satisfaction measure, which are described later. The order of all items was randomised (refer to Appendix D, Part C, question 11). Westbrook's (1980) delighted-terrible scale was the third attribute specific measure employed (question 12).

**Table 6-8: Attribute Specific Measures**

8. A number of features of the American Homebanking Service are listed below. Please imagine how Mr Smith would feel about them. (please tick)

Range of transactions possible	_____	_____	_____	_____	_____	_____
	Like					Dislike
Ease of use	_____	_____	_____	_____	_____	_____
	Like					Dislike
Visual appeal of the system	_____	_____	_____	_____	_____	_____
	Like					Dislike
Security of the system	_____	_____	_____	_____	_____	_____
	Like					Dislike
Response time of the system	_____	_____	_____	_____	_____	_____
	Like					Dislike
Reliability of the system	_____	_____	_____	_____	_____	_____
	Like					Dislike

11. Below are some statements which describe how Mr Smith might feel after he used the American Homebanking Service for the first time. Please read each statement carefully then circle the number which best describes your agreement or disagreement with the statement (e.g. 1=Strongly Agree, 5=Strongly Disagree).

	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
Mr Smith feels that the reliability of the system is very poor	1	2	3	4	5
Mr Smith feels that the ease of use of the system is excellent	1	2	3	4	5
He feels that the response time of the system is very bad	1	2	3	4	5
He feels that the range of possible transactions is very wide	1	2	3	4	5
He feels that the security of the system is very high	1	2	3	4	5
He feels that the visual appeal of the system is very poor	1	2	3	4	5

cont...

---

12. A number of features of the American Homebanking Service are listed below. Please imagine how Mr Smith would feel about them. (please tick like this:  :  : )

Ease of use	delighted	<input type="checkbox"/>	terrible						
Security of the system	delighted	<input type="checkbox"/>	terrible						
Range of transactions possible	delighted	<input type="checkbox"/>	terrible						
Response time of the system	delighted	<input type="checkbox"/>	terrible						
Reliability of the service	delighted	<input type="checkbox"/>	terrible						
Visual appeal of the service	delighted	<input type="checkbox"/>	terrible						

---

### Single-Item Overall Satisfaction Measures:

Single-item overall measures are widely applied in satisfaction research. Therefore, the most common of them were included primarily to assess their performance in comparison to the apparently better multi-item measures. The following measures were selected: Churchill and Surprenant's (1982) like-dislike scale (question 8) and Westbrook's (1980) delighted-terrible scale (question 12). Finally, the frequently used 5-point bipolar scale ranging from "very dissatisfied" to "very satisfied" was employed (question 14).

---

**Table 6-9: Single-Item Satisfaction Measures**

---

8. The American Homebanking System as a whole \_\_\_\_\_  
Like \_\_\_\_\_ Dislike

12. The Homebanking Service as a whole delighted \_\_\_\_\_:\_\_\_\_\_:\_\_\_\_\_:\_\_\_\_\_:\_\_\_\_\_:\_\_\_\_\_:\_\_\_\_\_ terrible

14. Finally, considering everything, how satisfied would Mr Smith be with the American Homebanking Service? (please tick):

\_\_\_\_\_      \_\_\_\_\_      \_\_\_\_\_  
very      neither      very  
dissatisfied      satisfied      satisfied  
nor dis-  
satisfied

---

### Overall Multi-Item Satisfaction Measures:

Overall multi-item measures seem to be superior to single-item and to weighted and added attribute specific measures. In empirical comparisons, the following two measures have performed best: Westbrook and Oliver's (1981) verbal scale (question 5, 6 and 13) and the 12-item Likert-type scale (question 11). These measures are presented in Table 6-10.



---

11. Below are some statements which describe how Mr Smith might feel after he used the American Homebanking Service for the first time. Please read each statement carefully then circle the number which best describes your agreement or disagreement with the statement (e.g. 1=Strongly Agree, 5=Strongly Disagree).

	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
If he could do it over again, he'd buy a different service	1	2	3	4	5
He truly enjoys this homebanking service	1	2	3	4	5
He feels bad about his decision to buy this service	1	2	3	4	5
He is <u>not</u> happy that he bought this service	1	2	3	4	5
His choice to buy this service was a wise one	1	2	3	4	5
Sometimes, he will have mixed feelings about keeping it	1	2	3	4	5
Mr Smith is satisfied with his decision to buy this service	1	2	3	4	5
Owning this service system has been a good experience for Mr Smith	1	2	3	4	5
This system hasn't worked out as well as he thought it would	1	2	3	4	5
This homebanking service is exactly what Mr Smith needs	1	2	3	4	5
This is one of the best homebanking services Mr Smith could have chosen	1	2	3	4	5
Mr Smith is sure that it was the right thing to choose this service	1	2	3	4	5

---

### 6.5.3 Measures of Affect

The Russell (1978, 1980; Russell and Pratt 1980; Russell, Ward and Pratt 1981) circumplex model of affect and its scales were used in this study (see Chapter III). Russell and Pratt operationalized the two dimensions of affect with two scales. Scale I encompassed the two dimensions of pleasure/displeasure and arousal/sleepy. Each dimension was measured with five positively and five negatively keyed items. Scale II consisted of the dimensions exciting/gloomy and distressing/relaxing. Again, each dimension was measured with ten items.

Both scales are metrics for the same two-dimensional space in which each point is the operationalization of a person's internal affective state (compare Figure 4-1). X and y axes are commonly defined as pleasure/displeasure and arousal/sleepy, respectively. Coordinates can then be obtained from the following theoretical equations:

$$\text{Scale I: } x = 1.00 (P-U) \quad \text{and} \quad y = 1.00 (A-S),$$

where P-U is the score on the pleasant-unpleasant score and A-S the scale on the arousing-sleepy scale.

The same x and y coordinates can also be determined with the exciting-gloomy (E-G) and distressing/relaxing (D-R) scales:

$$\text{Scale II: } x = .707 (E-G) - .707 (D-R) \quad \text{and} \quad y = .707 (E-G) + .707 (D-R)$$

To reduce stress and avoid subject fatigue, and because there was to be a second affect measure included in the questionnaire, it was decided to use only the pleasure and the arousal scales in this study. Another advantage of using these two scales is that they are directly comparable with the other affect measure employed in the questionnaire. The Russell and Pratt (1980) scale of the affective quality of social and human/environment interactions used is presented in Table 6-11 (question 9).

The second measure of affect employed was the Mehrabian (1980) semantic differential scale of emotional response to social situations and environment (question 7).

Table 6-11: Russell and Pratt's Measure of Affect Adapted to Homebanking Services

9. Below is a list of words that can be used to describe situations and how people feel in those situations. Please assess how accurately each of the words below describes Mr Smith's situation and his feelings while using the American Homebanking Service for the first time in his office.

You may need some imagination and there are no right or wrong answers. Use the following 1-8 rating scale for your answer, and please make sure that you have given an answer for each word. (please circle)

	extremely <u>inaccurate</u>	very <u>inaccurate</u>	quite <u>inaccurate</u>	slightly <u>inaccurate</u>	slightly accurate	quite accurate	very accurate	extremely accurate
nice	1	2	3	4	5	6	7	8
forceful	1	2	3	4	5	6	7	8
dissatisfying	1	2	3	4	5	6	7	8
unpleasant	1	2	3	4	5	6	7	8
intense	1	2	3	4	5	6	7	8
pretty	1	2	3	4	5	6	7	8
pleasant	1	2	3	4	5	6	7	8
lazy	1	2	3	4	5	6	7	8
alive	1	2	3	4	5	6	7	8
uncomfortable	1	2	3	4	5	6	7	8
drowsy	1	2	3	4	5	6	7	8
inactive	1	2	3	4	5	6	7	8
displeasing	1	2	3	4	5	6	7	8
arousing	1	2	3	4	5	6	7	8
active	1	2	3	4	5	6	7	8
pleasing	1	2	3	4	5	6	7	8
idle	1	2	3	4	5	6	7	8
repulsive	1	2	3	4	5	6	7	8
slow	1	2	3	4	5	6	7	8
beautiful	1	2	3	4	5	6	7	8

---

Table 6-12: Mehrabian's Measure of Affect Adapted to Homebanking

---

7. Now think about the feelings, moods and emotions Mr Smith might have experienced while using the American Homebanking Service for the first time in his office. Take about half a minute to get into the mood of the situation. Then rate the feelings with the adjective pairs below. Some of the pairs might seem unusual, but you will probably feel more one way than the other.

For each pair, put a tick (Example : \_\_\_ :  : \_\_\_) closer to the adjective which you believe best describes Mr Smith's feelings, moods and emotions. The more appropriate that adjective seems, the closer you should put your tick to it. Please make sure that you have given a tick for each pair.

Bored	_____	:	_____	:	_____	:	_____	:	_____	:	_____	:	_____	Relaxed
Hopeful	_____	:	_____	:	_____	:	_____	:	_____	:	_____	:	_____	Despairing
Calm	_____	:	_____	:	_____	:	_____	:	_____	:	_____	:	_____	Excited
Happy	_____	:	_____	:	_____	:	_____	:	_____	:	_____	:	_____	Unhappy
Jittery	_____	:	_____	:	_____	:	_____	:	_____	:	_____	:	_____	Dull
Melancholic	_____	:	_____	:	_____	:	_____	:	_____	:	_____	:	_____	Contented
Aroused	_____	:	_____	:	_____	:	_____	:	_____	:	_____	:	_____	Unaroused
Pleased	_____	:	_____	:	_____	:	_____	:	_____	:	_____	:	_____	Annoyed
Relaxed	_____	:	_____	:	_____	:	_____	:	_____	:	_____	:	_____	Stimulated
Wide awake	_____	:	_____	:	_____	:	_____	:	_____	:	_____	:	_____	Sleepy
Unsatisfied	_____	:	_____	:	_____	:	_____	:	_____	:	_____	:	_____	Satisfied
Sluggish	_____	:	_____	:	_____	:	_____	:	_____	:	_____	:	_____	Frenzied

---

#### **6.5.4 Measures of the Demographic Background and Other Individual Differences of Subjects**

As mentioned earlier, measuring the subjects' personal backgrounds is a way to solving the problems arising from heterogeneous samples. However, the number of potentially relevant background variables can be very large and, in its current state, consumer satisfaction theory cannot guide selection of subject variables which impact upon satisfaction. In spite of this lack of guidance, standard demographic data were collected in this study to test at least whether the subjects were actually assigned randomly to the treatment conditions (Appendix D, Part D).

Furthermore, certain consumption habits (use of automatic teller machines, homebanking and online data base services) and computer literacy (use of computer hardware and software) were controlled for. Psychological traits of subjects that might influence the dependent variables of the study were also measured. These were trait arousal and trait pleasure (Mehrabian and Russell 1974).

Finally, the subject's relationship with the London Business School was controlled for. All subjects who were either full or part-time employees or full-time students were categorised as members of LBS.

## **7 Summary**

In this chapter, methodological considerations were discussed and the method and measures employed in this study were determined. In sections one to three, four basic types of validity and the main threats to them, their relative importance in relation to the research objectives, and three classical research designs were discussed.

The objective of this study is theory rather than effects application, and the hypothesized relationships are primarily causal in nature. The review in sections one to three suggests that for this type of research, the emphasis on the different types of validity should be in the following order of decreasing importance: internal, construct, statistical conclusion and, finally, external validity. An experimental design was chosen as it ranks higher on internal validity than any other design. This is because it minimizes the possibility that other variables influence the observed correlations between independent and dependent variables.

In section four, a controlled laboratory setting was chosen because it minimizes extraneous sources of variation and thereby maximizes internal validity. The three main reasons for manipulating response time expectations and performance of a new PC-based homebanking service were also presented.

In section five, subject selection was discussed. As this study aimed at theory application, any sample that allowed a falsification of the theory could be selected.

In the context of this study this meant that subjects had to be selected so that they were able to understand and operate a PC-based homebanking service with the information provided in the experiment.

Finally, all aspects of the experimental work were discussed. They were (1) design and production of the homebanking system, (2) description of the experimental sessions, (3) reason, system description and instructions given to subjects, (4) manipulation of the independent variables, and (5) the selection and description of all measures employed.

# Chapter VII

## Data Analysis

The analyses of the data obtained from the experimental study are presented in this chapter. In the first section all the operationalizations of the variables and descriptions of their distributions are presented. The second section presents results of the preliminary analyses that were conducted. These analyses were examination of the demographic background of subjects, check of random assignment to experimental conditions, assessment of reliability of measures and manipulation check. In the third section results of traditional hypothesis tests are presented. The proposed model<sup>1)</sup> as implied by the hypotheses, the disconfirmation-of-expectations model and the results of the traditional statistical analyses was tested as a whole using structural equation modelling. The results are presented in section four.

---

<sup>1</sup> The hypothesized model is presented in Figure 7-2.

## 1 Independent and Dependent Variables Included in the Data Analysis

### Independent Variables

The independent variables of the study were the three treatment variables: expected performance (fast and slow), expected variance (low and high) and performance (fast and slow). Chapter VI, Section 6.4 provided a detailed description of all manipulations. The following constructs were measured and used to assess whether the experimental manipulations were successful: expected mean performance, expected variance in performance and perceived performance,

### Dependent Variables

The following dependent variables were analyzed as a function of the experimental treatments: disconfirmation-of-expectations, magnitude of disconfirmation, pleasure, arousal and satisfaction. A summary of all measures is provided in Table 7-1. Chapter VI, Section 6.5 provided a detailed description of all measures employed. A description of the distributions of all dependent variables is presented in Table 7-2.

Table 7-1: Operationalizations of the Treatment and Dependent Variables

Variable	Operationalization	Symbol
<b><u>Independent Variables:</u></b>		
Expectation Manipulation of Mean Performance	(1) Description of the Homebanking Service (2) Trial Run with a Simulator	MEXPECT
Expectation Manipulation of Variance in Performance	(1) Description of the Homebanking Service (2) Test Report of the American Bank  (Two Levels: Point-Expectations and a Possible Range of Performance Between Immediate Response and 15 Seconds)	MVAR
Performance Manipulation	Response Time in Seconds of the Homebanking Service to Each Data Input or Command	MPERF
<b><u>Manipulation Checks:</u></b>		
Performance Expectation	(1) A One-Item Continuous Scale (2) A Multichotomous Question	EPERF1 EPERF2
Variance Expectation	(1) A Multichotomous Question (2) The Subtracted Value of Two Continuous Scales Asking for the Minimum and Maximum Performance Expected	EVAR1 EVAR2
Perceived Performance	(1) A One-Item Continuous Scale (2) A One-Item Multichotomous Question	PPERF1 PPERF2
cont ...		

Variable	Operationalization	Symbol
<b><u>Dependent Variables:</u></b>		
Disconfirmation Time Performance	(1) A One-Item Multichotomous Question for the Disconfirmation of Time Performance (e.g. Oliver 1980, Bitner 1987)	DTIME1
	(2) Churchill and Surprenant's (1982) One-Item Multichotomous Question for Time Performance	DTIME2
	(3) $DTIME = (DTIME1 + DTIME2)/2$	DTIME
Disconfirmation Overall Performance	(4) A One-Item Multichotomous Question for Overall Disconfirmation (e.g. Oliver 1980, Bitner 1987)	DOQ1
	(5) Churchill and Surprenant's (1982) One-Item Multichotomous Question for Overall Disconfirmation	DOQ2
	(6) $DOQ = (DOQ1 + DOQ2)/2$	DOQ
	(7) A Linear Compensatory Multiattribute Measure Consisting of a Multichotomous Scale of Disconfirmation with six attributes (e.g. Oliver 1980, Bitner 1987), weighted by importance scores of a comparative rating scale according to Churchill (1991)	DOW1
	(8) A Linear Compensatory Multiattribute Measure Consisting of Churchill and Surprenant's (1982) Multichotomous Scale of Disconfirmation with six attributes, weighted by importance scores of a comparative rating scale according to Churchill (1991)	DOW2
Pleasure	(1) Mehrabian and Russell's (1974) Scale of Pleasure: A Six-Item Semantic Differential Scale	PLEAS1
	(2) Russell and Pratt's (1980) Scale of Pleasure: A Ten-Item Likert-Type Scale	PLEAS2
Arousal	(1) Mehrabian and Russell's (1974) Scale of Arousal: A Six-Item Semantic Differential Scale	AROUS1
	(2) Russell and Pratt's (1980) Scale of Arousal: A Ten-Item Likert-Type Scale	AROUS2
cont ...		

Variable	Operationalization	Symbol
Satisfaction	(1) One-Item 7-Point Like-Dislike Bipolar Scale (Churchill and Surprenant 1982)	SLDQ
	(2) One-Item 7-Point Delighted-Terrible Bipolar Scale (Westbrook 1980)	SDTQ
	(3) One-Item 5-Point Bipolar Scale (e.g. Tse and Wilton 1989)	SD
	(4) A Linear Compensatory Multiattribute Measure Consisting of a 7-Point Bipolar Like-Dislike Scale With Six Attributes (e.g. Churchill and Surprenant 1982, Oliver 1980, Bitner 1987), Weighted by Importance Scores of a Comparative Rating Scale According to Churchill (1991)	SLDW
	(5) A Linear Compensatory Multiattribute Measure Consisting of a 7-Point Bipolar Delighted-Terrible Scale With Six Attributes (Westbrook 1980), Weighted by Importance Scores of a Comparative Rating Scale According to Churchill (1991)	SDTW
	(6) A Linear Compensatory Multiattribute Measure Consisting of a 7-Point Likert-Type Belief Scale With Six Attributes (Churchill and Surprenant 1982, Tse and Wilton 1989), Weighted by Importance Scores of a Comparative Rating Scale According to Churchill (1991)	SBELW
	(7) Westbrook and Oliver's (1981) Three-Item Verbal Scale of Satisfaction Consisting of A 7-Point Delighted-Terrible Scale (SVERBAL1), an 11-Interval Percentage Scale (SVERBAL2) and an 11-Interval Behaviour Tendency Scale (SVERBAL3)	SVERBAL
	(8) Westbrook and Oliver's (1981) 12-Item Likert-Type Scale	SWLS12

Table 7-2: Description of the Distribution of all Dependent Variables

Measure	Possible Range	Mean	SD	Min	Max
<u>Manipulation Checks</u>					
EPERF1	0-30	7.1	5.40	1	25
EPERF2	1-6	2.8	1.33	1	6
EVAR1	1-5	2.8	1.27	1	5
EVAR2	0-30	8.2	6.24	0	25
PPERF1	0-30	7.3	4.08	1	20
PPERF2	1-6	2.7	1.06	1	6
<u>Dependent Variables</u>					
DTIME1	1-7	4.0	1.98	1.0	7.0
DTIME2	1-7	4.0	2.02	1.0	7.0
DOQ1	1-7	4.5	1.39	1.0	7.0
DOQ2	1-7	4.3	1.32	1.0	7.0
DOW1	1-7	4.2	.87	2.2	6.8
DOW2	1-7	4.1	.78	1.4	6.7
AROUS1	1-7	4.2	.92	1.8	6.8
AROUS2	1-8	4.9	1.05	2.1	7.2
PLEAS1	1-7	5.0	.97	1.8	6.8
PLEAS2	1-8	5.4	.99	1.6	8.0
SLDQ	1-7	5.2	1.18	2.0	7.0
SDTQ	1-7	5.1	1.14	2.0	7.0
SD	1-5	3.9	.80	2.0	5.0
SLDW	1-7	4.9	.97	2.1	7.0
SDTW	1-7	4.7	.99	1.5	6.9
SBELW	1-5	3.4	.60	1.3	4.9
SWLS12	1-5	3.5	.57	1.5	4.8
SVERBAL	0-10	7.1	1.52	1.2	9.7

## **2 Analyses Conducted Prior to Hypothesis Testing**

A number of preliminary analyses were conducted before hypothesis testing. First, the demographic background of the subjects was examined. Second, the cell sizes for each experimental condition were calculated and a check on the random assignment of subjects to each cell was performed. As a correlation between two variables can be blurred by unreliable measures, Cronbach alpha values were calculated for all multi-item variables in the third section. Fourth, the convergent and discriminant validity of all independent variables employed was examined. Fifth, checks were made to see whether the experimental manipulations had been successful and, finally, a summary is provided.

### **2.1 Usable Questionnaires and Demographic Background of Subjects**

Out of 134 questionnaires, two had to be discarded, because less than two-thirds of all questions had been answered. The remaining 132 questionnaires were used for all data analyses (excluding Lisrel) and, although in 10 questionnaires data on some variables were missing, analyses were conducted to obtain the maximum information from the data available. A Lisrel analysis requires a listwise deletion of all records with data on one or more variables missing, and this left data from 122 questionnaires for testing the model as a whole.

A summary of the demographic background of the sample is provided in Table 7-3. It should be noted that a relatively small proportion of subjects were students; 80% of all subjects were either employed or self-employed. About half of the subjects were neither students nor staff of the London Business School, and all subjects had some PC experience and had used an automatic teller machine at least once before.

Table 7-3: Demographic Background of the Subjects, N=132

Variable	Categories	N	%
Sex	Female	61	46.2
	Male	71	53.8
Age	Under 20	3	2.3
	21-25	31	23.5
	26-30	45	34.1
	31-35	25	18.9
	36-40	13	9.8
	41-45	6	4.5
	46-50	5	3.8
	51-55	4	3.0
Over 56	-	0.0	
Nationality	British	89	67.4
	Other Europeans	26	19.7
	Asian	11	8.3
	American	6	4.6
Marital Status	Single	62	47.0
	Living with Partner	9	6.8
	Married	52	39.4
	Separated/Divorced	8	6.1
	Widowed	1	.8

cont ...

Variable	Categories	N	%
Academic Qualification	GSE/O-Level/GCSE	8	6.1
	A-Level	16	12.1
	HND or Equivalent	12	9.1
	Degree	69	52.3
	Higher Degree	27	20.5
Occupational Status	Full-Time Student	26	19.7
	Part-Time Student	1	.8
	Full-Time Employed	96	72.7
	Part-Time Employed	4	3.0
	Self-Employed	4	3.0
	Unemployed	1	.8
Annual Income	Less than £5,000	19	14.4
	£5,000 to £9,999	6	4.5
	£10,000 to £19,999	52	39.4
	£20,000 to £29,999	16	12.1
	£30,000 to £40,000	25	18.9
	More than £40,000	14	10.6
Member of LBS	Yes	67	50.8
	No	64	48.5
	Missing	1	.8
Computer Literacy (Scale from 1 to 5, 1 equals very low level of computer literacy)	1	3	2.3
	2	19	14.4
	3	59	44.7
	4	38	28.8
	5	6	4.5
	Missing	7	5.3
Consumption Habits (Scale from 1 to 5, 1 equals very infrequent use of banking self services)	1	11	8.7
	2	14	10.6
	3	79	59.8
	4	22	16.7
	5	1	0.8
	Missing	5	3.8

## **2.2 Number of Subjects Per Experimental Condition and Randomization Checks**

The number of subjects randomly assigned to each treatment condition is shown in Table 7-4. A check was conducted to assure that the assignment of subjects had in fact been random. This random assignment is necessary to control for any background variables that might impact on the independent variables of interest. Four-way cross tabulations were run between the three treatment variables (mean expectation, variance expectation and performance) and the subjects' demographic backgrounds (sex, age, nationality, marital status, education, occupational status, income and membership of LBS), psychological traits (trait pleasure and trait arousal), computer literacy, and consumption of banking self-services.

All chi-square values indicated independence (non-significant at .05 level) between the experimental conditions and the measured background variables. The null-hypothesis that all groups had been equivalent prior conducting the experiment was not refuted. Therefore, all rival hypotheses that observed effects on the independent variables might have been due to prior group differences along the dimensions controlled for could be rejected.

Table 7-4: Number of Subjects Per Experimental Condition

		<u>Fast Expectations</u>		<u>Slow Expectations</u>	
		<u>Performance</u>		<u>Performance</u>	
		Fast	Slow	Fast	Slow
<u>Expected Variance</u>	Low	17	16	16	18
	High	15	17	17	16

### 2.3 Reliability of Measures

Since unreliable scales can mask the existence of a significant correlation between variables, the reliability of all multi-item measures employed was first examined before hypothesis testing. In each case, a Cronbach Alpha ( $\alpha$ ) was computed to examine the reliability of the measure. All  $\alpha$ -values are presented in Table 7-5. All items of the scales employed were retained. For each, the modification index showed that by excluding one or more items, the  $\alpha$ -values would have been lower in most cases. A marginal improvement (up to .01) could have been achieved for the scales AROUS2 and SBELA.

According to Nunally (1981), an  $\alpha$ -value of .80 or above is to be considered ideal, because any measurement error will then have little impact on correlations between variables. The scales of PLEAS1, PLEAS2, AROUS2, SWLS12 and SVERBAL showed high reliability with  $\alpha$ -values of at least .80. The  $\alpha$ -value of AROUS1 ( $\alpha=.78$ ) was somewhat lower than Nunally's "ideal" threshold. However, it can still be considered as reasonably high, especially as it was substantially above .70, which is the level Nunally suggests as the minimum acceptable level for basic research. The  $\alpha$ -values obtained were comparable with the values reported in previous studies (see Table 7-5).

The  $\alpha$ -values of the satisfaction scales based on multiattribute measures were substantially lower than those found in previous studies (Churchill and Surprenant

1982). They were .77, .77 and .43 for SLDA, SDTA and SBELA<sup>2)</sup>, respectively. To understand these results, one has to keep in mind that the items in these scales were supposed to measure consumers' satisfaction with a number of different attributes of a service or product. For example, item A may measure consumers' satisfaction with the response time of a service, whereas item B is designed to measure consumers' satisfaction with the range of transactions available. A customer may be very satisfied with attribute A and, at the same time, very dissatisfied with attribute B (Oliver and Bearden 1985). Hence, a multiattribute satisfaction measure combining items A and B may not produce high  $\alpha$ -values<sup>3)</sup>.

There may be two reasons why many researchers actually encountered high Cronbach Alpha-values with their multiattribute based measures. One potential reason is that a number of attributes were manipulated in the same direction. For example, in Churchill and Surprenant's (1982) study all attribute performances were either high, medium or low. Therefore, it is not surprising that the  $\alpha$ -value of the attribute-based overall disconfirmation or satisfaction measure was high. However, if the attributes are not manipulated in the same direction, one should not expect a high  $\alpha$ -value across the attribute measures, as they are each designed to measure different constructs, i.e. different attributes. Rather, for analytical purposes, marketers would want these multiattribute models to actually help them determine the attribute(s) that cause(s) a good or bad satisfaction outcome. The

---

<sup>2</sup> SLDA, SDTA and SBELA are the unweighted versions of SLDW, SDTW and SBELW respectively.

<sup>3</sup> The same argument also applies to the multiattribute based disconfirmation measures DA1 and DA2. DA1 and DA2 are the unweighted versions of DOW1 and DOW2 respectively.

second reason for high  $\alpha$ -values may be strong halo effects. For example, if someone is very satisfied with attribute A, he/she may also feel more satisfied with attributes B, C and D.

In conclusion, it seems surprising that researchers have actually computed and evaluated Cronbach Alpha values of multiattribute based overall disconfirmation and satisfaction scales. Instead, the reliability of these scales should have been evaluated on their convergent validity with other overall measures. Furthermore, one would want these measures to have low halo effects, which in fact means lower  $\alpha$ -values, so that the attribute(s) that cause(s) disconfirmation and satisfaction/dissatisfaction can be identified.

In this study, the expectations and performance of only one single attribute was manipulated. One would therefore expect that the  $\alpha$ -values of multiattribute based measures to be lower than in studies in which all or several attributes were manipulated in the same direction.

Furthermore, the obtained multiattribute measures differed considerably in their  $\alpha$ -values. In particular, SBELA ( $\alpha = .43$ ) showed a far lower  $\alpha$ -value than all of the other multiattribute scales. The question thus arises as to whether SBELA is a better measure to pin-point the attribute, which was response time in this study, that causes the disconfirmation and satisfaction/dissatisfaction outcomes. SBELA was the only Likert-type scale; all other multiattribute measures used semantic differential scales. Furthermore, it was also the only measure whose items were

mixed with another scale, namely with SWLS12. The questions of whether SBELA discriminates better between satisfaction with different attributes, and whether that is due to the type of scale or to its presentation in the questionnaire are not within the scope of this thesis. However, the results obtained provide a stimulus for further analysis of the data and future research.

In summary, scales of PLEAS1, PLEAS2, AROUS1, AROUS2, SWLS12 and SVERBAL showed high reliability. The reliability of multiattribute based disconfirmation and satisfaction measures cannot be assessed using Cronbach Alpha values because their underlying questions measure different constructs, i.e. the disconfirmation-of-expectation and satisfaction with different attributes. Other methods should be used to test their reliability.

Table 7-5: Cronbach Alpha of All Multi-Item Scales

Measure	No Items	N	Standard Item Alpha	Hui (1988)	Russell and Pratt (1980)
<u>Affect</u>					
AROUS1	6	129	.78	.65 to .88	-
AROUS2	10	128	.80	-	.86
PLEAS1	6	129	.89	.92 to .93	-
PLEAS2	10	128	.85	-	.93
				Churchill and Surprenant (1982)	Westbrook and Oliver (1981)
<u>Disconfirmation</u>					
DA1	6	130	.73	-	-
DA2	6	130	.65	.81 and .85	-
<u>Satisfaction</u>					
SLDA	6	130	.77	.91 and .94	-
SDTA	6	130	.77	-	-
SBELA	6	130	.43	.87 and .94	-
SWLS12	12	130	.91	-	.93
SVERBAL	3	130	.87	-	.72

## 2.4 Convergent and Discriminant Validity

As mentioned in Chapter VI, one important criterion for the validity of any experiment is how well the measures employed correspond to the concepts they purport to measure (Calder et al 1981). Indicators for construct validity are convergent, discriminant and nomological validity. Convergent validity refers to whether one method of measuring a construct agrees with other methods of measuring the same construct. In other words, the scores obtained for the same construct should be independent of the applied method of measurement. To facilitate an assessment of the extent of variance caused by the method of measurement, Zaltman et al (1973) recommend the use of maximally different methods.

Discriminant validity refers to the extent to which one construct differs from other constructs. Constructs can be invalidated by too high correlations with other constructs from which they were intended to differ (Campbell and Fisk 1959). Finally, nomological validity addresses the theoretical relationships of the constructs of interest. As in the context of this study nomological validity is basically concerned with hypotheses testing, only convergent and discriminant validity are utilized to assess the construct validity of the measures employed.

The traditional way of determining convergent and discriminant validity is Campbell and Fisk's (1959) multitrait-multimethod analysis. The basic idea of this analysis is that when every construct is measured by at least two different

measures, the correlations between the different measures of the same construct represents the convergent validity, whereas measures of different constructs obtained through the same method of measurement represents the discriminant validity. Obviously, the correlations between different measures of the same construct should be higher than their correlations with measures of all other constructs.

In this study, every construct was operationalized in at least two different ways. Pleasure, for example, was measured on a six-item semantic differential scale and a ten-item Likert-type scale. As a result, an evaluation of the convergent and discriminant validity of the operationalizations was possible.

In Table 7-6 the correlation coefficients between all measures employed are presented. Several aspects of these correlations are worth noting. First, most of the measures appeared to have good convergent and discriminant validity. Their correlations with other measures of the same construct are higher than their correlations with measures of different constructs. This was true for the following measures: Response time disconfirmation (DTIME1 and DTIME2), arousal (AROUS1 and AROUS2) and pleasure (PLEAS1 and PLEAS2). Exceptions were the measures of overall disconfirmation and satisfaction, which all had correlations with measures of other constructs that were equal to or greater than their correlations with other measures of the same construct.

The overall disconfirmation measure DOW2 had a lower correlation ( $r=.64$ ) with DOQ1, another measure of overall disconfirmation, than with the measures of response time disconfirmation DTIME1 and DTIME2 ( $r=.66$  and  $.67$  respectively). Similarly, DOW1 was as highly correlated with DTIME1 as it is with DOQ2 ( $r=.68$ ). Both overall disconfirmation measures (DOW1 and DOW2) had in common the fact that they were based on a linear compensatory multiattribute scale, in which response time<sup>4</sup> was the only manipulated attribute. One would therefore expect the measures of response time disconfirmation (DTIME1 and DTIME2) and the multiattribute based overall disconfirmation measures (DOW1 and DOW2) to be highly correlated.

The two overall disconfirmation measures, DOQ1 and DOQ2, based on subjects' own assessments of experienced overall disconfirmation were highly correlated ( $r=.85$ ) and had lower correlations with the measures of other constructs. It therefore appears that subjects' own evaluations are a better measure of overall disconfirmation than the multiattribute based scores. In further analyses, DOQ1 and DOQ2 were used as operationalizations of overall disconfirmation.

There appeared to be some validity problems with the overall satisfaction measures. The three multiattribute based measures (SLDW, SDTW and SBELW) were more highly correlated with the multiattribute based disconfirmation measures (DOW1 and DOW2) than with a number of other satisfaction measures. Similar to multiattribute based disconfirmation measures, these overall measures

---

<sup>4</sup> Respectively, DTIME1 and DTIME2 were used as one of DOW1 and DOW2's six attributes.

were computed from answers to questions about subjects' satisfaction with single attributes and each attribute's importance. These measures can be very useful for analytical purposes, although scales that use direct questions about overall satisfaction probably capture the true underlying satisfaction construct better. As with the disconfirmation measures, only overall satisfaction measures based on subjects' own overall evaluations were used in further analyses.

Of the remaining satisfaction measures, three (SLDQ, SDTQ and SD) were single-item scales and two (SWLS12 and SVERBAL) were multi-item scales. The single-item scales showed less convergent validity ( $r=.64$ ,  $.54$  and  $.67$ ) than the multi-item scales ( $r=.83$ ). A high convergent validity of SWLS12 and SVERBAL was obtained despite the measures used being very different: SVERBAL was a three-item measure consisting of a 7-point delighted-terrible scale, an 11-interval percentage scale and an 11-item behaviour tendency scale, whereas SWLS12 was a 12-item Likert-type scale.

Furthermore, one of the three single-item measures (SD) showed problems with its discriminant validity. SD was more highly correlated with both pleasure measures (PLEAS1,  $r=.63$ ; PLEAS2,  $r=.57$ ) than with other satisfaction measures (SLDQ,  $r=.54$ ; SWLS12,  $r=.63$ ). It was somewhat surprising that SD seems to lack convergent and discriminant validity because it has been one of the most frequently employed measures in effects application research.

From all satisfaction measures SWLS12 and SVERBAL showed the highest convergent and discriminant validity and were therefore employed for the hypothesis testing.

In summary, the following measures had good convergent and discriminant validity: DTIME1, DTIME2, DOQ1, DOQ2, AROUS1, AROUS2, PLEAS1, PLEAS2, SWLS12 AND SVERBAL.

**Table 7-6: Pearson Correlation Coefficients, 127 < n < 132, p < .001 for all r > .28; p < .01 for all r > .20; p < .05 for all r > .14**

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
1 MEXPCT	1																					
2 MVAR	-	1																				
3 MPERF	-	-	1																			
4 DTIME1	.44	-	-.49	1																		
5 DTIME2	.52	.17	-.43	.93	1																	
6 DOQ1	.42	-	-.35	.58	.59	1																
7 DOQ2	.39	-	-.42	.59	.61	.85	1															
8 DOW1	.35	-	-.34	.68	.60	.70	.68	1														
9 DOW2	.32	-	-.36	.66	.67	.64	.70	.83	1													
10 AROUS1	-.16	-	-	-	-	.21	.23	.23	.20	1												
11 AROUS2	-	.14	-.17	.16	.15	.36	.38	.40	.36	.66	1											
12 PLEAS1	-	-	-.20	.26	.27	.42	.43	.40	.35	.42	.54	1										
13 PLEAS2	-	-	-	.23	.25	.36	.33	.37	.30	.28	.44	.73	1									
14 SLDQ	.19	.14	-.20	.37	.42	.54	.51	.55	.53	.42	.43	.60	.57	1								
15 SDTQ	-	-	-.16	.23	.27	.40	.43	.35	.37	.42	.40	.61	.58	.64	1							
16 SD	-	-	-.19	.26	.29	.36	.40	.41	.47	.33	.36	.63	.57	.54	.67	1						
17 SLDW	-	-	-.20	.43	.43	.47	.42	.59	.64	.44	.44	.53	.46	.74	.56	.56	1					
18 SDTW	-	-	-.25	.43	.40	.40	.37	.59	.63	.41	.43	.59	.47	.65	.58	.61	.84	1				
19 SBELW	-	-	-.22	.41	.39	.39	.35	.55	.62	.38	.43	.54	.48	.61	.52	.55	.82	.79	1			
20 SWLS12	-	-	-.19	.36	.36	.48	.49	.51	.49	.46	.57	.71	.69	.71	.68	.63	.65	.59	.65	1		
21 SVERBAL	-	-	-.25	.31	.33	.49	.50	.48	.47	.43	.48	.67	.64	.66	.75	.71	.60	.58	.60	.83	1	

## 2.5 Manipulation Check

In any experimental research it is important to check whether the experimental treatments have been successful. A correlation analysis and analysis of variance were employed to assess whether the three experimental manipulations were effective.

### 2.5.1 Correlation Analysis

The correlations between the experimental conditions and the corresponding perception measures all had high magnitudes (refer to Table 7-7) and were significant at the .001 level. The correlations between the expectation manipulation and expected performance were higher ( $r=.83$  and  $.83$ ) than the correlations between the performance manipulation and perceived performance ( $r=.61$  and  $.57$ ). There are probably two reasons for this difference. First, all subjects were informed in the instructions about the precise response time of the homebanking system, and thereafter the "Simulator" performed exactly as described in those instructions. When they used the "Real Homebanking System," however, subjects had to rely on their individual time perceptions. Second, there may have been assimilation effects between expectations and perceived performance. This proposition was further examined using analysis of variance.

The correlations between manipulated variance and perceived variance ( $r = .38$  and  $.55$ ) were somewhat lower than those between the other manipulations and their checks. This may have been due to two causes. First, it may have been the result of the weaker experimental manipulation of variance in comparison to the manipulations of expectations and performance. The variance was manipulated only through a written service description, whereas expectations and performance were manipulated by written description and were further reinforced when the subjects experienced the "Simulator" response time (reinforcement of expectations) and the "Real Homebanking Service" (experience of performance). Second, it is possible that some subjects based their variance perceptions not only on the service description, but also on the "Simulator" performance, which did not show any variance in response time.

In summary, all correlations were highly significant and the magnitudes obtained indicate that all manipulations were successful. However, the variance manipulation seems to be the weakest of the three experimental conditions. In the next section, results of analyses of variance are reported supporting the findings of the correlation analysis.

Table 7-7: Pearson Correlation Coefficients between Experimental Conditions and Perception Measures, n=132, p<.001

---

	<b>MEXPECT</b>	<b>EPERF1</b>
<b>EPERF1</b>	<b>.83</b>	
<b>PPERF2</b>	<b>.83</b>	<b>.82</b>
	<b>MVAR</b>	<b>EVAR1</b>
<b>EVAR1</b>	<b>.38</b>	
<b>EVAR2</b>	<b>.55</b>	<b>.45</b>
	<b>MPERF</b>	<b>PPERF1</b>
<b>PPERF1</b>	<b>.61</b>	
<b>PPERF2</b>	<b>.57</b>	<b>.86</b>

---

## 2.5.2 Cross-Tabulations and Analyses of Variance

Results obtained from a three-way analysis of variance on expected performance (Table 7-8) revealed a strong expectation manipulation main effect [ $F(1,124)=271, p<.001$ ]. No other significant main or interaction effect was found. The cell means were also as expected. In the fast response time condition, subjects expected a short response time (cell means between 2.3 and 2.8 seconds), and in the slow condition, subjects expected a long response time (cell means between 10.8 and 12.1 seconds). It could therefore be concluded that the manipulation of expectations was effective.

Table 7-8: Mean Ratings of Expected Performance across Experimental Conditions

---

Expected Performance (in seconds)

<u>Cell Means:</u>		MPERF= Fast		MPERF= Slow			
	MVAR:	Low	High	Low	High		
	Fast	2.3	2.8	2.4	2.8		
MEXPECT:	Slow	10.8	11.6	12.1	11.3		
 <u>Adjusted Means:</u>		MEXPECT=1	-4.5	MVAR=1	-.1	MPERF=1	-.2
		=2	4.4	=2	.1	=2	.1
 <u>Main Effects:</u>		MEXPECT:	F(1,124)=271.6; p<.001		Variance Explained: 68.3%		
		MVAR:	F(1,124)= .1; n.s.				
		MPERF:	F(1,124)= .3; n.s.				
No interaction effect reached significance.							

---

The analysis of variance on expected variance showed two significant main effects (Table 7-9). First, the variance manipulation showed a significant difference [ $F(1,124)=57.5$ ;  $p<.001$ ] in the intended direction confirming the successful manipulation of expected variance. Second, the expectation manipulation also had a significant main effect [ $F(1,124)=7.3$ ;  $p=.008$ ], although it was much smaller and explained only 3.8% of the variance. A cross-tabulation revealed that subjects expected more variance when they also expected a long response time, and that they expected less variance when they expected a short response time. It seems logical that people can expect a greater variance the longer the mean response time, because their variance perception of a long response time will be higher than that of a short response time. This means, a person will "know" that a reoccurring response time of 1 second shows little variance, whereas a person might well perceive some variance in a reoccurring response time of 12 seconds. 1 and 12 seconds were the times the subjects experienced when using the "Simulator", after which they were questioned about their expected variance of the response time of the real service. As the F-value of the expectation main effect was small and the explained variance negligible, the impact of the expectation main effect was not examined further.

The cell means were as expected and also indicated a successful manipulation. Subjects expected a higher possible difference in the response time performance in the high variance condition (cell means between 9.9 and 14.2 seconds) than in the low variance condition (cell means between 3.0 and 6.6 seconds).

Table 7-9: Mean Ratings of Expected Variance across Experimental Conditions

---

Expected Variance (in seconds)

Cell Means:

		MPERF = Fast		MPERF = Slow	
MVAR:		Low	High	Low	High
MEXPECT:	Fast	4.6	10.8	3.0	9.9
	Slow	5.2	14.2	6.6	11.8

Adjusted Means:

MEXPECT=1	-1.2	MVAR=1	-3.4	MPERF=1	.4
=2	1.2	2	3.5	=2	-.4

Main Effects:

MEXPECT:	F(1,124) = 7.3; p=.008	Explained Variance: 3.8%
MVAR:	F(1,124) = 57.5; p<.001	Explained Variance: 29.9%
MPERF:	F(1,124) = .9; n.s.	

No interaction effect reached significance.

---

Finally, the analysis of variance on perceived performance revealed two significant main effects (Table 7-10). First, the performance manipulation had a large and significant main effect [ $F(1,124)=105.9$ ;  $p<.001$ ], which showed that the manipulation was effective. Second, the expectation manipulation also showed a significant main effect [ $F(1,124)=49.6$ ;  $p<.001$ ]. The F-values and explained variance figures show that the expectation main effect was much weaker than the performance manipulation.

There are two potential explanations for the expectations manipulation main effect. First, it might have been an example of consistency effects in the satisfaction process. When evidence is ambiguous, people tend to encode information according to prior knowledge and expectations. If response time perceptions showed some ambiguity, the main effects could be due to consistency effects.

Second, subjects had to translate their internal experience of the time interval into a physical measure, namely seconds. Apart from using their experience and knowledge about time intervals in general, they might have used previous experience with the homebanking "Simulator" as an anchor value to translate their internal time perception into seconds. If this was the case, the expectations main effect could be explained as measurement error of the internal time perception rather than a consistency effect.

A consistency effect would reduce the level of disconfirmation experienced by the subjects, and a measurement error would have no impact on disconfirmation. The purpose of manipulating expectations and perceived performance was to create disconfirmation experiences for testing the effect of expected variance on disconfirmation and the effect of disconfirmation on arousal and pleasure. Independent of the cause of the expectation main effect, the standard deviation of the disconfirmation with response time still seemed sufficiently high for the proposed hypotheses to be tested (means 4.0, standard deviations 1.98 and 2.02 respectively for DTIME1 and DTIME2, refer to Table 7-2).

In addition to the large and highly significant main effect of the performance manipulation on perceived performance, the cell means also indicated an effective manipulation of performance. The cell means for perceived performance varied between 3.1 and 6.8 seconds in the fast response time condition, and between 7.5 and 11.7 in the slow response time condition.

In summary, analyses of variance, subsequent F-tests and the cross-tabulations showed that all three manipulations (expected performance, expected variance and perceived performance) were successfully effected. Two further main effects, which could be interpreted, were not large enough to obscure the intended manipulations.

Table 7-10: Mean Ratings of Perceived Performance across Experimental Conditions

---

Perceived Performance (in seconds)

<u>Cell Means:</u>		MPERF= Fast		MPERF= Slow	
	MVAR: Low	High		Low	High
	Fast	3.1	3.1	7.5	8.3
MEXPECT:	Slow	6.8	5.9	11.7	11.3

<u>Adjusted Means:</u>	MEXPECT=1	-1.7	MVAR=1	.0	MPERF=1	-2.5
	=2	1.7	=2	-.0	=2	2.5

<u>Main Effects:</u>	MEXPECT:	F(1,124)= 49.6; p<.001	Explained Variance: 17.7%
	MVAR:	F(1,124)= .0; n.s.	
	MPERF:	F(1,124)=105.9; p<.001	Explained Variance: 37.7%

No interaction effect reached significance.

---

## 2.6 Summary

Chi-square tests confirmed that the subjects were randomly allocated to the experimental conditions. Therefore, all rival hypotheses that observed effects on dependent variables could have been due to prior group differences associated with the background variables controlled for were rejected.

The reliability and validity of the measures employed were examined. The measures listed in Table 7-11 were used in further data analyses. They all had good convergent and discriminant validity, and all multi-item scales selected showed good reliability.

Correlation analyses, analyses of variance and cross-tabulations confirmed that all three experimental manipulations had been successful.

Table 7-11: Measures Selected for Hypothesis Testing

---

<u>Variables</u>	<u>Selected Measures</u>
Disconfirmation with Time Performance	DTIME1 and DTIME2
Overall Disconfirmation	DOQ1 and DOQ2
Arousal	AROUS1 and AROUS2
Pleasure	PLEAS1 and PLEAS2
Overall Satisfaction	SWLS12 and SVERBAL

---

### 3 Hypotheses Testing I

The hypotheses advanced in Chapters II, III and V are listed in Table 7-12. These hypotheses were tested using structural equation modelling (Lisrel) and traditional statistical techniques. In this section the details of traditional hypothesis testing are presented. Correlation analysis was used in an initial examination of all hypothesized linear relationships between variables (Hypotheses I, IV & VI). More extensive testing, conducted with a Lisrel analysis, is reported in section 4 of the chapter.

Both Lisrel and correlation analyses can only be used to test linear relationships. However, a number of hypothesized relationships were non-linear and, therefore, other statistical tests had to be used. The tests of these hypotheses are presented after the correlation analysis and in the following order: the effect of expected variance on disconfirmation (Hypotheses II & III), of disconfirmation on arousal (Hypothesis V), and of arousal and pleasure on satisfaction (Hypothesis VII).

Table 7-12: Hypotheses to be Tested

---

No	Hypothesis
----	------------

---

**Perceived Performance as a Direct Causal Antecedent of Satisfaction?**

I      Direct causal links between perceived performance and satisfaction do not exist; high positive correlations between the two constructs are merely an indication of inappropriate perceived performance measures.

**The Impact of Expected Heterogeneity in Performance on Disconfirmation**

II     In situations where perceived performance and the expected mean performance are approximately equal, consumers experience confirmation independent of the expected variance in performance.

III    In disconfirmation situations, consumers who hold mean performance expectations with some degree of variation around the mean, experience smaller magnitudes of disconfirmation than consumers who hold point expectations.

**The Role of Affect in the Satisfaction Process**

IV     The degree of pleasure experienced in a service encounter is an increasing function of the perceived disconfirmation-of-expectations.

V      The level of arousal experienced in a service encounter is an increasing function of the perceived magnitude of disconfirmation-of-expectations.

VI     Satisfaction is an increasing function of the pleasure experienced during the service consumption process.

VII    The strength of the impact of pleasure on satisfaction is an increasing function of the level of arousal experienced during the service consumption process.

---

For the traditional statistical tests, one measure of each theoretical construct was employed and the selection criterion was the Cronbach Alpha value of the multi-item scale. Mehrabian and Russell's (1974) scales of affect (AROUS1 and PLEAS1) together have a slightly higher  $\alpha$ -value ( $\alpha=.84$ ) than Russell and Pratt's (1981) scale ( $\alpha=.83$ ), and the  $\alpha$ -value of SWLS12 ( $\alpha=.91$ ) is slightly higher than that of SVERBAL ( $\alpha=.87$ ).

There were two single item measures for each of response time disconfirmation (DTIME1 and DTIME2) and overall disconfirmation (DOQ1 and DOQ2). As the reliability of any measure can be increased substantially if further items are added to a scale (Nunnally 1981), the two measures for each construct were combined to form a new scale for response time disconfirmation (DTIME) and a new scale for overall disconfirmation (DOQ). The  $\alpha$ -values of the new scales were very high (.96 and .92 for DTIME and DOQ, respectively).

In summary, the following five operationalizations were selected for traditional hypothesis testing (refer to Table 7-1 for a detailed description of all operationalizations employed in this study): DTIME, DOQ, AROUS1, PLEAS1 and SWLS12.

### 3.1 Correlation Analysis (Hypotheses I, IV & VI)

The correlation coefficients between variables with a hypothesized linear relationship were examined (Table 7-13).

Disconfirmation<sup>5</sup> was positively correlated to expectations ( $r=.41$ ;  $p<.001$ ) and negatively correlated to performance ( $r=-.39$ ;  $p<.001$ ). These results were in line with the classical disconfirmation-of-expectations model and provide some nomological validity for the variables.

There was no significant correlation between expectations and satisfaction, and performance had only a relatively low correlation with satisfaction ( $r=.19$ ;  $p<.05$ ). Disconfirmation had a far higher correlation with satisfaction ( $r=.49$ ;  $p<.001$ ) than performance. According to the assumption of the weak causal order closure, if a sequence of variables is intercorrelated then the correlations between the two preceding variables should be higher than the correlations between the two succeeding variables. This means the correlation coefficients obtained tentatively supported the hypothesis that a direct causal link did not exist between performance and satisfaction (Hypothesis I), and that the effects of performance and expectations were mediated through the disconfirmation-of-expectations variable.

---

<sup>5</sup> Disconfirmation is defined as being negative for Expectations > Performance, and positive for Expectations < Performance.

As predicted, disconfirmation was positively correlated to pleasure ( $r=.43$ ;  $p<.001$ ) and pleasure was positively correlated to satisfaction ( $r=.71$ ;  $p<.001$ ; Table 7-13). These correlations provide support for the hypotheses that positive disconfirmation causes pleasure (Hypothesis IV) and that pleasure causes satisfaction (Hypothesis VI).

Table 7-13: Pearson Correlation Coefficients for Overall Measures;  
 $128 < n < 132$ ;  $p < .001$ ,  $^*)p < .01$ ,  $^{**})p < .05$

	1	2	3	4	5	6	7
1 MEXPECT	1.000						
2 MVAR	-	1.000					
3 MPERF	-	-	1.000				
4 DOQ	.41	-	-.39	1.000			
5 AROUS1	-.16 <sup>**</sup> )	-	-	.22 <sup>*)</sup>	1.000		
6 PLEAS1	-	-	-.20 <sup>*)</sup>	.43	.42	1.000	
7 SWLS12	-	-	-.19 <sup>**</sup> )	.49	.46	.71	1.000

### 3.2 The Impact of Variance Expectations on Disconfirmation

#### (Hypotheses II & III)

Analysis of variance was used to test the hypothesized three-way interaction between variance expectations, performance expectations and performance on disconfirmation (Hypotheses II & III).

As shown in Table 7-14, expectations and performance both showed significant main effects on disconfirmation [ $F(1,123)=32.6$ ,  $p<.001$  and  $F(1,123)=29.4$ ,  $p<.001$ ]. These results were in line with the classical disconfirmation-of-expectations model.

The variance main effect and all interaction terms were very small and far from significant which suggested rejection of Hypotheses II and III. Alternatively, it might have been the case that the overall disconfirmation measure (DOQ) was not sensitive enough to capture the impact of the variance manipulation. Therefore, the analysis of variance was repeated with the time-specific disconfirmation measure (DTIME).

Table 7-14: Anova of Experimental Conditions on Overall Disconfirmation

---

**Overall Disconfirmation**

**Cell Means:**

	MPERF= Fast		MPERF= Slow		
	MVAR: Low	High	Low	High	
	MEXPECT:	Fast	4.3	4.6	3.1
	Slow	5.3	5.4	4.3	4.6

**Adjusted Means:**

	MEXPECT=1		MVAR=1		MPERF=1	
	=2	-.55	=2	-.10	=2	.51
		.52		.11		-.51

**Main Effects:**

MEXPECT:	F(1,123) = 32.6; p<.001	Variance Explained: 17.4%
MVAR:	F(1,123) = 1.3; n.s.	
MPERF:	F(1,124) = 29.4; p<,001	Variance Explained: 15.7%

No interaction effect reached significance.

---

Results are presented in Table 7-15. There was a significant main effect of variance on disconfirmation [ $F(1,123)=5.1, p=.03$ ] and a three-way interaction effect between expectations, performance and variance [ $F(1,123)=4.3, p=.04$ ].

Examination of the cell means showed that when there was a large discrepancy between manipulated expectations and performance, the subjects experienced high levels of disconfirmation, and variance did not seem to have any impact on that level. The cell means for the low and high variance manipulation were both 5.8<sup>6</sup> in the positive disconfirmation condition (the subjects expected a slow performance and were exposed to a fast performance). In the negative disconfirmation condition the cell means were 2.0 and 2.1 for the low and high variance manipulation, respectively. The subjects appear to have experienced very similar levels of disconfirmation independent of the variance manipulation.

However, when the discrepancy between manipulated expectations and performance was small, the variance manipulation had a significant impact on the level of disconfirmation experienced. In these circumstances, subjects experienced a more positive disconfirmation in the high variance condition than in the low variance condition. In the fast expectation and fast performance condition, the cell mean for the low variance manipulation was 3.5, whereas it was 4.7 for the high variance manipulation. The same effect was obtained in the slow expectation

---

<sup>6</sup> Disconfirmation was measured on 1 to 7 semantic differential scales (DTIME1 and DTIME2); 1= the response time was far worse than expected, 4= the response time was just as expected, and 7= the response time was a lot better than expected.

and slow performance condition. Here the cell means were 3.6 and 4.7 for the low and high variance manipulations, respectively.

Table 7-15: Anova of Experimental Conditions on Response Time Disconfirmation

---

Response Time Disconfirmation

Cell Means:

		MPERF= Fast		MPERF= Slow	
		MVAR: Low	High	Low	High
MEXPECT:	Fast	3.5	4.7	2.0	2.1
	Slow	5.8	5.8	3.6	4.7

Adjusted Means:

MEXPECT=1	-.97	MVAR=1	-.28	MPERF=1	.92
=2	.93	=2	.29	=2	-.91

Main Effects:

MEXPECT:	F(1,123)= 57.9; p<.001	Variance Explained: 23.7%
MVAR:	F(1,123)= 5.1; p=.03	Variance Explained: 2.1%
MPERF:	F(1,123)= 53.4; p<.001	Variance Explained: 21.9%
3-Way Interaction F(1,123)= 4.3; p=.04		Variance Explained: 1.8%
No other interaction effect reached significance.		

---

The results of the analysis of variance were the opposite of those expected. Hypothesis II predicted no variance main effect on disconfirmation when performance and expectations are approximately equal, and Hypothesis III predicted a variance main effect in disconfirmation situations. A T-test was conducted to test the two hypotheses directly (Table 7-16). First, the test confirmed the impression obtained from the analysis of variance that the variance manipulation had no significant impact (T-value =  $-0.19$ ; d.f. = 63; n.s.) on response time disconfirmation experienced in high disconfirmation situations (fast expectations and slow performance; slow expectations and fast performance). Therefore, Hypothesis III had to be rejected.

Second, the T-test showed that the variance manipulations had a highly significant effect (T-value =  $-3.07$ ; d.f. = 64;  $p = .003$ ) on disconfirmation experienced in low/no disconfirmation situations (fast expectations and fast performance; slow expectations and slow performance). Therefore Hypothesis II was rejected.

Table 7-16: T-Tests on the Impact of the Variance Conditions on Response Time

Disconfirmation

Experimental Conditions:		Cell Means of DTIME in Experimental Condition:		
Expectation	Performance	Low Variance	High Variance	T-Test
fast	fast	3.5	4.7	T-value= -2.34; d.f.=30; p=.03
slow	slow	3.6	4.7	T-value= -1.96; d.f.=32; p=.06
$\Sigma$ Low/No Disconfirmation Situations		3.6	4.7	T-value= -3.07; d.f.=64; p=.003
slow	fast	5.8	5.8	T-value= -.03; d.f.=31; n.s.
fast	slow	2.0	2.1	T-value= -.18; d.f.=30; n.s.
$\Sigma$ High Disconfirmation Situations		3.9	4.0	T-value= -.19; d.f.=63; n.s.

In conclusion, when performance was close to expected performance, subjects experienced more positive disconfirmation when they also expected a high variance in performance. However, when the discrepancy between expected and perceived performance was high, the variance manipulation had no impact on the level of disconfirmation experienced. These results were very interesting and are further explored in the concluding chapter. However, Hypotheses II and III predicted very different results and had to be rejected.

### 3.3 The Impact of Disconfirmation on Arousal (Hypothesis V)

It was hypothesized that arousal would be an increasing function of the magnitude of disconfirmation experienced. Regression analysis was employed, rather than analysis of variance, because the variables concerned were continuous.

After deducting the mean from the scores of overall disconfirmation DOQ (Howell 1987, p. 504), the hypothesized relationship between disconfirmation and arousal can be approximated by either of the following two equations:

$$[1] \quad \text{AROUS1} = \beta_0 + \beta_1(\text{DOQ})^2 + \text{error}$$

$$[2] \quad \text{AROUS1} = \beta_0 + \beta_1 |\text{DOQ}| + \text{error}$$

None of the beta values of the predictor variables in each of the two equations was significant ( $p < .10$ ) and, therefore, the experimental results did not support Hypothesis V.

Hypothesis V predicted that the level of arousal would be higher in both positive and negative disconfirmation situations than in confirmation situations. A simple tabulation of AROUS1 versus DOQ showed that this hypothesis had to be rejected. In negative disconfirmation situations the level of arousal was lower than in confirmation situations, rather than higher as hypothesized (cell means of 3.90 versus 4.11; Table 7-17). These results were already indicated in the correlation

analysis presented in Table 7-13 where disconfirmation was shown to be positively correlated with arousal ( $r=.22$ ;  $p<.01$ ).

In conclusion, Hypothesis V had to be rejected. Whether the observed significant correlation between disconfirmation and arousal indicated a linear causal relationship was further examined in the Lisrel analysis reported in section 4.

Table 7-17: Tabulation of Arousal with Confirmation

	Negative Disconfirmation DOQ < 3.5	Low and No Disconfirmation $3.5 \leq \text{DOQ} \leq 4.5$	Positive Disconfirmation $4.5 < \text{DOQ}$
AROUS1:	3.90	4.11	4.37
Cell Size:	25	54	50

### 3.4 The Effect of the Interaction of Arousal and Pleasure on Satisfaction

#### (Hypothesis VII)

The hypothesized effect of the interaction of arousal and pleasure on satisfaction (Hypothesis VII, Figure 5-2) was tested through multiple regression procedures because the variables included were continuous rather than discrete.

According to Jaccard et al (1990, p. 18), the most commonly used procedure for testing interaction effects with multiple regression methods involves the calculation of two  $R^2$  values, one on the basis of an additive (or "main effects") model, and the other on the basis of the main effects plus the product of the two predictor variables that are hypothesized to interact. If an interaction term is present, then the difference between the two  $R^2$  values should be significant. The formal significance test requires the test of the null hypothesis that the improvement of  $R^2$  through the inclusion of the interaction term is zero. An F-test was used.

Howell (1987, p. 503-4) suggests a transformation of the predictor variables, deviations from their means, to improve the interpretation of the main effects when interaction effects are present. Therefore, deviations from the mean rather than nominal scale values were used.

In a first equation for calculating  $R^2$  PLEAS1 and AROUS1 were the predictor variables, and SWLS12 was the criterion variable:

$$[1] \quad SWLS12 = \beta_0 + \beta_1(PEAS1) + \beta_2(AROUS1) + \text{error}$$

In a second equation a multiplicative term, the product of PLEAS1 and AROUS1, was included:

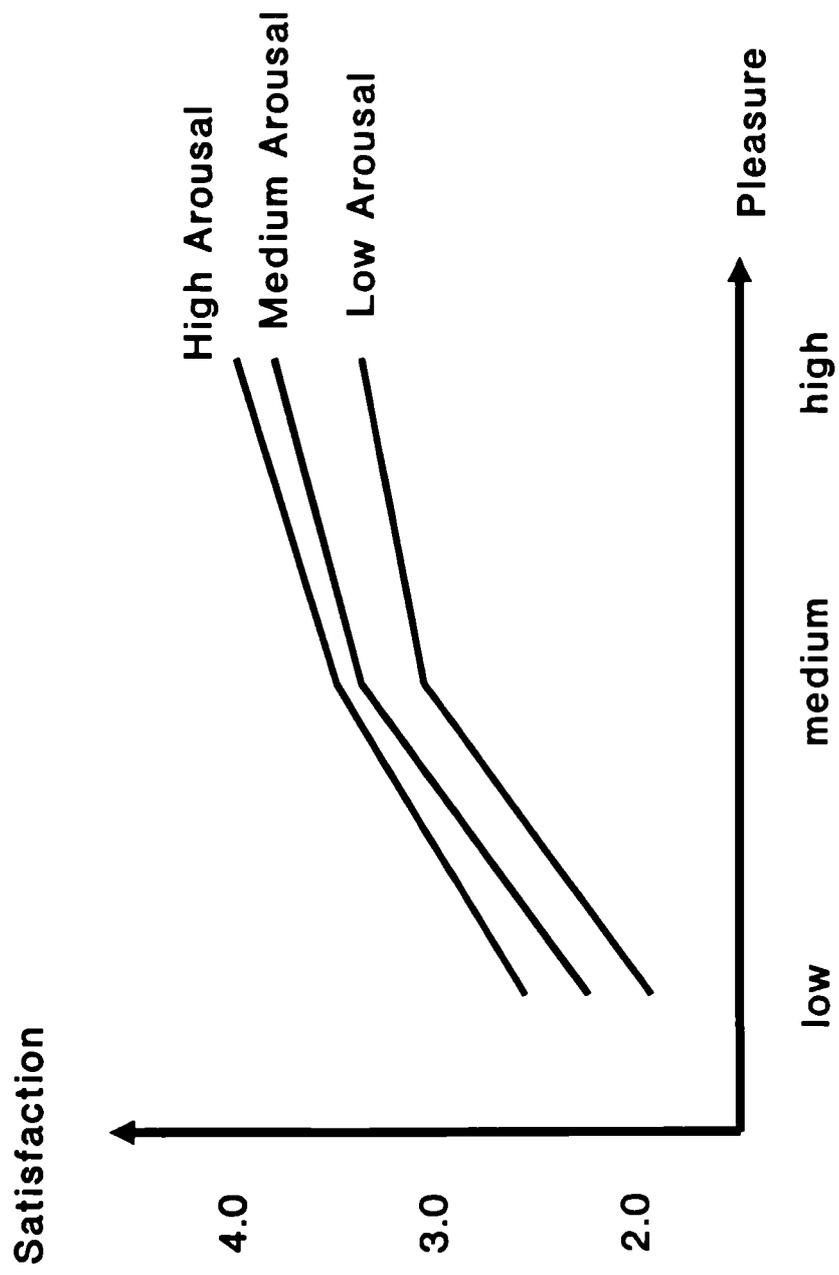
$$[2] \quad SWLS12 = \beta_0 + \beta_1(PEAS1) + \beta_2(AROUS1) + \beta_3(PEAS1 \times AROUS1) + \text{error}$$

The regression results led to rejection of Hypothesis VII (Table 7-18). No significant difference could be detected between the  $R^2$  values of equations [1] and [2] [ $F(1,128)=.519$ , n.s., for the calculation of the F-value refer to Jaccard et al 1990, p.19]. In a further test of Hypothesis VII, the arousal and pleasure scores were each divided into three categories (low, medium and high respectively) and their mean satisfaction scores were plotted (Figure 7-1). This plot also clearly shows that there was no effect of an interaction of arousal and pleasure on satisfaction. The hypotheses would have predicted a plot as shown schematically in Figure 5-2.

Table 7-18: Test of the Hypothesized Effect of an Interaction Between Pleasure and Arousal on Satisfaction

Criterion Variable	Predictor Variable	Beta	p	adj. R <sup>2</sup>
SWLS12	PLEAS1	.61	<.001	.507
	AROUS1	.20	=.004	
SWLS12	PLEAS1	.60	<.001	.509
	AROUS1	.20	=.004	
	PLEAS1 x AROUS1	-.08	>.10, n.s	

**Figure 7-1: Plot of the Effects of Arousal and Pleasure on Satisfaction**



Furthermore, the regression results in Table 7-18 showed highly significant positive beta values for pleasure ( $\beta_1=.61$  and  $\beta_1=.60$ ,  $p<.001$  for equations [1] and [2] respectively), consistent with the prediction of Hypothesis VI.

Although no direct impact of arousal on satisfaction was hypothesized, arousal had somewhat weaker but also significant positive beta values ( $\beta_2=.20$ ,  $p<.004$  for both equations). There were two potential explanations for this observation. First, it might have been due to a linear causal relationship between arousal and satisfaction. Second, the variation in both arousal and satisfaction might have been caused by a third variable, namely disconfirmation, which was significantly and positively correlated with both variables ( $r=.22$ ,  $p<.01$  and  $r=.49$ ,  $p<.001$  for arousal and satisfaction respectively, see Table 7-13). As Lisrel seemed particularly suited for examining linear causal relationships between three and more latent variables, this issue was further examined in the Lisrel analysis presented in section 4.

In conclusion, the predicted interaction effect between arousal and pleasure on satisfaction (Hypothesis VII) had to be rejected. Pleasure showed a strong and highly significant beta value which supported Hypothesis VI. Although not hypothesized, arousal also had a significant beta value. Whether arousal actually could be regarded as a direct causal antecedent of satisfaction was examined further in the structural equation analysis presented in section 4 of this chapter.

### 3.5 Summary

In this section, the results of the hypotheses testing with traditional statistical techniques were presented.

First, a correlation analysis tentatively supported that the effects of performance were mediated through disconfirmation and that no direct causal link exists between performance and satisfaction (Hypothesis I). A more conclusive test of the hypothesis is presented in the Lisrel analysis in section 4. The correlation analysis provided support to the hypotheses that positive disconfirmation causes pleasure (Hypothesis IV) and that pleasure causes satisfaction (Hypothesis VI).

Second, in low/no disconfirmation situations, subjects experienced more positive disconfirmation when they also expected a high variance in performance. In high disconfirmation situations, the variance manipulation had no impact on the level of disconfirmation experienced. These results were very interesting and were interpreted further in the conclusion chapter. However, Hypotheses II & III predicted very different results and had to be rejected.

Third, the hypothesis that arousal is an increasing function of the magnitude of disconfirmation had to be rejected (Hypothesis V). Whether the observed correlation between disconfirmation and arousal indicated a linear causal relationship was further examined in the Lisrel analysis in section 4.

Fourth, the predicted interaction effect between arousal and pleasure on satisfaction (Hypothesis VII) had to be rejected. In a regression analysis, an unexpected significant beta value for arousal on satisfaction was encountered. Whether arousal actually could be regarded as a direct causal antecedent of satisfaction was examined further in the Lisrel analysis in the next section.

#### 4 Hypotheses Testing II: Structural Equation Analysis

In section 3 of this chapter the hypothesized relationships were tested using traditional statistical techniques. However, these techniques have a number of shortcomings. In particular, examination of the correlation coefficients and analyses of variance cannot identify situations in which the observed correlation between two variables is due to the confounding effect of a third. Also, neither technique can be used to test the direction of causality, nor whether a variable has intervening effects on a causal relationship (e.g. pleasure on the relationship between disconfirmation and satisfaction). Finally, estimated regression parameters are often unreliable when the measurements contain measurement errors, and when the relationship of interest is among true or latent variables (Joereskog and Soerbom 1982). With structural equation modelling a hypothesized model can be tested as a whole. Confounding effects of variables, direction of causality and mediating effects can be detected, and measurement errors can be modelled.

As mediating effects had a central role in the hypotheses advanced, and as all independent variables were latent, structural equation modelling was used. First, a brief introduction to structural equation modelling is provided. Second, the hypothesized model is specified in equation form. Third and fourth, the fitting of the model is described and the results are discussed. Fifth, a number of rival models (hypotheses) are tested. A summary is provided at the end of the section.

#### **4.1 An Introduction to Structural Equation Modelling**

During the last fifteen years structural equation modelling has been increasingly employed in the social sciences. In marketing, scarcely an issue of the major journals is without one or more articles reporting results from structural equation models. The method is clearly a mainstay of multivariate statistical analysis in marketing (Bagozzi and Yi 1989).

Structural equation modelling addresses two basic issues of inference in the social sciences. First, it allows a realist perspective that causation exists only between theoretical variables rather than between observable measures of the theoretical constructs. Any observable relationship between the measures is merely a reflection of an unobservable theoretical relationship. Second, no measure can be a perfect operationalization of a theoretical construct, and hence it is important to consider measurement errors when testing causal relationships. Apart from its being designed to accommodate models that include latent variables and measurement errors, structural equation modelling also allows the modelling of reciprocal causation, simultaneity and interdependence between variables (Joereskog and Soerbom 1987).

A structural equation model has two parts: a measurement submodel and a structural equation submodel. The measurement submodel specifies how the latent variables are measured in terms of the observed variables, and it describes the measurement properties (validity and reliability) of the observed variables.

The structural equation submodel specifies the causal relationships among the latent variables. It describes the causal effects and estimates the amount of unexplained variance.

The equations of a hypothesized model are presented in matrix form. Elements of these matrices may be either free or fixed. The free elements are estimated using the maximum likelihood method. Joereskog and Soerbom (1989) have developed an iterative computer program for estimation which they called Lisrel. Lisrel calculates a value for each free parameter so that the estimated parameters are most similar to the observed variance-covariance matrix. Specifically, the value of the following function is minimized:

$$F = \log |\Sigma| + \text{tr}(S\Sigma^{-1}) - \log |S| - (m+n)$$

where  $S$  is the observed variance-covariance matrix of the measures,  
 $\Sigma$  is the estimated variance-covariance matrix of the measures<sup>7</sup>,  
 $m$  is the number of latent dependent variables, and  
 $n$  is the number of latent independent variables.

---

<sup>7</sup> This method is called Maximum Likelihood Estimation (ML). It provides standard errors for the parameter estimates, which is in contrast to the two other methods of calculating the parameter estimates, the Initial Estimates (IE) and the Unweighted Least Squares Estimates (ULS) (Joereskog and Soerbom 1982).

One advantage of using Lisrel is that the minimum value of  $F (F_0)$ , when multiplied by  $n$ , is approximately a  $\chi^2$  distribution with the following degrees of freedom:

$$\text{d.f.} = (m + n) (m + n + 1)/2 - t$$

where  $t$  is the number of estimated parameters.

A chi-square test is therefore readily available for examining the goodness-of-fit of the estimated parameters of the hypothesized model. As  $F$  is a positive function of the discrepancies between the elements of  $S$  and  $\Sigma$ , a low  $\chi^2$  indicates a good fit between  $S$  and  $\Sigma$ . It is generally accepted that  $\chi^2$  values of with probabilities .10 and higher indicate a good fit of one's data to the tested model.

Lisrel also provides an estimate of the standard error for every free parameter, which means that these parameters can be tested individually to see if they are significantly different from zero (t-test). Furthermore, Lisrel computes a modification index (MI), which determines the maximum reduction in  $\chi^2$  that can be achieved by relaxing a currently fixed element. Both the standard errors and modification index are good indicators of model misspecifications and can be helpful in respecifying the model and achieving an acceptable level of fit (Bone et al 1989, La Du and Tanaka 1989).

Despite the compelling advantages of Lisrel, a number of shortcomings need to be considered (Bagozzi 1981, p. 107-8). First, the model assumes interval scaled data. Although Lisrel can accommodate discrete exogenous variables, there are

problems of estimation when either a latent or observed dependent variable is binary. Second, Lisrel can only test linear relationships among variables. It may be inappropriate and misleading to employ linear structural equations to represent phenomena that are nonlinear in nature. Third, Lisrel assumes that observations are obtained from a multivariate normal distribution.

Lisrel is based on the variance-covariance matrix of the measures of the independent and dependent variables ( $x$ 's and  $y$ 's) included in the model structure. It can be shown that this matrix is a function of eight matrices that have to be specified for any structural equation analysis:  $\Lambda_x$ ,  $\Lambda_y$ ,  $\Gamma$ ,  $B$ ,  $\Phi$ ,  $\Psi$ ,  $\theta_\epsilon$  and  $\theta_\delta$ . Definitions of these matrices and all notations used in the following Lisrel analysis are provided in Table 7-17.

For a detailed introduction to structural equation or covariance structure modelling refer to one of the following books: Bagozzi (1981), Blalock (1985) or Hayduk (1987).

Table 7-19: Summary of all Notations Used in the Lisrel Analysis

Notation		Definition	Abbreviation
p		number of observed dependent variables	
q		number of observed <u>in</u> dependent variables	
m		number of latent dependent variables	
n		number of latent <u>in</u> dependent variables	
y		p x 1 vector of observed dependent variables	
$\eta$	eta	m x 1 random vector of latent dependent variables	
$\epsilon$	epsilon	p x 1 vector of measurement errors in y	
x		q x 1 vector of observed independent variables	
$\xi$	ksi	n x 1 random vector of latent independent variables	
$\delta$	delta	q x 1 vector of measurement errors in x	
$\Lambda_y$	lamda-y	p x m matrix of coefficients of the regression of y on $\xi$	LY
$\lambda_{y11}$	lamda-y <sub>11</sub>	element 11 of matrix LY	
$\Lambda_x$	lamda-x	q x n matrix of coefficients of the regression of x on $\eta$	LX
$\lambda_{x11}$	lamda-x <sub>11</sub>	element 11 of matrix LX	
$\Gamma$	gamma	m x n matrix of coefficients of the $\xi$ -variables in the structural relationship	GA
$\gamma_{11}$	gamma <sub>11</sub>	element 11 of matrix GA	
B	beta	m x m matrix of coefficients of the $\eta$ -variables in the structural relationship	BE
$\beta_{21}$	beta <sub>21</sub>	element 21 of matrix BE	
$\zeta$	zeta	m x 1 vector of equation errors in the structural relationships between the latent independent and latent dependent variables ( $\eta$ and $\xi$ )	
$\Phi$	phi	Cov( $\xi$ ), n x n covariance matrix of independent variables	PH
$\phi_{11}$	phi <sub>11</sub>	element 11 of matrix PH	
$\Psi$	psi	Cov( $\zeta$ ), m x m covariance matrix of equation errors	PS
$\psi_{11}$	psi <sub>11</sub>	element 11 of matrix PS	
$\Theta_\epsilon$	theta-epsilon	Cov( $\epsilon$ ), p x p covariance matrix of measurement errors in y	TE
$\theta_{\epsilon 11}$	theta-epsilon <sub>11</sub>	element 11 of matrix TE	
$\Theta_\delta$	theta-delta	Cov( $\delta$ ), q x q covariance matrix of measurement errors in x	TD
$\theta_{\delta 11}$	theta-delta <sub>11</sub>	element 11 of matrix TD	

## 4.2 Specification of the Hypothesized Model

Lisrel can only test additive linear relationships. All causal relationships suggested by the disconfirmation-of-expectations model are linear, i.e. the links from expectations and performance to disconfirmation, and the link from disconfirmation to satisfaction (Figure 2-2). Of the hypotheses advanced (Table 7-12), I, IV and VI predict linear causal relationships which could directly be included in the Lisrel model to be tested (Figure 7-2). The rationale for three further links included in this model is outlined in the following paragraphs.

Hypothesis V predicted that arousal would be an increasing function of the magnitude of disconfirmation. This hypothesis was rejected on the basis of results of the regression analysis presented in section 3.3. The data suggested a linear relationship between disconfirmation and arousal rather than the hypothesized (nonlinear) relationship between magnitude of disconfirmation and arousal. To test whether the linear relationship was causal, a link between disconfirmation and arousal was included in the model to be tested.

Hypothesis VII predicted that arousal would interact with pleasure in a way that an increase in arousal boosts the effect of pleasure on satisfaction (Figure 5-2). This hypothesis was rejected on the basis of the results of the multiple regression analysis and cross-tabulations reported in section 3.4. These analyses suggested that there was a linear relationship between arousal and satisfaction. To test

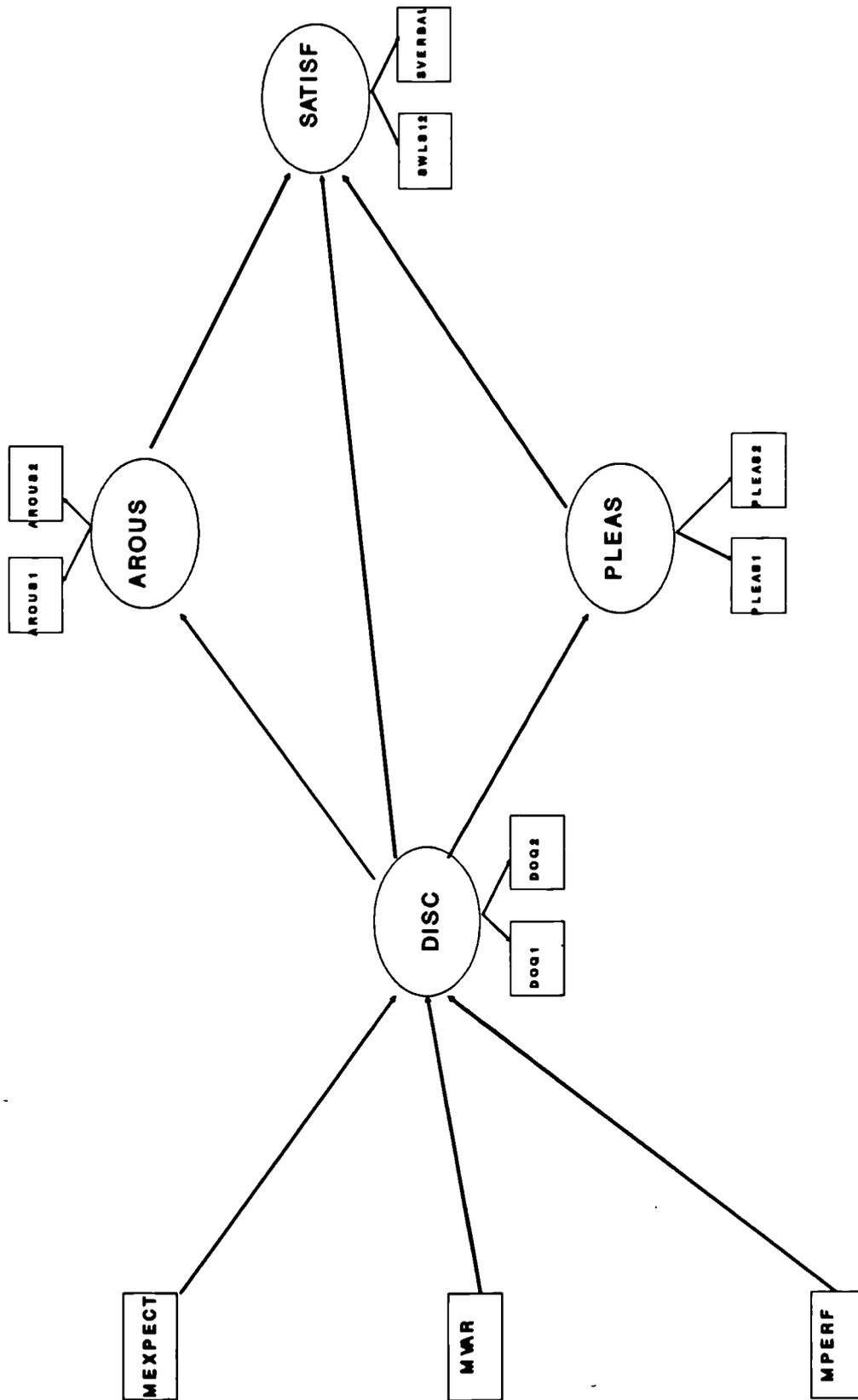
whether this relationship could be interpreted as causal it was also included in the Lisrel model.

The analysis of variance in section 3.2 showed that the variance manipulation had no significant effect on overall disconfirmation as measured by DOQ<sup>8</sup>). Nevertheless, to capture the effects of all manipulations, the variance manipulation was also included in the Lisrel model.

---

<sup>8</sup> This was in contrast to the effects of variance on disconfirmation of response time expectations (DTIME), where a significant main effect and a three-way interaction effect were observed (Table 7-16). Therefore, the overall measures were considered not sensitive enough to capture the impact of the variance manipulation.

Figure 7-2: Hypothesized Model



The structural equations that need to be defined for any Lisrel analysis are the following<sup>9</sup>:

Measurement Submodel: (1)  $y = \Lambda_y \eta + \epsilon$

(2)  $x = \Lambda_x \xi + \delta$

Structural Equation Submodel: (3)  $\eta = B\eta + \Gamma\xi + \zeta$

Covariance Matrices: (4)  $\Phi = \text{Cov}(\xi)$

(5)  $\Psi = \text{Cov}(\zeta)$

(6)  $\Theta_\epsilon = \text{Cov}(\epsilon)$

(7)  $\Theta_\delta = \text{Cov}(\delta)$

The specifications of the structural equations for the model shown in Figure 7-2 are presented in Table 7-20.

---

<sup>9</sup> Refer to Table 7-17 for the notations used.

Table 7-20: Specifications for the Hypothesized Model in Figure 7-2.

---

<u>Variable Specifications:</u>	$\xi_1 =$ Expected Performance Manipulation
	$\xi_2 =$ Expected Variance Manipulation
	$\xi_3 =$ Performance Manipulation
	$\eta_1 =$ Disconfirmation $y_1 =$ DOQ1
	$y_2 =$ DOQ1
	$\eta_2 =$ Arousal $y_3 =$ AROUS1
	$y_4 =$ AROUS2
	$\eta_3 =$ Pleasure $y_5 =$ PLEAS1
	$y_6 =$ PLEAS2
	$\eta_4 =$ Satisfaction $y_7 =$ SWLS12
	$y_8 =$ SVERBAL

Equation Specifications<sup>10)</sup>:

$$\begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 2 & 2 & 2 & 0 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} + \begin{bmatrix} 2 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} * \xi + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix} * \eta + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \\ e_8 \end{bmatrix}$$

$$\Psi = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

$$\Phi = \text{diag} [ 1 \ 1 \ 1 ]$$

$$\Theta_1 = \text{diag} [ 2 \ 2 \ 2 \ 2 \ 2 \ 2 \ 2 \ 2 ]$$

$$\Theta_2 = 0$$


---

<sup>10</sup> '0s' and '1s' are fixed parameters with the values of 0 and 1 respectively, and '2s' are parameters to be estimated.

### 4.3 Fitting of the Lisrel Model

Joereskog and Soerbom (1990) recommend that a model should be modified in steps, with changes made only to one parameter at a time and with inspection of the modification indices at each step. Application of this respecification procedure to the hypothesized model shown in Figure 7-2 is described in this section.

The hypothesized model  $M_0$ , depicted in Figure 7-2, produced a  $\chi^2$ -value of 41.1 with 41 degrees of freedom ( $p=.48$ ). This result indicated an excellent fit (Table 7-21), as the probability level was well above the generally accepted critical value of .10. In spite of this good fit, the modification index (MI) suggested that through the introduction of a causal path between expectations and arousal ( $\gamma_{21}$ ), the  $\chi^2$ -value could be further reduced by approximately 6.8.

A significance test is available for testing whether a modified model fits the data better than the original model. The  $\chi^2$  difference ( $\Delta\chi^2$ ) between two nested models (one model is part of another) is also a  $\chi^2$  distribution, with degrees of freedom equal to the difference between the degrees of freedom of the two models (Breckler 1990, Long 1983). Since the estimated  $\Delta\chi^2=6.8$ , with  $p<.01$ , the modification suggested by the modification index seemed to be a significant improvement over  $M_0$ .

The standardized residuals<sup>11)</sup> for the correlations between expectations ( $\xi_1$ ) and AROUS1 ( $y_3$ ), and between expectations ( $\xi_1$ ) and AROUS2 ( $y_4$ ) had values of -3.5 and -2.9, respectively. Furthermore, both residuals were clearly higher than the expected values as indicated by a Q-plot.

In conclusion, both the standardized residuals and the modification index located a potential misspecification of the model involving the two variables expectations ( $\xi_1$ ) and arousal ( $\eta_2$ ), which suggested the introduction of a causal link between these two variable. Therefore, model  $M_0$  was respecified: a causal link between expectations and arousal ( $\gamma_{21}$ ) was introduced in model  $M_1$ .

The modified model  $M_1$  was estimated and a  $\chi^2$ -value of 33.5 with 40 degrees of freedom ( $p=.76$ ) obtained. The difference of the  $\chi^2$ -values between  $M_0$  and  $M_1$  indicated that  $M_1$  fitted the data significantly better than  $M_0$  ( $\Delta\chi^2=7.6$ , d.f.=1,  $p<.01$ ). However, the path coefficients between arousal and satisfaction ( $\beta_{42}$ )

---

<sup>11</sup> These residuals are discrepancies between the correlations implied by the estimated model and the corresponding correlations as observed in the data. Standardizing helps in the identification of residuals that are larger than one would expect as a result of sampling variation alone. Usually, a standardized residual of 2 and larger indicates a misspecification (at approximately the .05 level).

However, these residuals are expected to be normally distributed around zero, which means one should not be surprised to find some of them to be larger than 2 if one has a large number of residuals. For example, for 40 standardized residuals, one would expect to find 5% or 2 of them to be more than 2 standard deviations away from the mean. For a set of standardized residuals, a plot of these expected intervals on the vertical axis against the actual standardized on the horizontal axis would yield a straight line with a slope of 1. This would be the expected pattern given no misspecifications in the model and residuals greater than zero simply because of random sampling variation. However, if the larger residuals on the positive side fall clearly below the ideally expected 45-degree line, or if the smaller residuals on the negative side fall clearly below it, than some of the residuals are larger than one would expect as a result of random variation. The residuals most clearly "out of line" suggest possible locations for a model misspecification (Herting and Costner 1985).

were not significantly different from zero, either in model  $M_0$  nor in  $M_1$  ( $t = .91$ , n.s. and  $t = 1.15$ , n.s., for  $M_0$  and  $M_1$ , respectively). Therefore, a further modified model  $M_2$  was specified in which this causal link was deleted.

Model  $M_2$  had a  $\chi^2$ -value of 34.7 with 41 degrees of freedom ( $p = .75$ ). The  $\Delta\chi^2$  between  $M_1$  and  $M_2$  was far from significant ( $\Delta\chi^2 = 1.2$ , d.f. = 1,  $p < .25$ ), suggesting that deletion of the link in question would not result in significant deterioration in the fit of the model. Also, a difference in  $\chi^2$  close to the difference in number of degrees of freedom indicates that the improvement in fit through the additional link may have been obtained by "capitalizing on chance" (Joereskog and Soerbom 1982, p. 408). In summary, the data did not justify the specification of a causal link between arousal and satisfaction. Therefore, this link was removed in all further models.

Some indices of fit suggested that further improvement of the model  $M_2$  might be possible. Two standardized residuals were slightly higher than 2, namely those for the correlations between expectations and AROUS1 (-2.28), and between expectations and PLEAS1 (-2.00). A Q-plot indicated that these two, and a number of other standardized residuals, were higher than one would have expected from sampling effects alone. The modification indices for the links between arousal and pleasure ( $\beta_{32}$ ), and between expectations and pleasure ( $\gamma_{31}$ ), were the same (MI = 4.59), indicating that the fit of the model could be further improved by introducing either one of the two links (estimated  $\Delta\chi^2 = 4.6$ , d.f. = 1,  $p < .05$ ).

Introducing a causal link between arousal and pleasure did not make substantive sense, as it has been repeatedly shown that pleasure and arousal are orthogonal dimensions of affect. The cause for the proposed change (MI) might have been some collinearity between the two variables. Also, the residuals did not point to a misspecification involving these two variables. In conclusion, the residuals and the modification index suggested a causal link between expectations and pleasure ( $\gamma_{31}$ ), which was specified in  $M_3$ .

Model  $M_3$  had a  $\chi^2$ -value of 30.0 with 40 degrees of freedom ( $p=.88$ ). The  $\Delta\chi^2$  between  $M_2$  and  $M_3$  indicated that  $M_3$  had a significantly better fit than  $M_2$  ( $\Delta\chi^2=4.7$ , d.f.=1,  $p<.05$ ).

$M_3$  was accepted as the final model and is presented in Figure 7-3. All of the focused indicators, which look at specific relationships between observed variables or specific parameters, indicated a good specification of the model. In particular, the t-values of all paths were significant at the 5% level (refer to Table 7-22), the standardized residuals did not indicate any misspecification (refer to Q-plot of residuals in Appendix F), and the modification index did not indicate further changes that could have significantly improved the fit<sup>12</sup>). The overall indices, which measure the fit for the entire model, also suggested that the model  $M_3$  fits the data very well. Apart from the excellent  $\chi^2$ -value and probability level, the

---

<sup>12</sup> The only marginally significant proposed change was a link from PLEAS2 ( $y_2$ ) to arousal ( $\eta_2$ ) (MI=4.16,  $p<.05$ ), which had to be rejected on the basis on theoretical reasoning.

adjusted goodness-of-fit (adjGFI=.93) and the root mean square residual (RMSR=.04) also indicated a very good fit<sup>13</sup>).

Table 7-21: Goodness-of-Fit Indicators for Alternative Models

Model	d.f.	$\chi^2$	p	Adj GFI	RMSR	Changes suggested by MI and T-values <sup>14)</sup>	$\Delta\chi^2$	p <sup>15)</sup>
M <sub>0</sub>	41	41.1	.48	.91	.07	$\beta_{42}$ : insignificant t-value $\gamma_{21}$ : maximum MI of 6.8	-	-
M <sub>1</sub>	40	33.5	.76	.92	.05	$\beta_{42}$ : insignificant t-value	+ 7.6	<.01
M <sub>2</sub>	41	34.7	.75	.92	.05	$\gamma_{31}$ : maximum MI of 4.6	- 1.2	n.s.
M <sub>3</sub>	40	30.0	.88	.93	.04	final model	+ 4.7	<.05

<sup>13</sup> See Joereskog and Soerbom (1982) for a detailed description of the various fit indices.

<sup>14</sup> The suggested changes printed in bold were implemented in the subsequent model.

<sup>15</sup> The probability level in this column indicates whether the change of the  $\chi^2$  value to the previous model is significantly different from zero.

#### 4.4 Discussion of the Results

The final model  $M_3$  is shown in Figure 7-3. The Lisrel estimates of the parameters and their associated t-values are summarized in Table 7-22.

##### Determinants of Disconfirmation

The highly significant gamma-values of .44 for  $\gamma_{11}$  ( $t=5.6$ ,  $p<.001$ ), and -.40 for  $\gamma_{13}$  ( $t=-5.1$ ,  $p<.001$ ) showed that both the expectations and performance manipulations had strong effects on disconfirmation. The results were consistent with the results of the correlation analysis and analysis of variance reported in section 3.

The variance manipulation had a small ( $\gamma_{12}=.14$ ) and marginally significant ( $p=.05$ ) effect on disconfirmation. This could be explained as it was shown in section 3.2 that on average a high variance expectation led to a more positive disconfirmation than a low variance expectation (Table 7-15). However, this observation was driven only by effects of variance expectations in low or no disconfirmation situations, where they had a strong and significant impact on disconfirmation. In high disconfirmation situations, no significant impact of variance expectations was observed.

Hypotheses II and III could not be tested using Lisrel analysis as they did not predict linear causal relationships (refer to section 3.2 for tests of these hypotheses and an interpretation of the results). The reason for including variance in the Lisrel model was to control for potential effects of the variance manipulation on other independent variables. Neither the modification indices nor the residuals gave any indication of variance having an impact on any other variable but disconfirmation.

### The Role of Affect in the Satisfaction Process

In line with the results of the correlation analysis, the Lisrel results showed that both pleasure and arousal were a direct function of disconfirmation. The impact of disconfirmation was about equally strong on pleasure ( $\beta_{31}=.56$ ) and arousal ( $\beta_{21}=.58$ ), and highly significant ( $t=5.2, p<.001$ , and  $t=4.7, p,.001$  for  $\beta_{31}$  and  $\beta_{21}$  respectively). This means that Hypothesis IV, a positive causal relationship between disconfirmation and pleasure was confirmed.

Hypothesis V, which proposed that arousal is driven by the magnitude of disconfirmation, was rejected following the analysis reported in section 3.3. The data showed a correlation between disconfirmation, rather than its magnitude, and arousal. The Lisrel results support the interpretation of this correlation as being causal in nature. This means that people became more aroused when performance was better than expected, and less aroused when it was worse than expected. It

seems possible that response time had an "intrinsic" arousal potential and that any arousing effect of a disconfirmation experience with response time was overruled by the effect of "slowing down" from one's expectations. A more detailed discussion of this interpretation is presented in the concluding chapter.

Pleasure had a strong and direct affect on satisfaction ( $\beta_{43}=.79$ ,  $t=8.9$ ,  $p<.001$ ), which confirmed Hypothesis VI.

Hypothesis VII, which predicted that arousal and pleasure would interact and impact on satisfaction, was rejected in section 3.4. Rather the analysis showed a significant positive correlation between arousal and satisfaction. The Lisrel results led to an interpretation of this correlation as being spurious rather than causal in nature. It seemed that the variation in both arousal and satisfaction was caused by a third variable, namely disconfirmation. Disconfirmation was a causal antecedent of arousal and it drove satisfaction directly as well as indirectly through pleasure.

Disconfirmation had a significant positive effect on satisfaction ( $\beta_{41}=.21$ ,  $t=2.9$ ,  $p<.01$ ). This link corresponds to the classical disconfirmation-of-expectations model. It shows that not all effects of disconfirmation were mediated through affect.

## Other Findings

Two significant causal paths between expectations ( $\xi_1$ ) and arousal ( $\gamma_{21} = -.34$ ;  $t = -3.2$ ;  $p < .001$ ) and expectations and pleasure ( $\gamma_{31} = -.22$ ;  $t = -2.2$ ;  $p < .02$ ) were encountered. As expectations were manipulated by letting the subjects work on a "Simulator," expectations were not only cognitively communicated through the information section of the questionnaire, but they were also experienced. During simulation the subjects experienced a fast (1 second) or a slow response time (12 seconds) condition. It seems plausible that the subjects in the fast response time simulation condition experienced higher levels of arousal and pleasure than subjects in the slow response time condition. This affective state during the first part of the experiment may have influenced how people felt when later experiencing the "real service."

The path coefficients between expectations and arousal ( $\gamma_{21} = -.34$ ), and between expectations and pleasure ( $\gamma_{31} = -.22$ ) were markedly smaller than between disconfirmation and arousal ( $\beta_{21} = .58$ ), and disconfirmation and pleasure ( $\beta_{21} = .58$ ). In other words, pleasure and arousal were mainly determined by disconfirmation, and this permitted examination of the hypothesized relationships in spite of unexpected effects arising from the experimental design.

Figure 7-3: Final Model

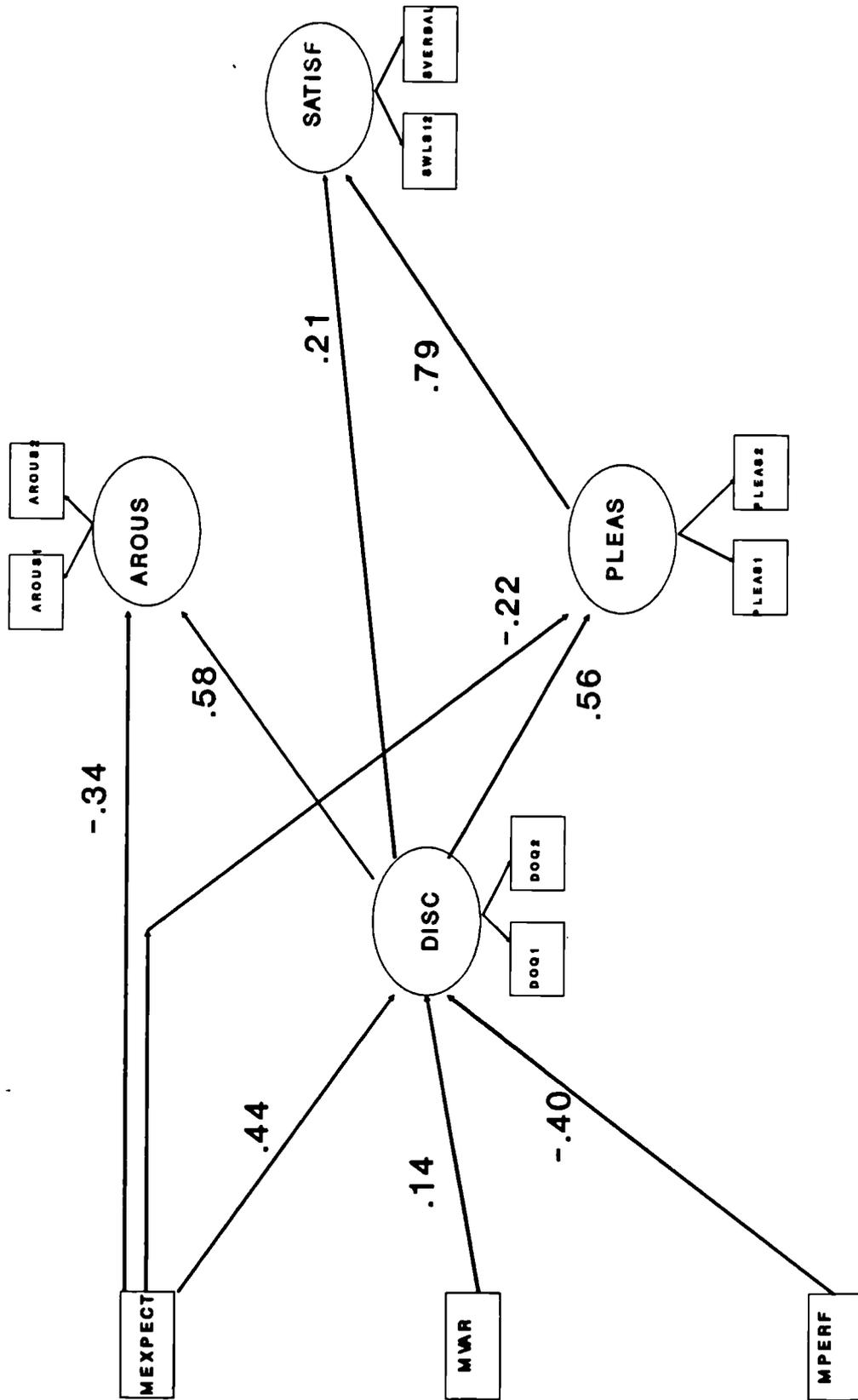


Table 7-22: Estimated Parameters for the Final Model  $M_3$

Parameter <sup>16)</sup>	Lisrel Estimate	Standard Error	T-Value <sup>17)</sup>	Standardized Value
$\lambda_{x11}$	1.00 <sup>18)</sup>			1.00
$\lambda_{x22}$	1.00			1.00
$\lambda_{x33}$	1.00			1.00
$\theta_{\epsilon11}$	0			0
$\theta_{\epsilon22}$	0			0
$\theta_{\epsilon33}$	0			0
$\phi_{11}$	1.00			1.00
$\phi_{22}$	1.00			1.00
$\phi_{33}$	1.00			1.00
$\lambda_{y11}$	1.00			.92
$\lambda_{y21}$	1.03	.07	14.4	.95
$\lambda_{y32}$	1.00			.73
$\lambda_{y42}$	1.25	.18	7.1	.92
$\lambda_{y53}$	1.00			.87
$\lambda_{y63}$	.94	.09	10.6	.82
$\lambda_{y74}$	1.00			.93
$\lambda_{y84}$	.95	.07	14.0	.88
$\theta_{\epsilon11}$	.17	.05	3.6	.17
$\theta_{\epsilon22}$	.12	.05	2.5	.12
$\theta_{\epsilon33}$	.46	.09	5.5	.46
$\theta_{\epsilon44}$	.16	.10	1.7	.16
$\theta_{\epsilon55}$	.24	.05	4.5	.24
$\theta_{\epsilon66}$	.33	.06	5.6	.33
$\theta_{\epsilon77}$	.14	.04	3.3	.14
$\theta_{\epsilon88}$	.22	.04	5.1	.22
$\gamma_{11}$	.41	.07	5.6	.44
$\gamma_{12}$	.13	.07	1.8	.14
$\gamma_{13}$	-.37	.07	-5.1	-.40
$\gamma_{21}$	-.25	.08	-3.2	-.34
$\gamma_{31}$	-.19	.09	-2.2	-.22

cont...

<sup>16</sup> The first number of the subscript defines the row and the second the column location of the parameter of the respective matrices.

<sup>17</sup> All t-values are significant at the .05 level (d.f.=121, one-tailed).

<sup>18</sup> The values 1.00 and 0 in this column were fixed by the analyst.

Parameter	Lisrel Estimate	Standard Error	T-Value	Standardized Value
$\beta_{21}$	.46	.10	4.7	.58
$\beta_{31}$	.53	.10	5.2	.56
$\beta_{41}$	.21	.07	2.9	.21
$\beta_{43}$	.84	.09	8.9	.79
$\psi_{11}$	.53	.09	6.1	.63
$\psi_{22}$	.39	.09	4.2	.73
$\psi_{33}$	.57	.11	5.4	.75
$\psi_{44}$	.16	.05	3.2	.19
$\psi_{32}$	.25	.07	3.7	.38
$\chi^2 = 30.0$	d.f. = 40	p = .88	AdjGFI = .93	RMSR = .04

## 4.5 Tests of Rival Models

Some care needs to be taken with Lisrel analysis because the function  $F$  may have several local minima<sup>19</sup>). Although experience indicates that multiple solutions seldom occur (Joereskog and Soerbom 1982), a number of potential rival models were tested and compared with the final model  $M_3$ .

### Rival Hypotheses About the Role of Pleasure in the Satisfaction Process

The first three rival models tested examined the role of pleasure in the satisfaction process. The alternative model  $M_{a1}$  incorporated the hypothesis that pleasure did not cause satisfaction, rather that pleasure was only a correlate of the satisfaction construct. To test this argument, the causal link between pleasure and satisfaction ( $\beta_{43}$ ) was deleted. The  $\chi^2$ -value was significantly worse than that of model  $M_3$  ( $\Delta\chi^2=81.0$ , d.f.=1,  $p<.001$ , Table 7-23). The rival hypothesis was not supported.

One might also argue that pleasure and satisfaction were highly correlated, not because pleasure causes satisfaction, but because a higher level of satisfaction is associated with a higher level of experienced pleasure. This argument was tested in model  $M_{a2}$  by reversing the hypothesized causal relationship between pleasure and satisfaction by setting  $\beta_{43}$  to zero and relaxing  $\beta_{34}$ . Again, the  $\chi^2$ -value was

---

<sup>19</sup> For a definition of  $F$  refer to Section 4.1 of this chapter.

significantly worse than that of  $M_3$  ( $\Delta\chi^2=23.6$ , d.f.=1,  $p<.001$ ) leading to the rejection of the rival hypothesis.

Furthermore, one might argue that all the effects of disconfirmation were mediated through pleasure and that there is no direct causal link between disconfirmation and satisfaction. This hypothesis was tested in model  $M_{a3}$  by deleting the link between disconfirmation and satisfaction ( $\beta_{41}$ ). The  $\chi^2$ -value was, again, significantly worse than that of  $M_3$  ( $\Delta\chi^2=7.4$  d.f.=1,  $p<.01$ ), and the rival hypothesis rejected.

In conclusion, the rival hypotheses of models  $M_{a1}$ ,  $M_{a2}$  and  $M_{a3}$  were not supported.

#### Performance as a Direct Determinant of Satisfaction - Test of Hypothesis I

Hypothesis I proposed that there was no direct causal link between performance ( $\xi_3$ ) and satisfaction ( $\eta_4$ ). To test this hypothesis, an alternative model  $M_{a4}$ , in which  $\gamma_{43}$  was relaxed, was estimated. The  $\Delta\chi^2$  between  $M_{a4}$  and  $M_3$  was not significant ( $\Delta\chi^2=.1$ , d.f.=1, n.s.), suggesting that the improvement of  $\chi^2$  through the introduction of this link capitalized on chance. Furthermore, the t-value of  $\gamma_{43}$  was far from significant ( $t=.19$ , n.s.). The results did not support the view of causal link between performance and satisfaction.

To test Hypothesis I, it had to be shown not only that there was no direct causal link between the performance manipulation and satisfaction, as tested in model  $M_{a4}$ , but also that perceived performance as operationalized by the measures PPERF1 and PPERF2 did not directly influence satisfaction.

So far, the manipulation checks have not been included in the fitted model  $M_3$ . In model  $M_{b1}$  (see Appendix G, Figure A-1, Table A-1) all of the manipulation checks were included in  $M_3$ . The model was fitted in a similar fashion as reported for model  $M_3$ . The modification index and standardized residuals suggested the introduction of a causal link between expected performance ( $\eta_1$ ) and perceived performance ( $\eta_3$ ) - incorporated in model  $M_{b11}$  - and a direct link between expected performance and expected variance ( $\eta_2$ ) - incorporated in model  $M_{b12}$ . Both of these effects had been observed in the analysis of variance discussed in section 2.5.2 of this chapter.

After the model was fitted, the disputed link between perceived performance and satisfaction was included in model  $M_{b2}$ . Again, the  $\Delta\chi^2$  between  $M_{b2}$  and  $M_{b12}$  ( $\Delta\chi^2=0.0$ ; d.f.=1; n.s.) and the t-value of this link ( $\beta_{73}$ ) ( $t=.14$ ; n.s.) were both far from significant.

In summary, the data neither supported a direct causal link from the performance manipulation nor from perceived performance to satisfaction, which provided support for Hypothesis I.

### Expectations as a Direct Determinant of Satisfaction

Not only performance, but also expectations have sometimes been thought to directly influence satisfaction (e.g. Churchill and Surprenant 1982, Oliver and DeSarbo 1988). Model M<sub>a5</sub> tested for an affect of expectations manipulation on satisfaction. The  $\Delta\chi^2$  between M<sub>a5</sub> and M<sub>3</sub> was far from significant ( $\Delta\chi^2=0.1$ ; d.f.=1; n.s.), which lead to the rejection of M<sub>a5</sub>. Model M<sub>b3</sub> tested whether satisfaction was a direct function of expected performance. Again, the  $\Delta\chi^2$  between M<sub>b3</sub> and M<sub>b21</sub> was not significant ( $\Delta\chi^2=0.1$ ; d.f.=1; n.s.)

In conclusion, neither a direct causal link from the expectation manipulation to satisfaction nor from expected performance to satisfaction was supported.

### Perception Measures Instead of Manipulations as Key Variables

The analysis of variance in section 2.5.2 showed two main effects for perceived performance (Table 7-10). The performance manipulation had a strong and highly significant main effect [ $F(1,124)= 105.9, p<.001$ ] which confirmed a successful manipulation. Although somewhat weaker, the expectations manipulation also showed a highly significant main effect [ $F(1,124)= 49.6, p<.001$ ].

The following interpretations were provided in section 2.5.2. First, the expectations main effect might have indicated a consistency effect. Second, subjects might have simply used the experienced response time with the "Simulator" (=expectations manipulation) as an anchor value to translate their internal time perception into seconds. If the latter was the case, the expectations main effect was mainly a measurement error and, in consequence, the performance manipulation would have been a better indicator of the response time perceived than the perception measures. This was the implicit assumption in the Lisrel models tested so far. If, however, there was a consistency effect between expectations and perceived performance, and the perceived performance measures were a better indicator of the performance perceived by the subjects, than the perception measures should be modelled as antecedents of the disconfirmation process rather than the manipulations themselves.

To test this alternative proposition, the manipulation checks were modelled as antecedents of the disconfirmation and satisfaction process. The hypothesized model  $M_{b4}$  is depicted in Appendix H, Figure A-3.

The model was fitted in a similar fashion as reported for Model  $M_3$ . The t-value for the link between expected variance (EVAR) and disconfirmation was found not to be significant ( $t= 1.1$ , n.s.). This link was removed in model  $M_{b41}$ . The modification index suggested the introduction of a link from manipulated performance (MPERF) to disconfirmation ( $MI= 5.0$ ), which was incorporated in model  $M_{b42}$ . In  $M_{b42}$  the link between the perceived performance (PPERF) was

not significant ( $t = -.9$ , n.s.) and was therefore removed in model  $M_{b43}$ . The fitted model  $M_{b43}$  is shown in Figure A-4.

The results clearly show that the performance manipulation (MPERF) was in fact a better indicator of the true performance perceived by the subjects than the perception measures (PPERF). The model  $M_{b42}$  showed that it was the manipulation rather than the manipulation check (PPERF) which drove disconfirmation. The link between MPERF and DISC was significant and strong ( $t = -2.3$ ,  $p < .05$ ,  $\gamma_{43} = -.30$ ), whereas the link between PPERF and DISC was not significant ( $t = -0.9$ , n.s.,  $\beta_{43} = -.14$ ). Furthermore, the overall fit of the model  $M_{b43}$  was significantly worse than the fit of model  $M_{b12}$  which used manipulations rather than manipulation checks as determinants of the disconfirmation and satisfaction process ( $\Delta\chi^2 = 4.3$ ; d.f. = 1;  $p < .05$ ).

In conclusion, the performance manipulation seemed to be a better indicator of the true perceptions of subjects than the perceived performance measures (manipulation checks). This provided support for using the manipulations rather than their checks as key variables in the Lisrel models.

Table 7-23: Goodness-of-Fit Indicators for Rival Models

Model	d.f.	$\chi^2$	p	Adj GFI	RMSR	Changes suggested by MI and T-values	$\Delta\chi^2$	p
<u>M<sub>3</sub></u>	<u>40</u>	<u>30.0</u>	<u>.88</u>	<u>.93</u>	<u>.04</u>	<u>final model</u>		
M <sub>a1</sub>	41	111.0	.00	.81	.14		-81.0 <sup>20)</sup>	p<.001
M <sub>a2</sub>	40	53.6	.07	.88	.09		-23.6	p<.001
M <sub>a3</sub>	41	37.4	.05	.92	.05		7.4	p<.01
M <sub>a4</sub>	39	29.9	.85	.93	.04		0.1	n.s.
M <sub>a5</sub>	39	29.9	.85	.93	.04		0.1	n.s.
M <sub>b1</sub>	111	175.7	.00	.82	.11	$\beta_{31}$ : maximum MI of 38.5	-	-
M <sub>b11</sub>	110	128.7	.11	.86	.07	$\beta_{21}$ : maximum MI of 8.9	67.0	p<.001
M <sub>b12</sub>	109	118.2	.26	.87	.05	fitted model	10.5	p<.001
M <sub>b2</sub>	108	118.2	.26	.87	.05		0.0 <sup>21)</sup>	n.s.
M <sub>b3</sub>	108	118.1	.24	.86	.05		0.1	n.s.
M <sub>b4</sub>	109	125.6	.13	.86	.05	$\beta_{42}$ : insignificant t-value	-	-
M <sub>b41</sub>	110	126.8	.13	.86	.06	$\gamma_{43}$ : maximum MI of 5.0	-1.2	n.s.
M <sub>b42</sub>	109	121.8	.19	.86	.06	$\beta_{43}$ : insignificant t-value	5.0	p<.05
M <sub>b43</sub>	110	122.5	.19	.86	.06	fitted model	-0.7	n.s.

<sup>20</sup> The  $\Delta\chi^2$  values for the rival models M<sub>a1</sub> to M<sub>a4</sub> are based on the comparison with the final model M<sub>3</sub>.

<sup>21</sup> The  $\Delta\chi^2$  value for the rival models M<sub>b2</sub> and M<sub>b3</sub> are based on a comparison with the fitted model M<sub>b12</sub>.

## 4.6 Summary

In this section the results of the model testing with Lisrel are presented.

First, the final model  $M_3$  is shown in Figure 7-3. All focused indicators suggested a good specification of the model. In particular, the t-values of all causal paths were significant at the 5% level, the residuals did not indicate any misspecification and the modification index indicated that any further changes would not have improved the fit significantly. All overall indices also suggested that the model fitted the data very well. Furthermore, in Lisrel analyses multiple solutions can occur in rare cases. Therefore, a number of potential rival models were tested all of which could be rejected.

Second, the Lisrel analyses were used to examine the hypothesized linear causal relationships between variables. The results clearly rejected a direct causal link from the performance manipulation and from perceived performance to satisfaction. This supported Hypothesis I. In accordance with the hypotheses advanced, the analysis showed that positive disconfirmation caused pleasure (Hypothesis IV) and that pleasure caused satisfaction (Hypothesis VI).

Third, Hypotheses V and VII concerning the determinants and antecedents of arousal were rejected in section 3 using traditional statistical methods. These hypotheses proposed non-linear relationships. However, the results of the traditional analysis suggested linear relationships. Lisrel was used to examine

whether these linear relationships were causal in nature. Concerning the determinants of arousal, the Lisrel results showed that disconfirmation was a linear causal antecedent of arousal. This link was explained with disconfirmation-of-response-time-expectations having an "intrinsic" arousal potential and that any arousing effect of a disconfirmation experience with response time was overruled by the effect of "slowing down" from one's expectations.

Concerning the consequences of arousal, the Lisrel results supported the interpretation of the correlation between arousal and satisfaction as being spurious rather than causal. It seemed that the variations in both arousal and satisfaction were caused by disconfirmation. Disconfirmation was a causal antecedent of arousal and of satisfaction directly and indirectly through pleasure.

Fourth, in section 3 the relationship between the variance expectations and disconfirmation (Hypotheses II & III) was shown to be non-linear. Therefore, it could not be tested using Lisrel.

# Chapter VIII

## Conclusions and Implications

In this final chapter, conclusions are drawn from the experimental results and related to the literature in consumer satisfaction, services marketing and psychology. In sections one and two the findings and implications of the test of two extensions to the traditional disconfirmation-of-expectations model are discussed. These extensions address two features typical of services. One extension takes account for the first time of the role of variance expectations in the disconfirmation process. The other extension includes affect (as conceptualized by Russell, 1980) in the disconfirmation model to better capture the experiential nature of services and to reconcile the disconfirmation-of-expectations model with recently developed perspectives on the service encounter that use affect as a mediating variable between stimuli and behaviour (Bateson and Hui 1988, Donovan and Rossiter 1982).

The third section covers the findings and research implications of the methodological issue of whether perceived performance can be regarded as a direct causal antecedent of satisfaction. The limitations of the study are discussed in the fourth section, and a summary is provided in the last section.

## 1 The Impact of Expected Performance Heterogeneity on Disconfirmation

High degrees of performance heterogeneity are a frequently cited feature of services (e.g. Berry 1980, Krughoff 1981), and their implications for management of the service encounter have been discussed extensively (e.g. Chase 1978, 1981; Levitt 1972, 1976; Mills and Moberg 1982). Furthermore, performance heterogeneity has received much attention in the pre-choice literature on perceived risk (e.g. Diamond 1988) and multiattribute models (e.g. Meyer 1981, Sarel 1978). However, it has been largely ignored in the satisfaction literature. For the satisfaction process, researchers have always implicitly assumed point-expectations, disregarding the potential impact of expected performance heterogeneity on the disconfirmation process.

Two hypotheses concerning the impact of expected performance heterogeneity on disconfirmation were developed on the basis of the disconfirmation-of-expectations model. In Chapter III, the following hypotheses were put forward:

### Hypothesis II:

In situations where perceived performance and the expected mean performance are approximately equal, consumers experience confirmation independent of the expected variance in performance.

### Hypothesis III:

In disconfirmation situations, consumers who hold mean performance expectations with some degree of variation around the mean, experience smaller magnitudes of disconfirmation than consumers who hold point expectations.

#### 1.1 Summary of Findings - Hypotheses II & III

Hypotheses II and III had to be rejected. In the two disconfirmation conditions, i.e. slow expected and fast perceived performance and vice versa, subjects perceived the same level of disconfirmation independent of the variance condition.

However, when the expected mean performance was very close to the actual performance, i.e. fast expected and fast perceived performance, or slow expected and slow perceived performance, the subjects experienced more positive disconfirmation when they expected a high variance in performance.

These results were quite surprising. They suggest that people disregard the variance in their expectations when the discrepancy between the expected mean performance and perceived performance is large. However, when the discrepancy is small, a high variance expectation causes the people to perceive a "better-than-expected" performance. One explanation for this effect may be that people are

influenced by the possibility that the performance could have been worse and, therefore, perceive a more positive disconfirmation than people who held point expectations.

## **1.2 Managerial Implications**

In the majority of consumption experiences, expectations are met by performance, and consumers experience confirmation and satisfaction. Therefore, it is an important finding that a large variance expectation has a positive impact on the level of perceived disconfirmation, when the discrepancy between expectations and perceived performance is small. This means that a service firm can make its customers perceive a more positive disconfirmation and higher levels of satisfaction just by communicating the inherent heterogeneity in its service process.

For example, the airport services supplier BAA Plc advertises in its baggage halls the number of suitcases handled every day and its commitment to fast and reliable service. The sheer magnitude of the task that has to be managed indicates to the traveller that there is potential for heterogeneity in performance, at least in the form of delays or missing baggage. The results of this thesis suggest that, in this situation, customers who know about heterogeneity will be more satisfied when they receive their suitcases as expected than customers who do not hold a variance expectation.

### Consumer Choice versus Satisfaction Processes

Expected performance heterogeneity also has impact on the consumer choice process. In particular, in circumstances of high expected performance heterogeneity consumers are likely to perceive higher levels of risk which they then try to reduce (Bauer 1960). Consumers do this, for example, by choosing service providers with smaller unit-to-unit variability. Furthermore, uncertainty about the performance of attributes leads to a discounting of expected utility (Meyer 1981). Both the increase in perceived risk and discounting of attribute utility can have, for the service firm, negative effects on the consumer choice process. In consequence, the advantage in the satisfaction process of having customers who expect heterogeneous performances has to be traded off against detrimental effects in the consumer choice process.

This trade-off may not be relevant in the monopoly supply of services (e.g. airports services supplier BAA Plc) or where the level of uncertainty about performance is the same for all suppliers in a service sector and this can be communicated to consumers (e.g. aircraft departure times). In these cases the service firm may not impair its relative competitive position in the consumer's choice process.

Furthermore, Gelb and Smith (1988) have shown that consumer expectations can be changed not only before but also during the service encounter, and it is the revised expectations that then constitute the comparison standard in the disconfirmation process. This means that it may be possible to increase the

expected performance variance during the service encounter after the choice has been made, thereby avoiding the detrimental impact on the choice process.

#### Customized Services have a High Inherent Heterogeneity

Customized services have a higher inherent heterogeneity in performance than standardized offers (Mills and Moberg 1982, Surprenant and Solomon 1987). For example, the taste of a hamburger meal at McDonalds shows very little unit-to-unit variability. However, a dish in an upmarket Chinese restaurant can vary significantly over time, depending on the current chef, the quality and kind of ingredients available, and the number and structure of orders of other customers.

As past experiences are an important determinant of expectations (LaTour and Peat 1977), it seems plausible that consumers also expect more heterogeneous performances from a highly customized service than from a highly standardized one. If this is the case, and the service firm can perform close to expectations, then the results of this study suggest that consumers of customized services should experience higher levels of satisfaction than consumers of highly standardized services. It might therefore be easier to have more satisfied customers with customized than with standardized services.

### 1.3 Further Research

No other study has examined the impact of expected performance heterogeneity on disconfirmation. To gain more confidence in the external validity of the findings, this study should be replicated in real-world service encounters. Furthermore, three extensions are proposed. First, in this study the perceived performance manipulation was either close to or far from the expected mean. It was therefore not possible to test whether the variance impact on disconfirmation would actually become smaller as the discrepancy between performance and expectations grows larger, or whether there is some threshold value above which variance expectations have no impact on disconfirmation. If there was a threshold value, it would be particularly important for quality specifications. To benefit from the positive effects of variance expectations, the specifications would have to assure performance below the threshold level.

Second, the variance effect was tested on only one fairly unambiguous attribute, i.e. response time of a homebanking service. It would be of interest to research the affect of variance expectations on more ambiguous attributes, such as the hygiene conditions in a restaurant or the soundness of tax advice given by an accountant. In Chapter II, it was proposed that consistency effects occur when consumers evaluate ambiguous attributes. That is, consumers experience what they expect. It remains to be tested whether and what impact variance expectations can have on this process. For example, it may be that consistency effects do not assimilate the level of perceived performance to the expected mean

performance, but to some level within a confidence range of the variance expectation, such as its upper or lower boundaries.

Finally, because the hypotheses developed on the basis of the disconfirmation-of-expectations model were rejected, the findings are largely empirical and not supported by theory. For developing a theoretical framework and linking it to existing theories an understanding of the underlying process is required. A first step would be exploratory research (e.g. with focus groups and personal in-depth interviews) determining potential dimensions underlying the process. As a further stage, existing theories related to these dimensions could be used to generate hypotheses to be tested.

## 2 The Role of Affect in the Satisfaction Process

The main feature of services is their experiential nature (Hui 1988, p.15). Several researchers have suggested that the commonly applied cognition based multiattribute and disconfirmation models are inadequate for capturing this feature (e.g. Bateson 1985b, Holbrook and Hirschman 1982). Therefore, alternative perspectives on the service encounter have been developed. These alternative perspectives use affect as a mediating variable between stimuli, cognitive processes and subsequent behaviour. The main objective of this thesis was to examine whether affect should also be included in satisfaction models for services.

Affect was conceptualized as having two dimensions, namely pleasure and arousal (Russell 1980). On the basis of the satisfaction literature it was hypothesized that disconfirmation drives the affective state of consumers during the service encounter. In particular, it was hypothesized that positive disconfirmation causes pleasure, negative disconfirmation causes displeasure (Hypothesis IV), and that the magnitude of disconfirmation is a determinant of arousal (Hypothesis V).

Research on the service encounter has shown that the affective state of consumers during the service experience has causal effects on their behaviour (Bateson and Hui 1988, Donovan and Rossiter 1982). This, and findings in satisfaction research which support the notion that affect drives satisfaction in general (Westbrook 1987, Westbrook and Oliver 1990), served as a basis for the hypothesis that the

affective state during the service encounter drives consumer satisfaction. Specifically, it was hypothesized that pleasure has a direct causal impact on satisfaction (Hypothesis VI), and that arousal works as an amplifier on this relationship (Hypothesis VII).

Theoretically, affect can perform the same role in the consumption of products as it performs in the consumption of services. However, the tested model is particularly valuable for services for two reasons. First the consumption experience, including the affective state of the consumer, can be managed by the service firm, whereas the goods firm has little control over the consumption experience of its products. For example, the service firm can influence customers' affective states through the design of the service encounter (e.g. its human interactions, built-in perceived control, colours, shapes, spacial layout, background music, etc.), whereas the goods firm has hardly any influence over the environment in which its products are consumed. Second, the environment in which consumption takes place is usually not a part of a product. However, both the environment and the experience constitute crucial parts, if not the main components, of service (Bitner 1990, Hui 1988, Knisely 1989).

In Chapter V, the following four hypotheses were advanced:

Hypothesis IV:

The degree of pleasure experienced in a service encounter is an increasing function of the perceived disconfirmation-of-expectations.

**Hypothesis V:**

The level of arousal experienced in a service encounter is an increasing function of the perceived magnitude of disconfirmation-of-expectations.

**Hypothesis VI:**

Satisfaction is an increasing function of the pleasure experienced during the service consumption process.

**Hypothesis VII:**

The strength of the impact of pleasure on satisfaction is an increasing function of the level of arousal experienced during the service consumption process.

## 2.1 Summary of Findings - Hypotheses IV to VII

Pleasure was shown to be driven by disconfirmation. Subjects experienced pleasure during the consumption process when they perceived a positive disconfirmation, and they experienced displeasure when they perceived negative disconfirmation. This result confirmed Hypothesis IV. Furthermore, the more pleasure the subjects experienced during the consumption process, the higher was their level of satisfaction with the consumption experience. Therefore, Hypothesis VI was also confirmed.

Concerning arousal, Hypotheses V and VII had to be rejected. In line with the psychology literature it was proposed that any disconfirmation, positive or negative, would cause arousal. The data confirmed this only for positive disconfirmation. Subjects were more aroused when the response time was faster than expected. When it was slower than expected (negative disconfirmation), the subjects felt less aroused. It seems that the arousing "surprise effect" from disconfirmation was probably overruled by the effect of "slowing down" from one's expectations. This relationship becomes clearer when one looks at the affective state instead of arousal by itself. One can see that people felt bored (low arousal and displeasure) in the "slower-than-expected" condition and felt excited (high arousal and pleasure) in the "faster-than-expected" condition.

The observed effect of disconfirmation of response time expectations on arousal was not anticipated when the study was designed. The subjects felt this way

probably because, apart from feeling bored, the longer-than-expected response time did not matter to them. This might have been partially an experimental artifact, as the subjects anticipated the study to take some time, and the actual conduct of the banking transactions was only a part of it - more time was spent on the introduction and answering of the questionnaire. In a real-world situation, people might have experienced higher levels of arousal.

The hypothesized role of arousal as an amplifier on the causal path between pleasure and satisfaction could not be confirmed, and Hypothesis VII had to be rejected. Since arousal and pleasure were significantly correlated, these results might have been due to multicollinearity between the two variables caused by disconfirmation and expectations. Disconfirmation had a strong and positive impact on both pleasure and arousal. Furthermore, expectations had a negative impact on both variables. As a consequence, pleasure and arousal were correlated which indicated that the affective state of subjects moved along the boring (low arousal and displeasure) and exciting axis (high arousal and pleasure) within the circumplex model of affect (Figure 4-3). The correlation between arousal and pleasure is likely to be due to the experimental design as substantial evidence exists for the orthogonality of the two variables (Mehrabian and Russell 1974; Russell 1979, 1980; Russell et al 1981, 1983).

Finally, the results provided further support for Westbrook's (1987) findings that disconfirmation and affect both contribute independently to satisfaction. The

analysis showed that disconfirmation had direct impact on satisfaction and that not all effects of disconfirmation on satisfaction were mediated through pleasure.

In summary, the data clearly supported the hypothesized role of pleasure in the satisfaction process. The relationship between disconfirmation and arousal seems to be more complex than originally hypothesized, and it may well be that some attributes have intrinsic arousal potential. The data supported the hypothesis that there is no direct causal relationship between arousal and satisfaction, and the hypothesized interaction effect between arousal and pleasure on satisfaction still needs to be confirmed.

## **2.2 Further Research: The Role of Arousal in the Satisfaction Process**

There are two main areas which require further research. First, some propositions about arousal in the satisfaction process remain to be confirmed (i.e. Hypotheses V and VII), and the issue is discussed in this subsection. Second, an extended model of consumer satisfaction with services proposed in the next subsection requires further development and testing. The extended model draws together new perspectives that use affect to model consumer behaviour in the service encounter, affect theories in psychology, and the findings of this thesis.

The hypothesized relationship between disconfirmation and arousal (Hypothesis V) could not be confirmed. It remains to be shown empirically that the generally proposed causal link between magnitude of disconfirmation and arousal holds.

Furthermore, the results seem to indicate that the arousing effect of a disconfirmation experience can be offset by the intrinsic arousal potential of a stimulus. This type of phenomenon has been observed in other contexts. For example, in a literature review on arousal creating stimuli, Mehrabian and Russell (1974, p.64 et seq.) made the observation that arousal increases directly with the loudness of music. If a pub that usually plays very loud music would suddenly on one evening play soft music, it is conceivable that the arousal level of its customers would be significantly reduced through the intrinsic arousal potential of music. In this case, the impact of disconfirmation on arousal may not be strong enough to counter the impact of the loudness of the music itself. In the context of the service encounter, very little research has been directed to the impact of environmental and situational stimuli that can drive arousal (e.g. music, colours, light intensity, room temperature and air humidity; Mehrabian and Russell 1974), which can, in turn, influence consumer behaviour in the encounter (Bruner 1990, Donovan and Rossiter 1982) and satisfaction (Westbrook 1987).

Environmental psychologists have described arousal as amplifying the impact of pleasure on behaviour and have demonstrated the effect (e.g. Donovan and Rossiter 1982, Mehrabian and Russell 1974). The existence of this amplification in the context of consumer satisfaction has been proposed (e.g. Westbrook 1987,

Westbrook and Oliver 1990) but, until now, has not been empirically tested. This study tested for and rejected the proposition that arousal amplifies the impact of pleasure on satisfaction (Hypothesis VII). The unexpected results can probably be attributed to multicollinearity between pleasure and arousal. This correlation between pleasure and arousal was likely to be due to the common determinants of the variation in both variables (Figure 7-3). To clarify the role of arousal in the satisfaction process, it therefore seems particularly important to conduct studies that examine separately the outcomes of pleasure and arousal (Hui 1988). Little has been done so far to investigate the impact of arousal on consumer behaviour in general, and on satisfaction in particular, when pleasure is being controlled for.

### **2.3 Further Research: Opening Up the Disconfirmation Paradigm**

In this study, it was shown for the first time that the affective state of people during a service encounter can be included in the standard disconfirmation-of-expectations model of consumer satisfaction. This is an important finding. Affect is the missing link that allows us to open up the disconfirmation paradigm and draw together recently proposed perspectives that aim to capture the experiential nature of the service encounter (Chapter I), the literature on consumer satisfaction (Chapter II) and the literature on affect in psychology (Chapter IV).

This thesis showed that disconfirmation is a powerful determinant of the affective state of consumers during the service encounter and of satisfaction outcomes thereafter. However, in the psychology and services marketing literatures many other causes of affect have been identified and these may also impact upon the consumer satisfaction process. Therefore it seems essential, as a next step, to investigate other potential determinants of affect in the service encounter.

The literature in cognitive, social and environmental psychology (reviewed in Chapter IV) shows that affect can be caused by cognitive processes of any degree of complexity. For instance, extremely simple cognitive processes that can cause affect include sensations and the perception of stimuli (e.g. colours, temperatures, light intensities, sounds and odours). These processes can happen consciously or even unconsciously. Examples of very complex cognitive processes that can cause affect are attribution of causes (Weiner 1980), appraisals (Hoffman 1986, Smith et al 1990) and schema processing (Fiske and Pavelchak 1986).

In the context of the service encounter little evidence exists, but this is consistent with the findings in the psychology literature. Researchers have tested the following stimuli: perceived control, perceived choice, crowding (Bateson and Hui 1988, Hui 1988), the physical design of service encounters (Donovan and Rossiter 1982), background music (Yalch and Spangenberg 1988) and interpersonal touch of hands (Fisher et al 1976). All of these stimuli were shown to be causal antecedents of affect.

Specifically, perceived control has been proposed and empirically corroborated as an important causal antecedent of affective responses in service encounters by Bateson (1985b) and by Bateson and Hui (1988). These researchers suggest that perceived control, mediated by affect, is a major dimension of the service experience<sup>1)</sup>. Furthermore, perceived choice and crowding in the service encounter were shown to significantly influence affect (Hui 1988). Perceived choice was shown to influence pleasure directly, and indirectly through perceived control as a mediating variable.

Donovan and Rossiter (1982) applied the Mehrabian and Russell (1974) model of affect to assess retail store atmosphere and to predict consumer behaviour within the store<sup>2)</sup>. Donovan and Rossiter conceptualized atmosphere as the affective response (pleasure and arousal) of customers to different retail store environments. The researchers suggest that commonly considered in-store variables like aisle-width, brightness and crowding are clearly antecedents of store atmosphere (or affective quality) rather than alternatives for it.

Yalch and Spangenberg (1987) varied background music in two clothing sections of a department store. They used a background music condition ("easy listening" without lyrics), a popular music condition (popular Top 40 with lyrics) and a no music condition. Subjects filled in self-administered questionnaires when in the

---

<sup>1</sup> For a review of the perceived control perspective on the service encounter refer to Chapter I, Section 3.2.

<sup>2</sup> For a review of Donovan and Rossiter's (1982) study refer to Chapter I, Section 3.1.

store. The results showed no significant impact of the different music conditions on pleasure but a significant impact ( $p > .03$ ) on arousal. Background music was found to be the least arousing condition. This finding was consistent with those of other researchers, establishing the fact that tempo determines the arousing quality of music (Mehrabian and Russell 1974). The finding that the different music conditions had no significant impact on pleasure is somewhat surprising, as other studies (for a review refer to Mehrabian and Russell 1974) showed that music is generally perceived as being pleasant.

Fisher et al (1976) conducted a study on the impact of a short touch of hands on the customer's affective state and his/her evaluation of a library clerk. Fisher et al based their conceptual framework on research in psychology on tactile stimulation. The manipulations were (1) tactile contact between a library clerk and a customer while in the process of handing back the customer's library card, and (2) no tactile contact. Data analysis revealed significant main effects for perceived pleasure [ $F(1,93)=4.21, p<.04$ ] and for library clerk evaluation [ $F(1,93)=6.02, p<.02$ ]. That is, more pleasure was perceived by subjects who were touched than by those who were not touched. Furthermore, subjects who were touched gave the clerk a significantly more favourable rating than those who were not touched. The subjects in the touch condition were asked later whether they perceived the touch. Only 57% said that they had. Interestingly, no differences were found between the "aware" and the "unaware" groups for any of the dependent measures. Fisher et al concluded that touch improved perceived pleasure and clerk evaluation, independent of whether the touch was perceived

or not. This finding also corroborates the notion of some psychologists that affect can be elicited without, with only minimal, or even with unconscious cognitive processing (e.g. Zajonc 1980).

In conclusion, the results of the studies on the service encounter reviewed here are consistent with results reported in the psychology literature. Both bodies of literature show that cognitive processes of any degree of complexity can cause affective responses. More complex cognitive processes may have a greater influence on affect than simpler processes (Hoffman 1986). Therefore, disconfirmation in a service experience may have a stronger impact on affect than simpler cognitive processes such as the perception of background music or touch of hands. However, in a situation with no disconfirmation (which is the usual consumption experience!) these processes with lower complexity might significantly influence affective outcomes.

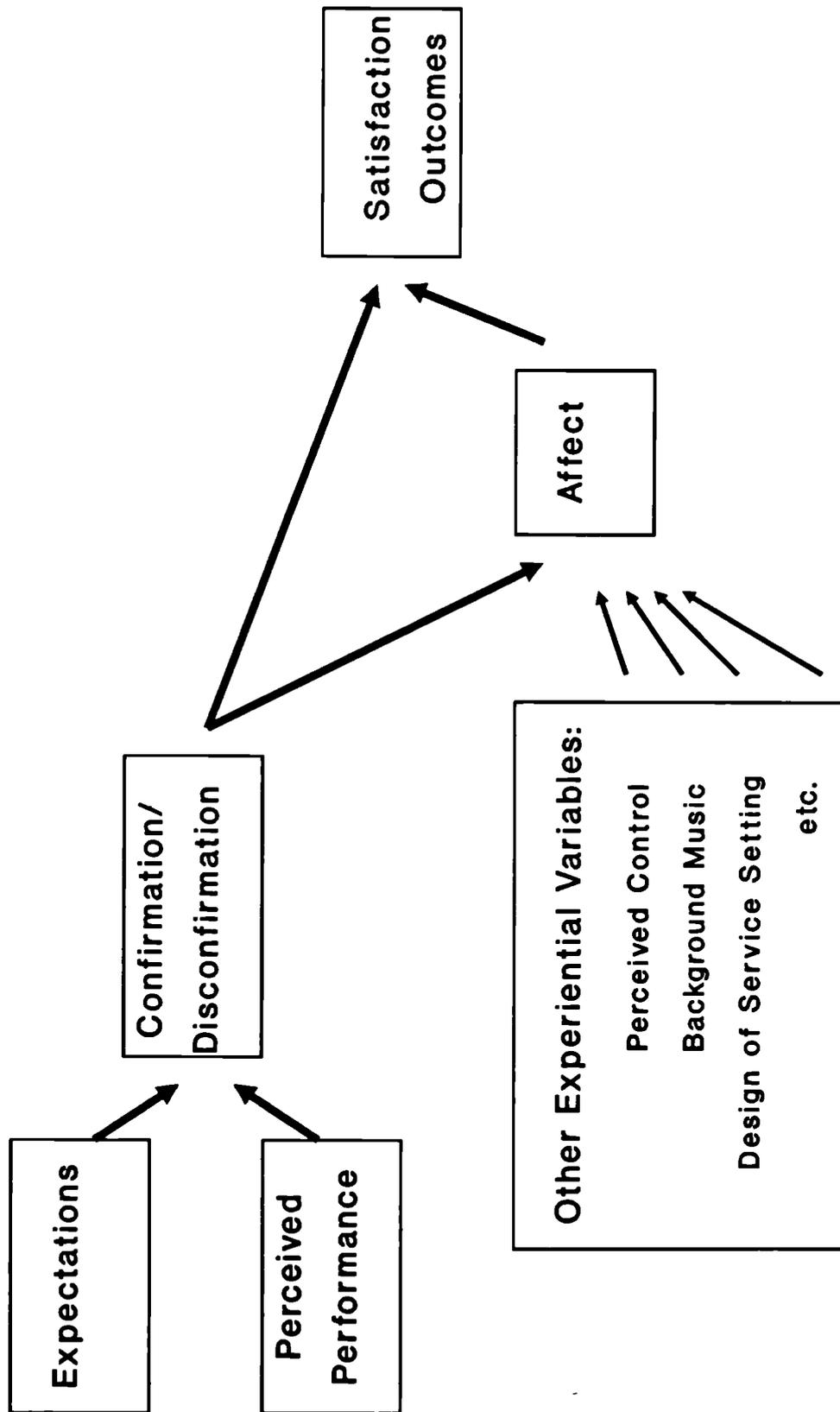
In the model tested in this study (Figure 5-1), the confirmation/disconfirmation variable was the only determinant of affect, and all other background factors were either controlled for or randomized across the experimental conditions. Further research effort should be directed toward extension of this model by incorporating into it other potential causes of affect identified in the psychology and services marketing literature. An extended model is presented in Figure 8-1<sup>3</sup>.

---

<sup>3</sup> This model was first presented at the TIMS College on Marketing Special Interest Conference on Services Marketing, 16-18 September 1990 at Vanderbilt University, Nashville, Tennessee (Wirtz and Bateson 1990).

The extended model seems intuitively appealing, not least because strong evidence exists for its component relationships. The link between affect and satisfaction has been supported by Westbrook (1987), and Westbrook and Oliver (1991). The links between a variety of sensual, perceptual and cognitive stimuli and affect have been shown in the services marketing (e.g. Bateson and Hui 1988, Fisher et al 1976) and the psychology literatures (e.g. Fiske and Taylor 1981, Mandler 1975, Zajonc 1980). The missing links between disconfirmation and affect, and between affect, as conceptualized by Russell (1980), and satisfaction have been established in this thesis. Whether the entire model is empirically supported when several stimuli are manipulated simultaneously is an intriguing question to be addressed by further research.

**Figure 8-1: Inclusion of Experiential Variables  
in Satisfaction Models**



## 2.4 Managerial Implications

The proposed "opened-up" disconfirmation model (Figure 8-1)<sup>4)</sup> has at least three important implications for services marketing management.

First, the model provides a framework for examining potential causes of dissatisfaction. Only disconfirmation with key attributes of the service encounter could be examined with the traditional disconfirmation model<sup>5)</sup>. However, as Bateson and Hui (1988) have demonstrated, a lack of perceived control can have a detrimental impact on the affective state of consumers and their behaviour in the service encounter. This potential cause of dissatisfaction could not have been detected by examining disconfirmation with key attributes. Apart from perceived control, there are many other experiential variables that can drive the affective state of consumers during the service encounter. Inclusion of affect in the satisfaction model, as tested in this thesis (Figure 5-1), can provide a "warning signal" that will indicate when disconfirmation with key attributes is probably not the cause of dissatisfaction with a service encounter. The search for the cause of dissatisfaction can then be extended to the experiential variables of the service encounter.

---

<sup>4</sup> Although there is strong empirical evidence for all the links proposed in this model, some caution needs to be raised because the model has not yet been tested when several stimuli vary or are manipulated simultaneously.

<sup>5</sup> Usually, this analysis includes expectations and perceived performance, which together drive disconfirmation.

Second, the model opens up new ways of thinking about how increasing consumer satisfaction. Contributors to the marketing literature have repeatedly argued that it is difficult to increase consumer satisfaction by increasing performance perceptions, simply because consumers adjust their expectations to the higher performance levels and do not then experience positive disconfirmation thereafter. Furthermore, expectations can rarely be lowered because the service provider would then lose out in the consumer choice process. The proposed model gives alternative ways of thinking about increasing consumer satisfaction by including affect as an additional variable for manipulation. For example, empirically tested determinants of affect capable of being manipulated are perceived control (Bateson and Hui 1988) and background music (Yalch and Spangenberg 1988).

Third, the model provides a conceptual framework for assessing the impact of experiential variables on satisfaction. Most practitioners already work with at least of some of these experiential variables (e.g. music, lighting, colours, employee-customer interactions), but without being able to judge their impact on consumer satisfaction. Instead, the impact of background music and colours has been assessed in terms of time or money spent in the service setting (e.g. Belizzi et al 1983, Milliman 1982, 1986). However, even if no significant impact of the different music conditions on time or money spent can be determined, does this mean that background music does not matter? With the proposed model managers will be able to make such assessments more easily by determining the impact of background music on the affective state of their customers in the service setting and on their satisfaction afterwards.

### 3 The Role of Perceived Performance in the Satisfaction Process

In a number of studies, extremely high correlations between perceived performance and satisfaction measures were observed. This led to the notion that perceived performance is a direct causal antecedent of satisfaction. However, the review of these studies in Chapter II suggested that high correlations may have been caused by perceived performance measures that captured a high proportion of the satisfaction construct. This has often appeared as a problem with discriminant validity between the two measures (e.g. Churchill and Surprenant 1982). It is a general methodological issue, relevant to satisfaction research concerned with both products and services.

In this study great care was taken to measure value-free performance perceptions, rather than evaluations. The aim was to show that there is no direct causal relationship between perceived performance and satisfaction. In Chapter II the following hypothesis was advanced:

Hypothesis I:

Direct causal links between perceived performance and satisfaction do not exist; high positive correlations between the two constructs are merely an indication of inappropriate perceived performance measures.

### **3.1 Summary of Findings - Hypothesis I**

As expected, the results strongly supported the classical disconfirmation-of-expectations model with disconfirmation as the mediating variable between expectations, performance and satisfaction. These results are consistent with those of mainstream research in consumer satisfaction (e.g. Bitner 1990, Cadotte et al 1987, Oliver 1980a, Woodruff et al 1983, Tse and Wilton 1988).

However, in contrast to a number of studies (e.g. Churchill and Surprenant 1982, Oliver 1985, Oliver and DeSarbo 1988, Tse and Wilton 1988), a direct causal link between performance and satisfaction was clearly rejected. The data provided support for Hypothesis I. This is an important result, as it corroborates the proposed reinterpretation of the high correlations between perceived performance and satisfaction observed in numerous studies as being caused by inappropriate measures, rather than indicating a causal relationship. The implications of this reinterpretation are discussed in the next subsection.

### **3.2 Research Implications**

In many satisfaction-related studies, neither expectations nor perceived performance are measured directly. These measures may be superfluous if the research is not concerned either with examining the causes of disconfirmation (Olson and Dover 1979) or with perceived performance itself. However, when a

researcher aims at examining the causes of the level of perceived performance and disconfirmation experienced by consumers, or when the objective is to research the disconfirmation process in the context of theory testing, direct perceived performance and expectation measures are needed (Cadotte et al 1987).

The standard scaling approaches for perceived performance use semantic differential or Likert-type scales, which often include anchor words such as good/bad, friendly/unfriendly and terrible/excellent (Cadotte et al 1987). The results of this study confirm the need to employ scales that measure true beliefs rather than evaluations. Where possible, measures that cue the respondent to an objective level of an attribute performance should be developed. For instance, the speed of a service can be measured in time units rather than with a fast/slow scale, and menu variety can be measured in number of items on the menu instead of with a few/many scale.

This approach is difficult to implement when specific attributes do not have physical correlates (e.g. the taste of a meal, employee friendliness). Therefore, the aim of further research should be to develop performance measures that cue respondents to more objective levels of attribute performance than do the measures currently employed. Until improved measures are available, researchers should not rely solely on current measures. Instead, wherever possible, researchers should use performance manipulations, independent observations or

expert judgment as approximations for perceived performance.<sup>6)</sup> In cases where this is not possible, caution needs to be exercised when the interpretation of the data relies on perceived performance measures, especially those which apparently explain a large part of the variance in satisfaction [e.g. 88% in Churchill and Surprenant's (1982) and 65% in Tse and Wilton's study (1988)].

---

<sup>6</sup> This of course, has the drawback that consistency processes with expectations cannot be observed directly (see Chapter II, Section 1.1.3.1). However, these effects probably can be inferred from the relative impact of expectations and performance on disconfirmation.

## 4 Limitations

As with any research, there are limitations associated with the study reported here. Limitations inherent in the research design and in the model tested are discussed in this section.

### Design Limitations

The objectives of this study, specified in Chapter VI, were theory testing, rather than effects application. This led to selection of a true experimental design that put maximum emphasis on internal and construct validity. External validity was only of secondary importance. This means that caution needs to be exercised in any generalization of the results beyond the experimental setting, stimuli and population.

This limitation most seriously affects the results from manipulation of variance expectations. No previous work has been conducted into the effects of variance expectations on disconfirmation. In consequence, the findings of this study cannot be integrated into a well researched theoretical framework: its conclusions in relations to variance expectations have limited external validity. Replication and extension of the work, as proposed in the section on further research, are essential to overcoming this limitation.

The results concerning the other two principal research issues of this thesis, perceived performance measures and affect in the satisfaction process, fit well

within existing and established theories in marketing and psychology. Therefore, the generalizability of those findings to real-world service encounters can be proposed with greater confidence.

In the experiment conducted, subjects were asked to project themselves into a situation (consuming a homebanking service) and to role play (imagine themselves as being Mr Smith). The advantages and disadvantages of projecting and role playing have been fully discussed elsewhere (Sawer 1977, Surprenant and Churchill 1977) and the reasons for selecting these methods were presented in Chapter VI. The main disadvantages are a greater likelihood that demand characteristics will affect the results, and subjects may not be able to project themselves into an imaginary situation and respond as they would in a real situation. Feedback to the last question of the questionnaire (Appendix D: open-ended question in Part D, number 11) and informal feedback from the subjects after the experimental sessions did not give any indication that subjects had problems with projecting themselves into the situation described, or that demand characteristics had been created.

### Model Limitations

No single model can hope to capture all causes and consequences of a given phenomenon (Bitner 1987). This is true for the extended satisfaction model tested in this thesis (Figure 7-2). The literature on affect in services marketing has shown that perceived control, the design of the service setting and background music can all drive the affective state of consumers during the service encounter.

Psychologists have shown that there are many other potential determinants of the affective states of people. In this study, it was assumed that all of the potential determinants of affect, with the exception of disconfirmation, were either controlled for through standardization (setting design, simulated and projected consumption experience, interactions with the homebanking system) or randomized across all conditions through the random allocation of subjects.

## 5 Summary

This chapter presented the findings and implications of the three issues researched. The first two issues concern extension of the traditional disconfirmation model to take account of features typical of services, namely variance expectations and the experiential nature of the service encounter. The third issue is a methodological issue relevant to satisfaction research in general.

First, it was found that variance expectations can have an impact on the disconfirmation process. When the discrepancy between the mean expectation and perceived performance was small, a high variance expectation caused subjects to perceive positive disconfirmation which, in turn, led to higher satisfaction. This is an important finding since, for the majority of consumption experiences, expectations are met by performance. The finding suggests that a service firm can make its customers perceive positive disconfirmation and higher satisfaction by communicating the inherent heterogeneity in its service offer. The implications of this proposal were discussed with regard to the consumer choice process and the classic trade-off between customization and standardization of the service offer.

Second, it was shown that affect can be included in the traditional disconfirmation-of-expectations model. The data clearly supported the proposition that disconfirmation drives the affective state of people during a consumption experience, and that this affective state is, in turn, a direct determinant of

satisfaction expressed after consumption. This finding allows us to open up the disconfirmation paradigm and bring together theories developed in services marketing and psychology. These theories hold that there are numerous other stimuli beside disconfirmation that can drive affect and probably also consumer satisfaction. A number of managerial implications of the opened-up disconfirmation model were discussed.

Third, the data clearly lead to the rejection of an often proposed causal link from perceived performance to satisfaction. This suggests a reinterpretation of the frequently observed high correlation between the two variables in terms of being caused by measures with insufficient discriminant validity. It was therefore proposed that perceived performance measures that cue respondents to more objective levels of perceived performance should be developed and used in place of current measures.

# Appendices

## Appendix A:

**A Sample Leaflet Used for the Recruitment of Subjects**

# Marketing Study for a New Homebanking Service

Dear MBA29,

An American bank is planning to launch a PC-based homebanking service in the UK. This service will allow you to handle most of your bank transactions from home or office. All that is required is a PC and a modem that can be connected to the bank via your normal telephone line. This service will be provided 24 hours a day on 7 days a week

We are looking for 35 people, who would like to try this new service on a PC and fill in a questionnaire on how they feel about the service. If you would like to try your hand at it, the following times have been arranged:

Date: Tuesday, 3 July 1990,

at 12.00 am, 13.30 pm, 15.00 pm, 16.30 pm and 18.00 pm

Location: Room E202 (PC Teaching Centre), E Wing

Duration: 40 minutes

**Remuneration: Every participant will get £10 cash!!**

I would very much appreciate your participation. Please arrive on the time, because there will be a short presentation on homebanking right at the beginning.

I am looking forward to seeing you on Tuesday.

With best wishes,



Jochen Wirtz

# Appendix B:

## Introductory Presentation to Subjects

# Marketing Study for a New Homebanking Service

## Experimenter Instructions

(show headline of slide) Good evening and thank you very much for participating in our study. We have asked you to participate in this study because most of you are familiar with PC's and automatic teller machines.

First, let me introduce myself. My name is Jochen Wirtz, and I am conducting this study on behalf of an American bank, which is planning to launch a homebanking service in the UK. This service will allow private, self-employed and small business users to handle most of their daily bank transactions from their home or office. All that is required is a slightly adapted PC that can be connected to the bank via your normal telephone line.

Four different homebanking services have been developed and each service offers a number of different transactions. The drawback of the more elaborate homebanking services is that such systems require a longer response time. In other words, the more complicated and numerous the possible transactions that a homebanking system offers, the more time it takes to process and respond to each command or data input.

(show definition of response time) If you look at the slide, our definition of response time is the time the system requires to respond to each strike of the "ENTER" key.

We would like to know how you feel about the features offered and the corresponding response time of this version of the homebanking service. However, please don't compare this homebanking service with PC software like wordprocessors, Lotus or dBase. The technology involved for homebanking is far more complex. This is also a first test on how the system performs under market conditions.

Here is a short description of a potential user, Mr Smith (show the slide and read the text).

We would like you to imagine yourself as Mr Smith, while testing this service to give you a better feel of its possible applications.

Now, if you please take the questionnaire out of the envelope. Read read the first two pages carefully, as they are important for understanding the study, and then proceed with the questionnaire. Please tell me when you have completed the whole questionnaire and I will come and collect it. Thank you very much.

# Appendix C:

**Slide: A Potential User of the Homebanking Service**

# Marketing Study for a New Homebanking Service

Response time = time the system needs to respond to each time you hit the "ENTER" key to input data or to give commands to the system.

Mr Smith runs a small property developing business. He has to pay a number of bills and check his account balance daily. Furthermore, active cash management is quite important to Mr Smith, because large sums of money pass through his account.

# Appendix D:

## A Sample Questionnaire

# Marketing Study for a New Homebanking Service

## This study is structured in three parts:

In Part A, we provide you with information about this new homebanking service.

In Part B, you can test the Homebanking Service on a Simulator so that you become familiar with the service. During this simulation no real transactions are executed. We will then ask you to answer 5 questions after using the Simulator.

In Part C, you can test the real American Homebanking Service. The PC will be connected to a mainframe and your transactions will be executed. After using the real service, we will ask you to answer 14 questions.

## Instructions:

1. Please do not ask any questions, the questionnaire is self-explanatory, and please do not consult with others while using the service. If you have any problems using the Homebanking Service please ask the instructor.
2. Please do not rush through the questionnaire and read the information on page 2 carefully. You can take as much time as you wish. Previous sessions lasted between 40 to 50 minutes.
3. Please answer all the questions.
4. Many questions may appear somewhat similar. This is not an attempt to "trip you up," but merely reflects the fact that alternative phrasings mean different things to different people.

## Part A - Information About The American Homebanking Service

- Available transactions:**
- Obtain account details (e.g. account balance, credit limit, interest accrued)
  - Make bill payments
  - Make inter-account transfers (e.g. from chequing account to investment account)
  - Order cheque books and account statements.

**Response time:** The response time of a homebanking service is very important. This is because firstly, the charges are time dependent (they are the same as telephone charges for local calls) and secondly, one of the main reasons for using a home-banking service is to save time.

In its advertisements, the American Bank claims that its system needs **on average only one second** to respond to each data input (= each ENTER strike), and that this response time **can vary between immediate response and 15 seconds** between each time the service is used. This variance is due to peak and low times. For example, the response time can vary between different times during the day and different days of the week. The American Bank published the following test results on the response time of their system:

Table 1: Test results on the response time

Response time	% of trials falling in this category
0 - 2 seconds	74 %
3 - 5 seconds	17 %
6 - 10 seconds	6 %
11 - 15 seconds	3 %
16 - 20 seconds	-

**"As you can see in Table 1, in around three-quarters of all tests the system needed less than 2 seconds to respond. In the remaining quarter however, the system was slower, in some rare cases as slow as 15 seconds."**

Please **do not compare** the response time of this homebanking service with PC standard software you might be familiar with. Unlike wordprocessors, Lotus or dBase, the homebanking system has to **access a mainframe via a telephone line** to execute your commands, check account numbers, account balances, etc. This technology is **far more complicated** than mere PC-based software. Furthermore, however long the response time of a homebanking service, it will always be quicker than going round to your local bank.

## Part B - Testing the Simulator

Before making a decision Mr Smith tries out the Simulator that functions in exactly the same way as the real American Homebanking Service. However, the PC is not connected to a mainframe and no real transactions are executed. Please imagine you are Mr Smith and try out the Simulator, following the instructions below.

### First, Mr Smith wants to access the Simulator.

1. Press "ENTER" to start the Homebanking Service.
2. Select "SIMULATOR" from the menu (use arrow keys to highlight choice and press "ENTER").
3. Enter your user number: "75472900" and press "ENTER" (if you made a mistake press the "DELETE" key and try again).
4. Enter your password: "LBSFHRBR" and press "ENTER". If the account number and password are correct, the system will go to the "Main Menu". Try again, if you made a mistake.

### Second, Mr Smith wants to check his chequing account balance.

1. Select "ACCOUNT DETAILS AND STATEMENT" and press "ENTER".
2. Select "CHEQUING" account (number 75472901) and press "ENTER" (The account balance and today's turnover will be shown after execution of this command).
3. Now select "RETURN TO MAIN MENU" (use arrow keys to highlight choice and press "ENTER").

### Third, Mr Smith wants to pay a bill to one of his suppliers, Mr Green.

1. Select "MAKE BILL PAYMENTS" and press "ENTER".
2. Select "PAYMENTS TO CUSTOMER SELECTED ACCOUNTS" and press "ENTER".
3. Select "CHEQUING ACCOUNT" to be debited and press "ENTER".

4. Enter name of receiver: "Mr Green" and press "ENTER". (If you made a mistake, use the "Delete" key and try again.
5. Enter name of receiver bank: "Midland Bank" and press "ENTER".
6. Enter receiver bank code: "40 04 26" and press "ENTER".
7. Enter amount in \$ to be paid: "45.90" and press "ENTER".
8. Enter reference for receiver: "no 45787" and press "ENTER".
9. If all data entered are correct select "COMPLETE TRANSACTION" and press "ENTER".

If you want to amend or cancel the operation, select "AMEND TRANSACTION DETAILS" or "CANCEL TRANSACTION, RETURN TO MAIN MENU", press "ENTER" and try again.

10. Check the Bill Payment Confirmation. Select "COMPLETE TRANSACTION" if the data are correct, if not select "AMEND TRANSACTION" and try again.
11. Now select "RETURN TO MAIN MENU" and press "ENTER".
12. To leave the Simulator select "EXIT AMERICAN HOMEBANKING SYSTEM" and press "ENTER".
13. Do not "Quit" from the next menu and remain in the homebanking programme.

This introduction to the homebanking service is to give you an idea on how it works. Of course, there are many other possible transactions. For example: ordering cheque books and transferring money between your chequing and investment accounts to finetune your cash management.

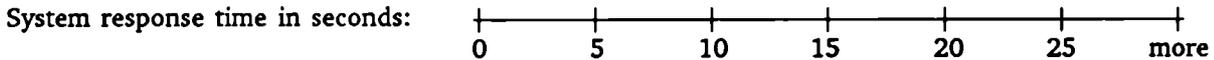
Please remain in the last menu while answering the questions on the next page.

- Please do not turn the page -  
- until you have concluded your test of the Simulator -

**Please answer the following 5 questions after using the Simulator:**

Please imagine that Mr Smith had also read the information page on the American Homebanking Service and that he also used the Simulator. Please answer the following six questions as if you were Mr Smith.

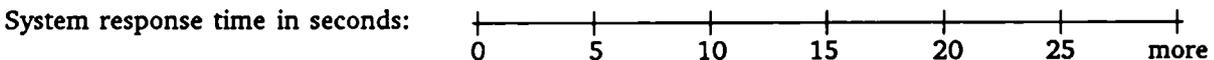
1. After reading the information page and using the Simulator, what would Mr Smith expect the average response time of the real American Homebanking Service after every "ENTER" strike to be? Please indicate on the following scale like this: (~~X~~)



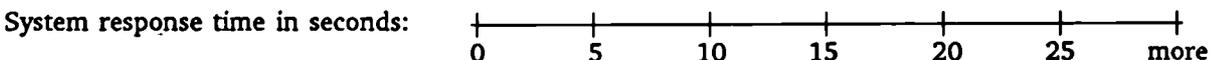
2. Please read the following statement carefully then circle the number which best describes your agreement or disagreement with the statement (e.g. 5=Strongly Agree)

	Strongly Agree	Agree	Agree nor Disagree	Disagree	Strongly Disagree
After reading the <u>information page</u> and using the <u>Simulator</u> Mr Smith believes that the response time of the <u>real American Homebanking System</u> could vary by more than 5 seconds each different time he uses the system	5	4	3	2	1

3. What would Mr Smith expect the maximum possible response time of the real American Homebanking System to be?



4. What would Mr Smith expect the minimum possible response time of the real American Homebanking System to be?



5. In what range would Mr Smith expect the most likely response time of the real American Homebanking System to be? Mr Smith would expect it to be between:

- |                          |                          |                          |                          |                          |                          |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> |
| 0-2<br>seconds           | 3-5<br>seconds           | 6-10<br>seconds          | 11-15<br>seconds         | 16-20<br>seconds         | 21 and more<br>seconds   |

## Part C - Testing the real American Homebanking Service

On the basis of the information provided on page 2 and his experience with the Simulator, Mr Smith has decided to buy the American Homebanking Service. Imagine that the Service has just been installed on the PC in his office and that he is now going to use it for the first time. Please imagine you are Mr Smith in his office and see how the real American Homebanking Service works.

This time the PC will be connected to a mainframe and the transactions will be performed. Please follow the instructions. They are very similar to the transactions tried with the Simulator.

### First, Mr Smith wants to access the American Homebanking Service.

1. Select "AMERICAN HOMEBANKING SYSTEM" from the menu and press "ENTER".
2. Enter your user number: "75472900" and press "ENTER" (if you made a mistake press the "DELETE" key and try again).
3. Enter your password: "LBSFHRBR" and press "ENTER". If the account number and password are correct, the system will go to the "Main Menu".

### Second, Mr Smith wants to check his chequing account balance.

1. Select "ACCOUNT DETAILS AND STATEMENT" and press "ENTER".
2. Select "CHEQUING" account (number 75472901) and press "ENTER" (The account balance and today's turnover will be shown after execution of this command).
3. Now select "RETURN TO MAIN MENU" and press "ENTER".

### Third, Mr Smith wants to pay a bill to the London Business School (LBS).

1. Select "MAKE BILL PAYMENTS" and press "ENTER".

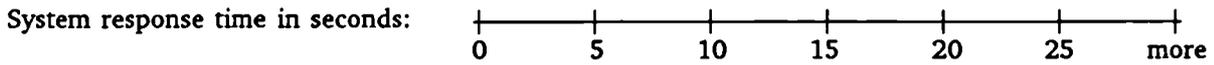
2. Select "PAYMENTS TO CUSTOMER SELECTED ACCOUNTS" and press "ENTER".
3. Select "CHEQUING ACCOUNT" to be debited and press "ENTER".
4. Enter name of receiver: "LBS" and press "ENTER".
5. Enter name of receiver bank: "Midland Bank" and press "ENTER".
6. Enter receiver bank code: "40 04 26" and press "ENTER".
7. Enter amount in \$ to be paid: "10.00" and press "ENTER".
8. Enter reference for receiver: "Test 4" and press "ENTER".
9. If all data entered are correct select "COMPLETE TRANSACTION" and press "ENTER".

If you want to amend or cancel the operation, select "AMEND TRANSACTION DETAILS" or "CANCEL TRANSACTION, RETURN TO MAIN MENU", press "ENTER" and try again.

10. Check the Bill Payment Confirmation. Select "COMPLETE TRANSACTION" if the data are correct, if not select "AMEND TRANSACTION" and try again.
11. Now select "RETURN TO MAIN MENU" and press "ENTER".
12. Now, please look at Today's Transactions of the chequing account. First select "ACCOUNT DETAILS AND STATEMENT". Then try to do the next steps without detailed instructions. After you have checked today's transactions, please return to the main menu and continue with the instructions below.
13. To leave the American Homebanking System select "EXIT AMERICAN HOMEBANKING SYSTEM" and press "ENTER".
14. Now select "QUIT AHBS" and press "ENTER".
15. To confirm that you want to quit the American Homebanking System select "YES" and press "ENTER".
16. Please do not "Quit" from the next menu and remain in the homebanking programme while answering the questions on the next pages.

Please answer the following 14 questions after testing the American Homebanking Service:

1. How long did the American Homebanking Service need approximately to respond to each "ENTER" strike. Please indicate on the following scale like this: (~~X~~)



2. How was the actual performance of the following features of the American Homebanking Service in comparison to Mr Smith's expectations? (Please tick)

	worse than expected		just as expected		better than expected
The visual appeal of the system was	___	___	___	___	___
The reliability of the service was	___	___	___	___	___
The ease of use was	___	___	___	___	___
The range of transactions possible was	___	___	___	___	___
The response time was	___	___	___	___	___
The security of the service was	___	___	___	___	___
Overall, the homebanking service was	___	___	___	___	___

3. In what range did the American Homebanking Service respond to each "ENTER" strike? (please tick)

- |                          |                          |                          |                          |                          |                          |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> |
| 0-2<br>seconds           | 3-5<br>seconds           | 6-10<br>seconds          | 11-15<br>seconds         | 16-20<br>seconds         | 21 and more<br>seconds   |

4. What feelings would Mr Smith have about his previous expectations after using the American Homebanking Service? (please tick)

In comparison to the American Homebanking Service, his expectations were

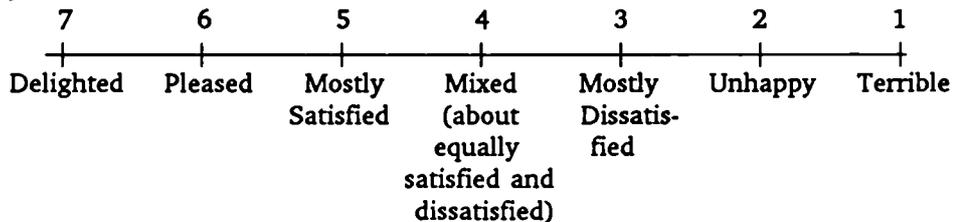
too high:  
It was poorer than he thought

Accurate:  
It was just as he expected

Too low:  
It was better than he thought

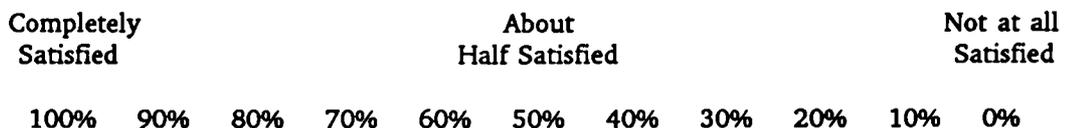
His expectation of the security of the system was	___	___	___	___	___	___
His expectation of the response time was	___	___	___	___	___	___
His expectation of the visual appeal of the system was	___	___	___	___	___	___
His expectation of the ease of use was	___	___	___	___	___	___
His expectation of the range of transactions possible was	___	___	___	___	___	___
His expectation of the reliability of the system was	___	___	___	___	___	___
Overall, his expectations about the homebanking service were	___	___	___	___	___	___

5. Overall, how would Mr Smith feel about the experience with the American Homebanking Service? (please circle)



- A Neutral (neither satisfied nor dissatisfied)
- B Mr Smith never thought about it

6. Overall, Mr Smith would feel? (please circle)



7. Now think about the feelings, moods and emotions Mr Smith might have experienced while using the American Homebanking Service for the first time in his office. Take about half a minute to get into the mood of the situation. Then rate the feelings with the adjective pairs below. Some of the pairs might seem unusual, but you will probably feel more one way than the other.

For each pair, put a tick (Example :      :  :     ) closer to the adjective which you believe best describes Mr Smith's feelings, moods and emotions. The more appropriate that adjective seems, the closer you should put your tick to it. Please make sure that you have given a tick for each pair.

Bored	<u>    </u> : <u>    </u>	Relaxed
Hopeful	<u>    </u> : <u>    </u>	Despairing
Submissive	<u>    </u> : <u>    </u>	Dominant
Calm	<u>    </u> : <u>    </u>	Excited
Guided	<u>    </u> : <u>    </u>	Autonomous
Happy	<u>    </u> : <u>    </u>	Unhappy
Jittery	<u>    </u> : <u>    </u>	Dull
Melancholic	<u>    </u> : <u>    </u>	Contented
Important	<u>    </u> : <u>    </u>	Awed
No Control	<u>    </u> : <u>    </u>	In Control
Aroused	<u>    </u> : <u>    </u>	Unaroused
Pleased	<u>    </u> : <u>    </u>	Annoyed
Relaxed	<u>    </u> : <u>    </u>	Stimulated
Wide awake	<u>    </u> : <u>    </u>	Sleepy
Influential	<u>    </u> : <u>    </u>	Influenced
Unsatisfied	<u>    </u> : <u>    </u>	Satisfied
Controlling	<u>    </u> : <u>    </u>	Controlled
Sluggish	<u>    </u> : <u>    </u>	Frenzied

8. A number of features of the American Homebanking Service are listed below. Please imagine how Mr Smith would feel about them. (please tick)

Range of transactions possible	<u>    </u> Like	<u>    </u> Dislike					
Ease of use	<u>    </u> Like	<u>    </u> Dislike					
Visual appeal of the system	<u>    </u> Like	<u>    </u> Dislike					
Security of the system	<u>    </u> Like	<u>    </u> Dislike					
Response time of the system	<u>    </u> Like	<u>    </u> Dislike					
Reliability of the system	<u>    </u> Like	<u>    </u> Dislike					
The American Homebanking System as a whole	<u>    </u> Like	<u>    </u> Dislike					

9. Below is a list of words that can be used to describe situations and how people feel in those situations. Please assess how accurately each of the words below describes Mr Smith's situation and his feelings while using the American Homebanking Service for the first time in his office.

You may need some imagination and there are no right or wrong answers. Use the following 1-8 rating scale for your answer, and please make sure that you have given an answer for each word. (please circle)

	extremely <u>in</u> accurate	very <u>in</u> accurate	quite <u>in</u> accurate	slightly <u>in</u> accurate	slightly accurate	quite accurate	very accurate	extremely accurate
nice	1	2	3	4	5	6	7	8
forceful	1	2	3	4	5	6	7	8
dissatisfying	1	2	3	4	5	6	7	8
unpleasant	1	2	3	4	5	6	7	8
intense	1	2	3	4	5	6	7	8
pretty	1	2	3	4	5	6	7	8
pleasant	1	2	3	4	5	6	7	8
lazy	1	2	3	4	5	6	7	8
alive	1	2	3	4	5	6	7	8
uncomfortable	1	2	3	4	5	6	7	8
drowsy	1	2	3	4	5	6	7	8
inactive	1	2	3	4	5	6	7	8
displeasing	1	2	3	4	5	6	7	8
arousing	1	2	3	4	5	6	7	8
active	1	2	3	4	5	6	7	8
pleasing	1	2	3	4	5	6	7	8
idle	1	2	3	4	5	6	7	8
repulsive	1	2	3	4	5	6	7	8
slow	1	2	3	4	5	6	7	8
beautiful	1	2	3	4	5	6	7	8

10. Now, please indicate how important each of the following features or characteristics of homebanking are to Mr Smith in general. Please divide 100 points between the following six features to show their importance to Mr Smith.

- Range of transactions possible \_\_\_\_\_
- Ease of use \_\_\_\_\_
- Visual appeal of the system \_\_\_\_\_
- Security of the system \_\_\_\_\_
- Response time of the system \_\_\_\_\_
- Reliability of the system \_\_\_\_\_

11. Below are some statements which describe how Mr Smith might feel after he used the American Homebanking Service for the first time. Please read each statement carefully then circle the number which best describes your agreement or disagreement with the statement (e.g. 1=Strongly Agree, 5=Strongly Disagree).

	Strongly Agree	Agree	Agree nor Disagree	Disagree	Strongly Disagree
Mr Smith felt in control of the situation	1	2	3	4	5
If he could do it over again, he'd buy a different service	1	2	3	4	5
Mr Smith feels that the reliability of the system is very poor	1	2	3	4	5
He truly enjoys this homebanking service	1	2	3	4	5
Mr Smith felt unable to influence the way things were	1	2	3	4	5
He feels bad about his decision to buy this service	1	2	3	4	5
Mr Smith feels that the ease of use of the system is excellent	1	2	3	4	5
He feels that the response time of the system is very bad	1	2	3	4	5
He is <u>not</u> happy that he bought this service	1	2	3	4	5
He feels that the range of possible transactions is very wide	1	2	3	4	5
His choice to buy this service was a wise one	1	2	3	4	5
Sometimes, he will have mixed feelings about keeping it	1	2	3	4	5
Mr Smith felt it was difficult to achieve what he desired	1	2	3	4	5
Mr Smith is satisfied with his decision to buy this service	1	2	3	4	5
Mr Smith felt it was easy to get his own way	1	2	3	4	5
Owning this service system has been a good experience for Mr Smith	1	2	3	4	5
This system hasn't worked out as well as he thought it would	1	2	3	4	5
He feels that the security of the system is very high	1	2	3	4	5
This homebanking service is exactly what Mr Smith needs	1	2	3	4	5
This is one of the best homebanking services Mr Smith could have chosen	1	2	3	4	5
He feels that the visual appeal of the system is very poor	1	2	3	4	5
Mr Smith is sure that it was the right thing to choose this service	1	2	3	4	5

12. A number of features of the American Homebanking Service are listed below. Please imagine how Mr Smith would feel about them. (please tick like this:   ::  )

Ease of use	delighted	<u>  </u> : <u>  </u>	terrible
The Homebanking Service as a whole	delighted	<u>  </u> : <u>  </u>	terrible
Security of the system	delighted	<u>  </u> : <u>  </u>	terrible
Range of transactions possible	delighted	<u>  </u> : <u>  </u>	terrible
Response time of the system	delighted	<u>  </u> : <u>  </u>	terrible
Reliability of the service	delighted	<u>  </u> : <u>  </u>	terrible
Visual appeal of the service	delighted	<u>  </u> : <u>  </u>	terrible

13. Thinking back . . . If Mr Smith had to do it all over again, what are the chances in 10 that he would have chosen this homebanking service again? (Please tick one number below to indicate the chance)

<u>  </u> 0	<u>  </u> 1	<u>  </u> 2	<u>  </u> 3	<u>  </u> 4	<u>  </u> 5	<u>  </u> 6	<u>  </u> 7	<u>  </u> 8	<u>  </u> 9	<u>  </u> 10
(No chance he'd do it again)					(50-50 Chance)			(Certain he'd do it again)		

14. Finally, considering everything, how satisfied would Mr Smith be with the American Homebanking Service? (please tick):

<u>  </u>	<u>  </u>	<u>  </u>	<u>  </u>
very <u>dissatisfied</u>		neither satisfied nor dis- satisfied	very satisfied



9. We are interested to know how YOU feel in general, that is most of the time. Each pair of words below describes a common feeling. Some of the pairs might seem unusual, but you may generally feel more one way than the other. So, for each pair, put a tick (Example: \_\_\_:  : \_\_\_) to show how YOU feel IN GENERAL. Please take your time so as to arrive to a real characteristic description of your feelings.

Sleepy	___:___:___:___:___:___:___	Wide awake
Despairing	___:___:___:___:___:___:___	Hopeful
No Control	___:___:___:___:___:___:___	In Control
Autonomous	___:___:___:___:___:___:___	Guided
Stimulated	___:___:___:___:___:___:___	Relaxed
Unhappy	___:___:___:___:___:___:___	Happy
Contented	___:___:___:___:___:___:___	Melancholic
Dull	___:___:___:___:___:___:___	Jittery
Annoyed	___:___:___:___:___:___:___	Pleased
Dominant	___:___:___:___:___:___:___	Submissive
Satisfied	___:___:___:___:___:___:___	Unsatisfied
Controlled	___:___:___:___:___:___:___	Controlling
Unaroused	___:___:___:___:___:___:___	Aroused
Frenzied	___:___:___:___:___:___:___	Sluggish
Awed	___:___:___:___:___:___:___	Important
Relaxed	___:___:___:___:___:___:___	Bored
Influential	___:___:___:___:___:___:___	Influenced
Excited	___:___:___:___:___:___:___	Calm

10. How frequently do you use the following computer equipment, software and services? (please tick)

	never	rarely	at least once a month	at least once a week	nearly every day
Personal computer (PC)	___	___	___	___	___
Wordprocessor or graphics software	___	___	___	___	___
Lotus 123 or dBase	___	___	___	___	___
Mainframe	___	___	___	___	___
Cash money from Automatic Teller Machines (ATM's)	___	___	___	___	___
Transfer money between current and savings or investment accounts on ATM's	___	___	___	___	___
Homebanking services	___	___	___	___	___
On-line services like database searches or ticket booking	___	___	___	___	___

11. If you have any further comment you would like to add, please use the space below and/or the other side of this page.

Thank you very much for your cooperation.

- End -

# Appendix E:

## Selection of Screens of the Homebanking System

# AMERICAN HOME BANKING SYSTEM

WELCOME TO AHBS

Press 'Enter' to start AHBS

DATE: 5-8-1991

PLEASE WAIT

# AMERICAN HOME BANKING SYSTEM

Please select your preference

## Simulator

American Homebanking System (AHBS)

Quit AHBS

Use arrow keys to select choice, then press "Enter"

DATE: 5-8-1991

PLEASE WAIT

TIME: 10:57:50

# AMERICAN HOME BANKING SYSTEM

## SIMULATION

### GATEWAY TO AMERICAN HOME BANKING

This gateway available to registered users only

Enter your account number, then press "Enter"

Enter your password

Press "Del" key to clear box and repeat entry

Type "exit" to return to last screen

DATE: 5-8-1991

PLEASE WAIT

# AMERICAN HOME BANKING SYSTEM

## SIMULATION

MAIN MENU - Select an activity:

ACCOUNT DETAILS AND STATEMENT

MAKE BILL PAYMENTS

INTER-ACCOUNT TRANSFER

SERVICE REQUESTS

CHANGE PASSWORD

EXIT AMERICAN HOME BANKING SYSTEM

Use arrow keys to select choice, then press "Enter"

DATE: 5-8-1991

PLEASE WAIT

TIME: 11:212

# AMERICAN HOME BANKING SYSTEM

## SIMULATION

Statement for which account?

- 75472901 Chequing
- 75472902 Savings
- 75472903 Investment
- 75472904 Currency
- 75472905 Joint

Use arrow keys to select choice, then press "Enter"

Use "Esc" to cancel operation. Press "F1" to return menu

DATE: 5:8:1991

PLEASE WAIT

TIME: 11:2:49

# AMERICAN HOME BANKING SYSTEM

## SIMULATION

Account No. 75472901 At close of business 5:8:1991

Date	Details	Amount	Balance
5:8:1991	Net accruals	\$ 7.34	\$ 50.00

Return to account details  
Return to main menu

Select choice, then press "Enter"

DATE: 5:8:1991

PLEASE WAIT

TIME: 11:4:34

# AMERICAN HOME BANKING SYSTEM

MAIN MENU - Select an activity.

ACCOUNT DETAILS AND STATEMENT  
**MAKE BILL PAYMENTS**  
INTER-ACCOUNT TRANSFER  
SERVICE REQUESTS  
CHANGE PASSWORD  
EXIT AMERICAN HOME BANKING SYSTEM

Use arrow keys to select choice, then press "Enter"

DATE: 5/8/1991

TIME: 11:10:1

# AMERICAN HOME BANKING SYSTEM

Transfer funds from which account?

75472901 Chequing  
75472902 Savings  
75472903 Investment  
75472904 Currency  
75472905 Joint

Use arrow keys to select choice, then press "Enter"

Use "Esc" to cancel operation or return to main menu

DATE: 5/8/1991

PLEASE WAIT

TIME: 11:11:9

# AMERICAN HOME BANKING SYSTEM

## Payment to customer selected accounts

Enter name of receiver: **LES**  
Receiver Bank: **MIDLAND BANK**  
Receiver Bank Code: **40 04 28**  
Amount: **10.00** Reference: **TEST 4**

Press " Del " key to clear box and repeat entry

Enter details then press "Enter" key

DATE: 5-6-1991

PLEASE WAIT

TIME: 11:14:41

# AMERICAN HOME BANKING SYSTEM

## Bill payment confirmation

From : 754 0001 (Customer)  
To : LES 00 01 28  
Amount : 10.00

### Complete transaction

Select choice, then press "Enter"

Amend transaction details

Cancel transaction, return to main menu

DATE: 5-6-1991

PLEASE WAIT

TIME: 11:17:31

# Appendix F:

The Printout of the Lisrel Solution for the Final Model  $M_3$

STANDARDIZED RESIDUALS

	DDQ1	DDQ2	AROUS1	AROUS2	PLEAS1	PLEAS2
DDQ1	-.298					
DDQ2	-.303	-.318				
AROUS1	-1.240	-1.515	.003			
AROUS2	.310	.796	.091	.013		
PLEAS1	.749	.223	.769	.793	-.048	
PLEAS2	-.280	-.908	-2.298	-1.472	.674	-.023
SJLS12	-.153	-.752	.934	1.355	-1.284	.868
SVERBAL	.659	.184	.619	-.577	-.558	-.073
MEXPECT	.064	-.360	-1.669	.670	-.486	.370
MUAR	-.621	.113	.519	1.365	.934	.726
MPERF	.536	-.071	.827	.762	.199	1.528

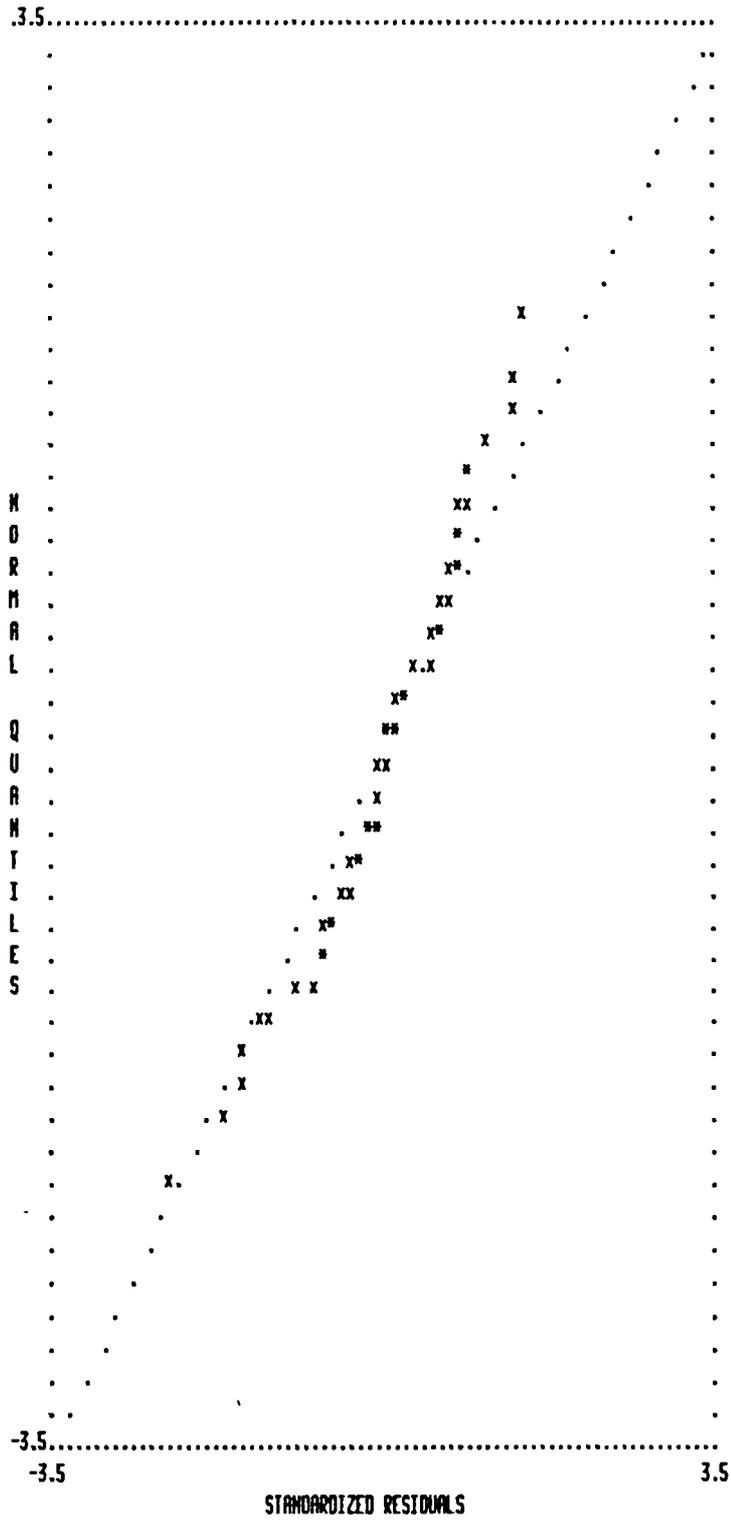
  

	SJLS12	SVERBAL	MEXPECT	MUAR	MPERF
SJLS12	-.048				
SVERBAL	-.248	-.105			
MEXPECT	-.459	-.030	.000		
MUAR	.838	-.608	.000	.000	
MPERF	1.127	.040	.380	.180	.000

SUMMARY STATISTICS FOR STANDARDIZED RESIDUALS:

SMALLEST STANDARDIZED RESIDUAL = -2.298  
 MEDIAN STANDARDIZED RESIDUAL = .008  
 LARGEST STANDARDIZED RESIDUAL = 1.528

Q P L O T O F S T A N D A R D I Z E D R E S I D U A L S



CORRELATION MATRIX OF ETA AND KSI

	<u>DISC</u>	<u>AROUS</u>	<u>PLEAS</u>	<u>SATISF</u>	<u>MEXPECT</u>	<u>MVAR</u>
DISC	1.000					
AROUS	.426	1.000				
PLEAS	.463	.640	1.000			
SATISF	.570	.591	.882	1.000		
MEXPECT	.440	-.086	.031	.115	1.000	
MVAR	.136	.078	.076	.087	.000	1.000
MPERF	-.401	-.231	-.223	-.258	.000	.000

MODIFICATION INDICES

MODIFICATION INDICES FOR LAMBDA Y

	<u>DISC</u>	<u>AROUS</u>	<u>PLEAS</u>	<u>SATISF</u>
DOQ1	.000	.013	.405	.405
DOQ2	.000	.015	.184	.190
AROUS1	2.533	.000	.036	.111
AROUS2	2.533	.000	.036	.116
PLEAS1	.588	1.097	.000	2.398
PLEAS2	.588	4.161	.000	.436
SALS12	.482	2.680	.482	.000
SVERBAL	.482	.924	.482	.000

MODIFICATION INDICES FOR LAMBDA X

	<u>MEXPECT</u>	<u>MVAR</u>	<u>MPERF</u>
MEXPECT	.000	.224	.457
MVAR	.000	.000	.033
MPERF	.144	.033	.000

MODIFICATION INDICES FOR BETA

	<u>DISC</u>	<u>AROUS</u>	<u>PLEAS</u>	<u>SATISF</u>
DISC	.000	.196	.317	.451
AROUS	.000	.000	.000	.821
PLEAS	.000	.000	.000	.847
SATISF	.000	.865	.000	.000

MODIFICATION INDICES FOR GAMMA

	MEXPECT	MUAR	MPERF
DISC	.000	.000	.000
AROUS	.000	1.168	.353
PLEAS	.000	.095	.295
SATISF	.036	.299	.034

MODIFICATION INDICES FOR PHI

	MEXPECT	MUAR	MPERF
MEXPECT	.000		
MUAR	.000	.000	
MPERF	.144	.033	.000

MODIFICATION INDICES FOR PSI

	DISC	AROUS	PLEAS	SATISF
DISC	.000			
AROUS	.032	.000		
PLEAS	.169	.000	.000	
SATISF	.210	.821	.847	.000

MAXIMUM MODIFICATION INDEX IS 4.16 FOR ELEMENT ( 6, 2) OF LAMBDA Y

# Appendix G:

## Further Lisrel Analysis for Testing Hypothesis I

Table A-1: Specifications of the Variables for Model  $M_{b1}$  With Included Manipulation Checks (Refer to Figure A-1)

---

$\xi_1$ = Expected Performance Manipulation	
$\xi_2$ = Expected Variance Manipulation	
$\xi_3$ = Performance Manipulation	
$\eta_1$ = Expected Performance	$y_1$ = EPERF1 $y_2$ = EPERF2
$\eta_2$ = Expected Variance	$y_3$ = EVAR1 $y_4$ = EVAR2
$\eta_3$ = Perceived Performance	$y_5$ = PPERF1 $y_6$ = PPERF2
$\eta_4$ = Disconfirmation	$y_7$ = DOQ1 $y_8$ = DOQ1
$\eta_5$ = Arousal	$y_9$ = AROUS1 $y_{10}$ = AROUS2
$\eta_6$ = Pleasure	$y_{11}$ = PLEAS1 $y_{12}$ = PLEAS2
$\eta_7$ = Satisfaction	$y_{13}$ = SWLS12 $y_{14}$ = SVERBAL

---

Table A-2: Estimated Parameters for the Rival Model  $M_{b12}$

Parameter	Lisrel Estimate	Standard Error	T-Value <sup>1)</sup>	Standardized Value
$\lambda_{x11}$	1.00 <sup>2)</sup>			1.00
$\lambda_{x22}$	1.00			1.00
$\lambda_{x33}$	1.00			1.00
$\theta_{\delta11}$	0			0
$\theta_{\delta22}$	0			0
$\theta_{\delta33}$	0			0
$\Phi_{11}$	1.00			1.00
$\Phi_{22}$	1.00			1.00
$\Phi_{33}$	1.00			1.00
$\lambda_{y11}$	1.00			.90
$\lambda_{y21}$	1.02	.07	15.3	.91
$\lambda_{y32}$	1.00			.55
$\lambda_{y42}$	1.50	.28	5.2	.80
$\lambda_{y53}$	1.00			.94
$\lambda_{y63}$	.97	.06	15.2	.91
$\lambda_{y74}$	1.00			.92
$\lambda_{y84}$	1.03	.07	14.4	.95
$\lambda_{y95}$	1.00			.73
$\lambda_{y105}$	1.25	.18	7.1	.92
$\lambda_{y116}$	1.00			.87
$\lambda_{y126}$	.94	.09	10.6	.82
$\lambda_{y137}$	1.00			.93
$\lambda_{y147}$	.95	.07	14.0	.88
$\theta_{\epsilon11}$	.20	.04	5.6	.20
$\theta_{\epsilon22}$	.17	.03	5.1	.17
$\theta_{\epsilon33}$	.70	.10	6.9	.70
$\theta_{\epsilon44}$	.35	.11	3.2	.35
$\theta_{\epsilon55}$	.12	.04	3.0	.12
$\theta_{\epsilon66}$	.17	.04	4.3	.17
$\theta_{\epsilon77}$	.17	.05	3.6	.17
$\theta_{\epsilon88}$	.12	.05	2.5	.12
$\theta_{\epsilon99}$	.46	.09	5.5	.46
$\theta_{\epsilon1010}$	.16	.10	1.7	.16
$\theta_{\epsilon1111}$	.24	.05	4.5	.24
$\theta_{\epsilon1212}$	.33	.06	5.6	.33
$\theta_{\epsilon1313}$	.14	.04	3.3	.14
$\theta_{\epsilon1414}$	.22	.04	5.1	.22

cont...

<sup>1</sup> All t-values are significant at the .05 level (d.f.=121, one-tailed).

<sup>2</sup> The values 1.00 and 0 in this column were fixed by the analyst.

Parameter	Lisrel Estimate	Standard Error	T-Value	Standardized Value
$\gamma_{11}$	.82	.05	15.8	.91
$\gamma_{41}$	.41	.07	5.6	.44
$\gamma_{51}$	-.25	.08	-3.2	-.34
$\gamma_{61}$	-.19	.09	-2.2	-.22
$\gamma_{22}$	.37	.07	5.1	.67
$\gamma_{42}$	.13	.07	1.8	.14
$\gamma_{33}$	.63	.06	10.9	.67
$\gamma_{43}$	-.37	.07	-5.1	-.40
$\beta_{21}$	.20	.06	3.2	.33
$\beta_{31}$	.49	.07	7.1	.47
$\beta_{54}$	.46	.10	4.7	.58
$\beta_{64}$	.53	.10	5.2	.56
$\beta_{74}$	.21	.07	2.9	.21
$\beta_{76}$	.84	.09	8.9	.79
$\psi_{11}$	.13	.03	4.2	.16
$\psi_{22}$	.11	.05	2.3	.37
$\psi_{33}$	.29	.05	5.5	.33
$\psi_{44}$	.53	.09	6.1	.63
$\psi_{55}$	.39	.09	4.2	.73
$\psi_{66}$	.57	.09	5.4	.75
$\psi_{77}$	.16	.05	3.2	.19
$\psi_{21}$	.07	.02	2.3	.11
$\psi_{65}$	.25	.07	3.7	.38
$\chi^2 = 118.2$ d.f. = 109    p = .26    AdjGFI = .87    RMSR = .05				

Figure A-1: Hypothesized Model (incl Manipulation Checks)

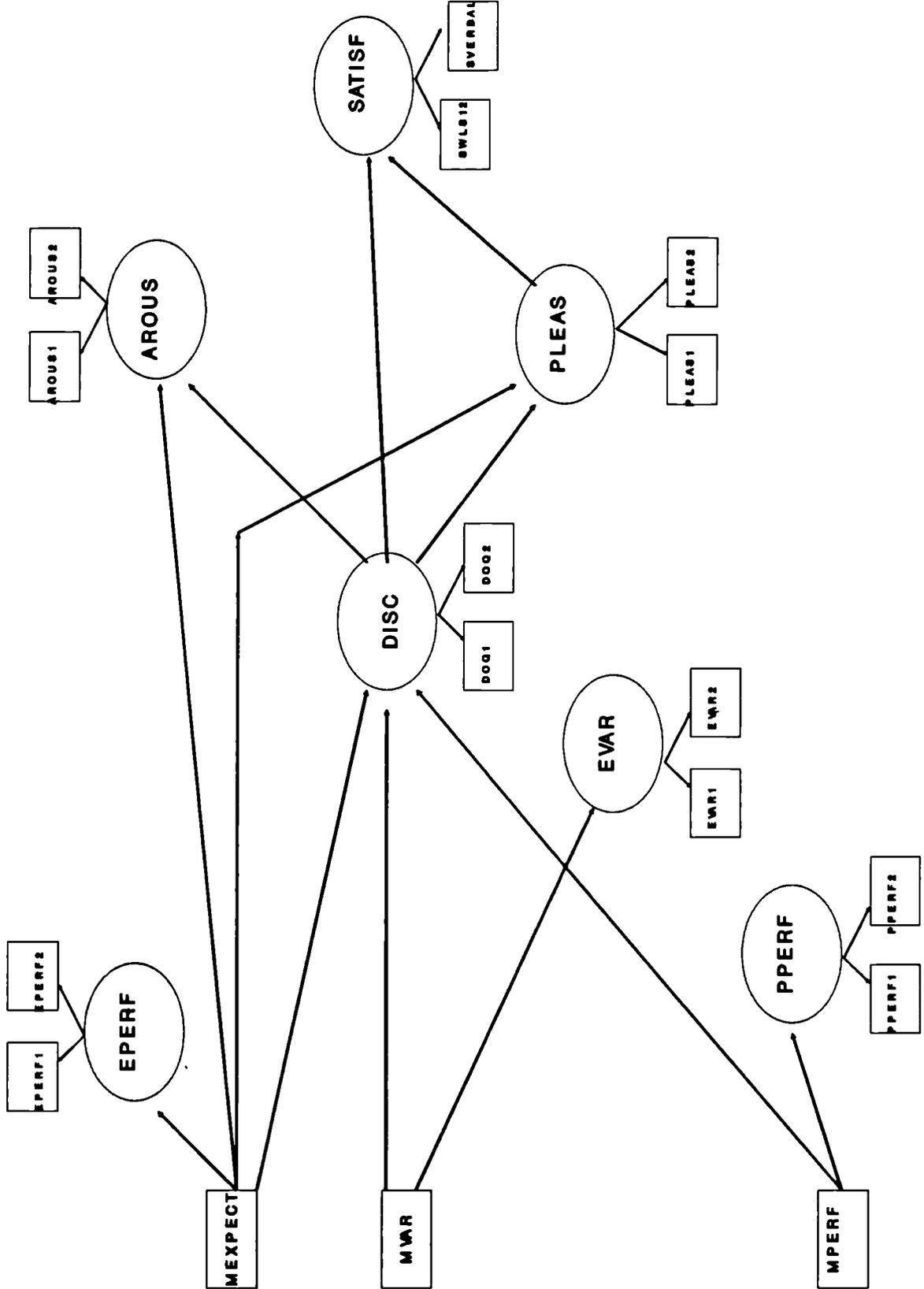
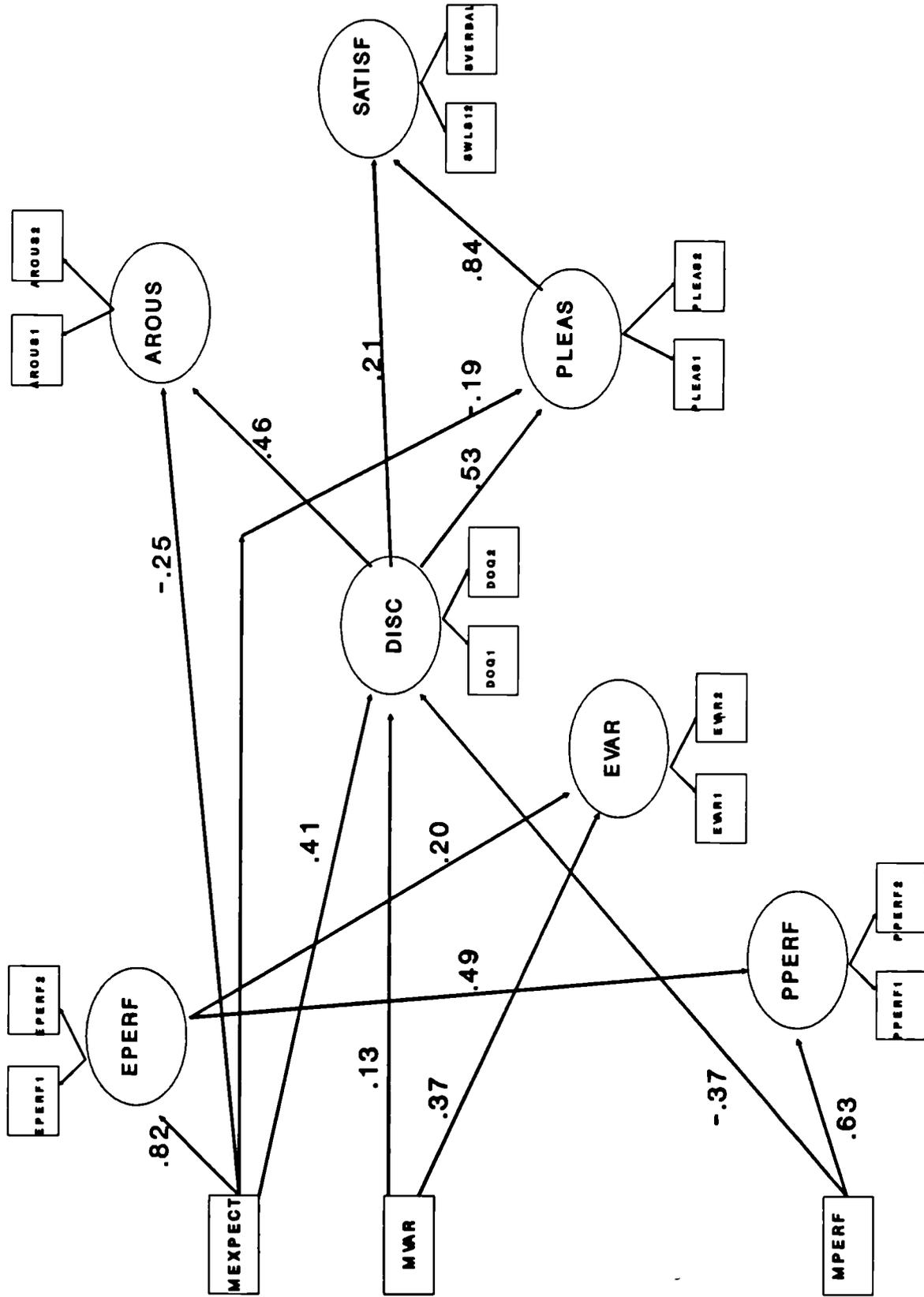


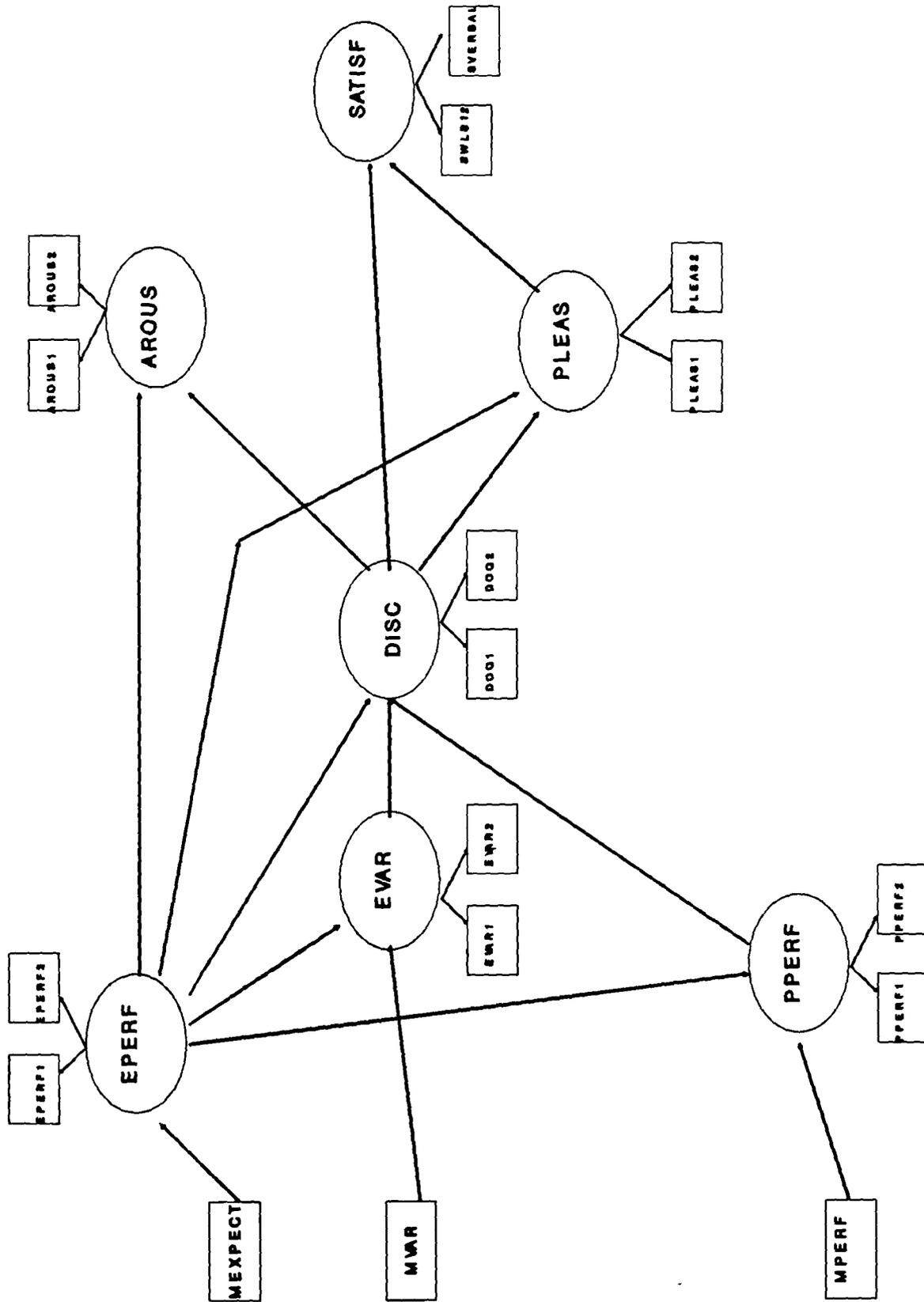
Figure A-2: Fitted Model (incl Manipulation Checks)



# Appendix H:

## Manipulation Checks as Key Variables in the Lisrel Model

Figure A-3: Rival Hypothesized Model





# References

- Abelson, Robert F. (1976), "Script Processing in Attitude Formation and Decision Making," in Cognition and Social Behavior, J.S. Carroll and J.S. Payne, eds., Hillsdale, NJ: Erlbaum.
- \_\_\_\_\_, E. Aronson, W.J. McGuire, T.M. Newcomb and M.J. Rosenberg (1968), Theories of Cognitive Consistency: A Sourcebook. Chicago: Rand McNally.
- Anderson, Rolph E. (1973), "Consumer Dissatisfaction: The Effect of Disconfirmed Expectancy on Perceived Product Performance," Journal of Marketing, Vol. 10 (February), 38-44.
- Anderson, N.H. (1974), "Algebraic Models in Perception," in Handbook of Perception 2, E.C. Carteremette and M.P. Friedman, eds., NY: Academic Press, 215-98.
- Andrews Frank M. and Stephen B. Withey (1976), Social Indicators of Well-Being. NY: Plenum Press.
- Averill, J.R. (1973), "Personal Control Over Aversive Stimuli and Its Relationship to Stress," Psychological Bulletin, Vol. 80, No. 4, 286-303.
- \_\_\_\_\_, L. O'Brien and G.W. DeWitt (1977), "The Influence of Response Effectiveness on the Preference for Warning and on Psychophysiological Stress Reactions," Journal of Psychology, Vol. 45, 395-418.
- Bagozzi, Richard P. (1981), Causal Models in Marketing. NY: John Wiley & Sons.
- \_\_\_\_\_, and Youjae Yi (1989), "On the Use of Structural Equation Models in Experimental Designs," Journal of Marketing Research, Vol. 26 (August), 271-84.
- Bateson, John E.G. (1977), "Do We Need Service Marketing?," in Marketing Consumer Services: New Insights, P. Eiglier, E. Langeard, C.H. Lovelock, J.E.G. Bateson and R.F. Young, eds., Cambridge, MA: Marketing Science Institute.
- \_\_\_\_\_, (1985a), "Researching the Service Consumer," in Services Marketing in a Changing Environment, Thomas M. Bloch, Gregory D. Upah and Vallerie A. Zeithaml, eds., Chicago: American Marketing Association.

- \_\_\_\_\_ (1985b), "Perceived Control and the Service Encounter," in The Service Encounter: Managing Employee/Customer Interaction in Service Businesses, J.A. Czepiel, M.R. Solomon and C. Surprenant, eds., Lexington, MA: Lexington Books.
- \_\_\_\_\_ (1989), Managing Services Marketing. Chicago: The Dryden Press.
- \_\_\_\_\_ and Michael Hui (1987), "A Model for Crowding in the Service Experience: Empirical Findings," in Integration for Competitive Advantage, J.A. Czepiel, C.A. Congram and J. Shanahan, eds., Chicago: American Marketing Association.
- \_\_\_\_\_ and \_\_\_\_\_ (1988), "Perceived Control as a Crucial Dimension of the Service Experience: An Experimental Study," in Add Value to Your Service, M. English, T. Gillett, T. Powell, T. Stephens, and C. Surprenant, eds., Chicago: American Marketing Association.
- Batra, Rajeev (1986), "Affective Advertising: Role, Process and Measurement," in The Role of Affect in Consumer Behaviour, Robert A. Peterson, Wayne D. Hoyer and William R. Wilson, eds., MA: Lexington Books, 87-118.
- \_\_\_\_\_ and Douglas M. Stayman (1990), "The Role of Mood in Advertising," in Journal of Consumer Research, Vol. 17, 203-14.
- Bauer, R.A. (1960), "Consumer Behaviour as Risk-Taking," in Dynamic Marketing for a Challenging World, R. S. Hancock, ed., Chicago: American Marketing Association.
- Bearden, William D. and Jesse E. Tell (1983), "Selected Determinants of Consumer Satisfaction and Complaint Reports," Journal of Marketing Research, Vol. 20 (November), 21-8.
- Belizzi, Joseph A., Ayn E. Crowley and Ronald W. Hasty (1983), "The Effects of Color in Store Design," in Journal of Retailing, Vol. 59 (Spring), No. 1, 21-45.
- Berlyne, D.E. (1970), "Novelty, Complexity and Hedonic Value," in Perception and Psychophysics, Vol. 8, 279-86.
- Berry, Leonard L. (1980), "Services Marketing is Different," Business, Vol. 30 (May- June), 24-9.
- \_\_\_\_\_ (1981), "The Employee as Customer," in Journal of Retail Banking, Vol. 3, No. 1, 33-40.
- Bitner, Mary J. (1987), "Contextual Cues and Consumer Satisfaction: The Role of Physical Surroundings and Employee Behaviors in Service Settings," Unpublished doctoral dissertation, University of Washington.

- \_\_\_\_\_ (1990), "Evaluating Service Encounters: The Effects of Physical Surroundings and Employee Responses," in Journal of Marketing, Vol. 54 (April), 69-82.
- \_\_\_\_\_, Bernard H. Booms and Mary Stanfield Tetreault (1990), "The Service Encounter. Diagnosing Favorable and Unfavorable Incidents," in Journal of Marketing, Vol. 54 (January), 71-84.
- Bone, Paula Fitzgerald, Subhash Sharma and Terence A. Shimp (1989), "A Bootstrap Procedure for Evaluating Goodness-of-Fit Indices of Structural Equation and Confirmatory Factor Models," Journal of Marketing Research, Vol. 26 (February), 105-11.
- Bowen, David E. and Benjamin Schneider (1985), "Boundary-Spanning-Role Employees and the Service Encounter: Some Guidelines for Management and Research," in The Service Encounter, John A. Czepiel, Michael R. Solomon and Carol F. Surprenant, eds., Lexington, MA: Lexington Books, 127-47.
- Bower, Gordon H. (1981), "Emotional Mood and Memory," American Psychologist, 36, 129-48.
- \_\_\_\_\_ and Paul R. Cohen (1982), "Emotional Influences in Memory and Thinking: Data and Theory," in Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition, M.S. Clark and S.T. Fiske, eds., Hillsdale, NJ: Erlbaum, 291-331.
- Breckler, Steven J. (1990), "Applications of Covariance Structure Modelling in Psychology: Cause for Concern?" in Psychological Bulletin, Vol. 107, No. 2, 260-73.
- Bringham, J.C. (1971), "Ethnic Stereotypes," in Psychological Bulletin, Vol. 76, 15-38.
- Brown, C. Marlin (1988), Human-Computer Interface Guidelines. New Jersey: Ablex Publishing Corporation.
- Bruner, Gordon C. (1990), "Music, Mood, and Marketing," in Journal of Marketing, Vol. 54 (October), 94-104.
- Cacioppo John T., Mary E. Losch, Louis G. Tassinary and Richard E. Petty (1986), "Properties of Affect and Affect-laden Information Processing as Viewed Through the Facial Response System," in The Role of Affect in Consumer Behaviour, Robert A. Peterson, Wayne D. Hoyer and William R. Wilson, eds., MA: Lexington Books, 87-118.

- Cadotte, Ernest R., Robert B. Woodruff and Roger L. Jenkins (1987), "Expectations and Norms in Models of Consumer Satisfaction," Journal of Marketing Research, 24 (August), 305-14.
- Calder, B.J., L.W. Phillips and A.M. Tybout (1981), "Designing Research for Application," Journal of Consumer Research, Vol. 8, 197-207.
- \_\_\_\_\_, \_\_\_\_\_ and \_\_\_\_\_ (1982), "The Concept of External Validity," Journal of Consumer Research, Vol. 9, 240-4.
- \_\_\_\_\_, \_\_\_\_\_ and \_\_\_\_\_ (1983), "Beyond External Validity," Journal of Consumer Research, Vol. 10, 112-4.
- Campbell, D.T. and D.W. Fisk (1959), "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix" Psychological Bulletin, Vol. 56 (March), 100-22.
- \_\_\_\_\_ and J.C. Stanley (1963), Experimental and Quasi-Experimental Designs for Research, Boston: Houghton Mifflin.
- Carlsmith, J. Merrill and Elliott Aronson (1963), "Some Hedonic Consequences of the Confirmation and Disconfirmation of Expectations," Journal of Abnormal Social Psychology, Vol. 66 (February), 151-56.
- Carman, James M. (1990), "Consumer Perceptions of Service Quality: An Assessment of the Servqual Dimensions," Journal of Retailing, Vol. 66, No. 1 (Spring), 33-55.
- Chase, Richard B. (1978), "Where Does the Consumer Fit in the Service Operation?" Harvard Business Review, Vol. 56, 137-42.
- \_\_\_\_\_ (1981), "The Consumer Contact Approach to Services: Theoretical Bases and Practical Extensions," Operations Research, Vol. 29, No. 4, 698-706.
- Churchill, Gilbert A. Jr., (1991), Marketing Research, Methodological Foundations, Hinsdale, Illinois: The Dryden Press.
- \_\_\_\_\_ and Carol Surprenant (1982), "An Investigation into the Determinants of Customer Satisfaction," Journal of Marketing Research, Vol. 19 (November), 491-504.
- Clark, Margaret Sydnor and Susan T. Fiske (1982), Affect and Cognition: The Seventeenth Annual Carnegie Symposium on Cognition. Hillsdale, NJ: Lawrence Erlbaum Associates.
- \_\_\_\_\_ and A.M. Isen (1982), "Toward Understanding the Relationship Between Feeling States and Social Behaviour," in Cognitive Social Psychology, A. Hastorf and A. Isen, eds., NY: Elsevier North-Holland.

- \_\_\_\_\_, S. Milberg and J. Ross (1984), "Arousal Cues Arousal-Related Material in Memory: Implications for Understanding Effects of Mood on Memory," in Journal of Verbal Learning and Verbal Behavior, Vol. 23, 78-86.
- \_\_\_\_\_ and B.A. Waddell (1983), "Effects of Moods on Thoughts about Helping, Attraction and Information Acquisition," in Social Psychology Quarterly, Vol. 46, 31-5.
- Cook, T.C. and D.T. Campbell (1979), Quasi-Experimentation: Design and Analysis Issues for Field Settings. Boston: Houghton Mifflin.
- Cook, Karen S. and David M. Messick (1983), "Psychological and Sociological Perspectives on Distributive Justice: Convergent, Divergent and Parallel Lines," in Equity Theory: Psychological and Sociological Perspectives, David M. Messick and Karen S. Cook, eds. New York: Praeger Publishers, 1-12.
- Coombs, C.H. (1975), "Portfolio Theory and the Measurement of Risk," in Human Judgement and Decision Processes, M.F. Kaplan and S. Schwartz, eds., NY: Academic Press.
- \_\_\_\_\_, R. Dawes and A. Tversky (1970), Mathematical Psychology: An Elementary Introduction. Eaglewood Cliffs: Prentice Hall.
- Cronbach, Lee J. and L. Furby (1970), "How Should We Measure Change or Should We," Psychological Bulletin, Vol. 74 (July), 68-80.
- Cunningham S.M. (1967), "Major Dimensions of Perceived Risk," in Risk Taking and Information-Handling in Consumer Behaviour. D.F. Cox, ed., Boston: Harvard University Press, 82-108.
- Currim, Imran S. and Rakesh K. Sarin (1983), "A Procedure for Measuring and Estimating Consumer Preferences under Uncertainty," Journal of Marketing Research, Vol. 20 (August), 249-56.
- Czepiel John A, Michael R. Solomon and Carol F. Surprenant (1985), The Service Encounter. New York: Lexington Books.
- Darby, M.R. and E. Karni (1973), "Free Competition and the Optimal Amount of Fraud," Journal of Law and Economics, Vol. 16 (April), 67-86.
- Day, Ralph L. (1977), "Towards a Process Model of Consumer Satisfaction," in Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction. H. Keith Hunt, ed. Cambridge, MA: Marketing Science Institute, 153-83.
- \_\_\_\_\_ (1984), "Modelling Choices Among Alternative Responses to Dissatisfaction," in Advances in Consumer Research, Thomas C. Kinnear, ed., Ann Arbor, MI: Association for Consumer Research, Vol. 11, 496-9.

- Day R.H., D.J. Aigner and K.R. Smith (1971), "Safety Margins and Profit Maximization in the Theory of Firm," Journal of Political Economy, Vol. 79, 1293-301.
- DeSouza, Glenn (1989), "Now Service Business Must Manage Quality," Journal of Business Strategy, (May/June), 21-5.
- Diamond, William D. (1988), "The Effect of Probability and Consequence Levels on the Focus of Consumer Judgments in Risky Situations," in Journal of Consumer Research, Vol. 15 (September), 280-3.
- Donovan, Robert J. and John R. Rossiter (1982), "Store Atmosphere: An Environmental Psychology Approach," Journal of Retailing, Vol. 58 (Spring), 34-57.
- Eiglier, P. and E. Langeard (1977), "Services as Systems: Marketing Implications," in Marketing Consumer Services: New Insights, P. Eiglier, E. Langeard, C.H. Lovelock, J.E.G. Bateson and R.F. Young, eds., Cambridge, MA: Marketing Science Institute.
- Einhorn, H.J. and Hogarth R. M. (1985), "Ambiguity and Uncertainty in Probabilistic Inference," Psychological Review, Vol. 92, 433-61.
- \_\_\_\_\_ and \_\_\_\_\_ (1986), "Decision Making Under Ambiguity," Journal of Business, Vol. 59, No. 4, Part 2, 225-50.
- Ekman, Paul (1982), Emotion in the Human Face. Cambridge: Cambridge University Press.
- \_\_\_\_\_ and Wallace.V. Friesen (1978), Facial Action Coding System (FACS): A Technique for the Measurement of Facial Action. Palo Alto, CA: Consulting Psychologists Press.
- \_\_\_\_\_, \_\_\_\_\_ and S. Ancoli (1980), "Facial Signs of Emotional Experience," in Journal of Personality and Social Psychology, 39 (June), 1125-34.
- \_\_\_\_\_, \_\_\_\_\_ and Phoebe Ellsworth (1982), "Conceptual Ambiguities," in Emotion in the Human Face, Paul Ekman, ed., Cambridge: Cambridge University Press, 7-21.
- Festinger, L. (1957), A Theory of Cognitive Dissonance. Stanford, CA: Stanford University Press.
- Fike, R.L. (1977), "Subjective Estimation of Variability," unpublished doctoral dissertation, University of Arizona.

- \_\_\_\_\_ and William R. Ferrell (1977), "Heuristics in Variability Estimation," Proceedings to the International Conference on Cybernetics and Society, Washington.
- Fishbein M. (1967), "A Behaviour Theory Approach to the Relations Between Beliefs About an Object and the Attitude Toward the Object," in Readings in Attitude Theory and Measurement. M. Fishbein, ed., NY: Wiley, 389-99.
- \_\_\_\_\_ and I. Ajzen (1975), Belief, Attitude, Intention and Behaviour: An Introduction to Theory and Research. Reading, MA: Addison-Wesley.
- Fisher, Jeffrey, Marvin Rytting and Richard Neslin (1976), "Hands Touching Hands: Affective and Evaluative Effects on Interpersonal Touch," Sociometry, Vol. 39, No. 4, 416-21.
- Fisk, Raymond P. (1981), "Toward a Consumption/Evaluation Model for Services," in Marketing of Services, J.H. Donnelly and W.R. George, eds., Chicago: American Marketing Association.
- Fiske, Susan T. (1981), "Social Cognition and Affect," In Cognition, Social Behaviour and the Environment, J. Harvey, ed., Hillsdale, NJ: Erlbaum, 227-64.
- \_\_\_\_\_ (1982), "Schema-triggered Affect: Applications to Social Perception," in Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition, M.S. Clark and S.T. Fiske, eds., Hillsdale, NJ: Erlbaum, 55-78.
- \_\_\_\_\_ and Mark A. Pavelchak (1986), "Category-Based versus Piecemeal-Based Affective Responses, Developments in Schema-Triggered Affect," in Handbook of Motivation and Cognition: Foundation of Social Behaviour, R.M. Sorrentino and E.T. Higgins, eds., NY: Guildford Press, 168-203.
- \_\_\_\_\_ and Shelley E. Taylor (1984), Social Cognition. MA: Addison-Wesley.
- Fox, K. and S. Benda (1986), "Behavior-Setting Analysis to Improve Bank Service Delivery," in Creativity in Services Marketing: What's New, What Works, What's Developing, M. Venkatesan, D.M. Schmalensee and C. Marshall, eds., Chicago: American Marketing Association.
- Gelb, Betsy D. and Samuel V. Smith (1988), "A Model of Post-Purchase Processes Determining Satisfaction with A Professional Service Provider," in Educators Proceedings, Efficiency and Effectiveness in Marketing, 111.
- Gibson, J.J. (1979), The Ecological Approach to Visual Perception. Boston: Houghton Mifflin.

- Groenroos, C. (1982), Strategic Management and Marketing in the Service Sector. Cambridge, MA: Marketing Science Institute.
- Grove, Stephen J. and Raymond P. Fisk (1983), "The Dramaturgy of Services Exchange: An Analytical Framework for Services Marketing," in Emerging Perspectives in Services Marketing, Leonard L. Berry, G. Lynn Shostack and Gregory D. Upah, eds. Chicago: American Marketing Association.
- Guseman, D.S. (1981), "Risk Perception and Risk Reduction in Consumer Services," in Marketing of Services, J.H. Donnelly and W.R. George, eds., Chicago: American Marketing Association.
- Havlena W.J. and M.B. Holbrook (1986), "The Varieties of Consumption Experience: Comparing Two Typologies of Emotion in Consumer Behaviour," Journal of Consumer Research, Vol. 13, 394-404.
- Hawes, Douglas K. (1979), "Time and Theories of Consumer Behaviour," Faculty Research Paper No 270, University of Wyoming.
- Hayduk, Leslie A. (1987), Structural Equation Modelling with Lisrel. Baltimore: The John Hopkins University Press.
- Herr, Paul, Stephen J. Sherman and Russell H. Fazio (1982), "On the Consequences of Priming: Assimilation and Contrast Effects," Journal of Experimental Social Psychology, Vol. 13, 141-54.
- Herting, Jerald R. and Herbert L. Costner (1985), "Respecification in Multiple Indicator Models," in Causal Models in the Social Sciences, H.M. Blalock, Jr., ed., NY: Aldine Publishing Company, 321-93.
- Herzberg, F. (1975), "One More Time: How Do You Motivate Employees," Harvard Business Review on Management, 361-78.
- Heskett, J.L. (1986), Managing in the Service Economy, Boston, MA: Harvard Business School Press.
- Higgins E.T., C.P. Herman and M.P. Zanna (1981), Social Cognition: The Ontario Symposium. Vol. 1, Hillsdale, NJ: Erlbaum.
- Hoch, Stephen J. and John Deighton (1989), "Managing What Consumers Learn from Experience," Journal of Marketing, Vol. 53, (April), 1-20.
- \_\_\_\_\_ and Young-Won Ha (1986), "Consumer Learning: Advertising and the Ambiguity of Product Experience," Journal of Consumer Research, Vol. 13 (October), 221-33.

- Hoffman, Martin L. (1982), "Development of Prosocial Motivation: Empathy and Guilt," in Development of Prosocial Behaviour, N. Eisenberg, ed., NY: Academic Press, 281-313.
- \_\_\_\_\_ (1983), "Empathy, Guilt and Social Cognition," in The Relationship Between Social and Cognitive Development, W.F. Overton, ed., Hillsdale, NJ: Erlbaum, 1-52.
- \_\_\_\_\_ (1986), "Affect, Cognition, and Motivation," in Handbook of Motivation and Cognition: Foundation of Social Behaviour, R.M. Sorrentino and E.T. Higgins, eds., NY: Guildford Press, 244-80.
- Hogarth, Robin M. (1987), Judgement and Choice. 2nd Edition, NY: John Wiley & Sons.
- Holbrook, Morris B. (1984), "Situation-Specific Ideal Points and Usage of Multiple Dissimilar Brands," in Research in Marketing, Vol. 7, Jagdish N. Sheth, ed. Greenwich, CT: JAI Press, Inc., 93-112.
- \_\_\_\_\_ (1986), "Emotion in the Consumption Experience: Toward a New Model of the Human Consumer," in The Role of Affect in Consumer Behaviour, Robert A. Peterson, Wayne D. Hoyer and William R. Wilson, eds., MA: Lexington Books, 17-52.
- \_\_\_\_\_ and Elizabeth C. Hirschmann (1982), "The Experiential Aspects of Consumption: Consumer Fantasies, Feelings and Fun," Journal of Consumer Research, Vol. 9 (September), 132-40.
- \_\_\_\_\_ and John O'Soughnessy (1984), "The Role of Emotion in Advertising," Psychology and Marketing, Vol. 1, 87-98.
- Hoover, Robert J., Robert T. Green and Joel Saegert (1978), "A Cross National Study of Perceived Risk," Journal of Marketing, Vol. 42 (July), 102-8.
- Howell C. David (1987), Statistical Methods for Psychology. 2nd Edition, MA: PWS Publishers.
- Hui, Michael (1988), "Perceived Control and Crowding in the Service Encounter," Unpublished Doctoral Dissertation, London Business School.
- Isen, Alice M. (1984), "Toward Understanding the Role of Affect in Cognition," in Handbook of Social Cognition, R. Wyer and T. Srull, eds., Hillsdale, NJ: Lawrence Erlbaum Associates, 179-236.
- \_\_\_\_\_, Barbara Means, Robert Patrick and Gary Nowicki (1982), "Some Factors Influencing Decision-Making Strategy and Risk Taking," in Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition, M.S. Clark and S.T. Fiske, eds., Hillsdale, NJ: Erlbaum, 243-62.

- Jaccard, James, Robert Turrisi and Choi K. Wan (1990), Interaction Effects in Multiple Regression. London: Sage Publications.
- Izard, Carroll E. (1977), Human Emotions. NY: Plenum Press.
- Jacoby, J. and L. Kaplan (1972), "The Components of Perceived Risk," in Proceedings 3 A.C.R., 382-93.
- Joereskog, Karl G. and Dag Soerbom (1982), "Recent Developments in Structural Equation Modelling," Journal of Marketing Research, Vol. 19 (November), 404-16.
- \_\_\_\_\_ and \_\_\_\_\_ (1984) LISREL VI: Analysis of Linear Structural Relationships by Maximum Likelihood, Instrumental Variables, and Least Square Methods. Mooresville, IN: Scientific Software.
- \_\_\_\_\_ and \_\_\_\_\_ (1987), Simplis: Estimating Linear Structural Relationships the Easy Way Using Two-Stage Least-Squares. Mooresville, IN: Scientific Software.
- \_\_\_\_\_ and \_\_\_\_\_ (1989), Lisrel 7: User's Reference Guide. Mooresville, IN: Scientific Software.
- \_\_\_\_\_ and \_\_\_\_\_ (1990), "Model Search with Tetrad II and Lisrel," in Sociological Methods and Research, Vol. 19, 1 (August), 93-106.
- Kahneman D. and A. Tversky (1979), "Prospect Theory: An Analysis of Decision Under Risk," Econometrica, Vol. 47, 263-91.
- Kataoka, S. (1963), "A Stochastic Programming Model," Econometrica, Vol. 31, 191-6.
- Klaus, P.G. (1985), "Face to Face Service Encounters: The Issue of Quality in the Mass-Delivery of Services by Public Service Enterprises," Unpublished doctoral dissertation, Boston University.
- Knisely, Gary (1988), "Comparing Marketing Management in Package Goods and Service Organizations," in Managing Services Marketing, John E.G. Bateson ed., Chicago: The Dryden Press, 58-77.
- Kidder Louise H. and Charles M. Judd (1986), Research Methods in Social Relations. NY: Holt Rinehard and Winston.
- Koepp, Stephen (1987), "Pul-eeze! Will Somebody Help Me?" in Time (February 2), 28-34.

- Korgaonkar, Pradeep K. (1982), "Non-Store Retailing and Perceived Product Risk," in An Assessment of Marketing Thought and Practice - 1982 Educators' Conference Proceedings, Bruce J. Walker, ed., 204-7.
- Kotler, Philip (1989), Principles of Marketing. NY: Prentice Hall.
- Krughoff, R. (1981), "Service Evaluation," in Marketing of Services, J.H. Donnelly and W.R. George, eds., Chicago: American Marketing Association.
- Kuhl, Julius (1986), "Motivation and Information Processing: A New Look at Decision Making, Dynamic Change and Action Control," in Handbook of Motivation and Cognition: Foundations of Social Behaviour, Richard M. Sorrentino and E. Tory Higgins, eds., NY: The Guildford Press, 404-34.
- La Du, Terence J. and J.S. Tanaka (1989), "Influence of Sample Size, Estimation Method, and Model Specification on Goodness-of-Fit Assessments in Structural Equation Models," in Journal of Applied Psychology, Vol. 74, No. 4, 625-35.
- Langer E.J. and S. Saegert (1977), "Crowding and Cognitive Control," Journal of Personality and Social Psychology, Vol. 35, 175-82.
- LaTour, Stephen A. and Nancy C. Peat (1978), "The Role of Situationally-Produced Expectations, Others' Experiences, and Prior Experience in Determining Consumer Satisfaction," in Jerry C. Olson, ed., Advances in Consumer Research, Vol. 7, Ann Arbor, Mich: Association for Consumer Research, 588-92.
- \_\_\_\_\_ and \_\_\_\_\_ (1979a), "Conceptual and Methodological Issues in Consumer Satisfaction Research," in Advances in Consumer Research, Vol. 6, William L. Wilkie, ed. Ann Arbor, MI: Association for Consumer Research, 431-7.
- \_\_\_\_\_ and \_\_\_\_\_ (1979b), "Determinants of Consumer Satisfaction: A Field Study," Proceedings of the Division 23 Program, 87th Annual Convention of the American Psychology Association, Ivan Ross, ed., 83-5.
- Lazarus, Richard S. (1982), "Thoughts on the Relations Between Emotion and Cognition," American Psychologist, Vol. 37, 1019-24.
- \_\_\_\_\_ (1984), "On the Primacy of Cognition," American Psychologist, Vol. 39, 124-9.
- Leventhal, H. (1974), "Emotions: A Basic Problem for Social Psychology," in Social Psychology: Classic and Contemporary Integrations, C. Nemeth, ed., Chicago: Rand McNally.

- Levitt, Theodore (1972), "Production Line Approach to Service," Harvard Business Review, Vol. 50 (September-October), 41-52.
- \_\_\_\_\_ (1976), "The Industrialization of Service," Harvard Business Review, 54 (September-October), 63-74.
- \_\_\_\_\_ (1981), "Marketing Intangible Products and Product Intangibles," Harvard Business Review, May-June, 94-102.
- Leyens, Jacques-Phillipe and Jean-Paul Codol (1988), "Social Cognition," in Introduction of Social Psychology, Miles Hewstone, Wolfgang Stroebe, Jean-Paul Codol and Geoffrey M. Stephenson, eds., Oxford: Basil Blackwell, 89-110.
- Liechty, M. and Gilbert A. Churchill, Jr. (1979), "Conceptual Insights into Consumer Satisfaction with Services," in Educator's Conference Proceedings, Series 94, Neil Beckwith, eds., Chicago: American Marketing Association, 509-15.
- Linville, Patricia W. (1982a), "Affective Consequences of Complexity Regarding the Self and Others," in Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition, M.S. Clark and S.T. Fiske, eds., Hillsdale, NJ: Erlbaum, 79-110.
- \_\_\_\_\_ (1982b), "The Complexity-Extremity Effect and Age-based Stereotyping," Journal of Personality and Social Psychology, 42, 193-211.
- Locke, Erwin (1967), "Relationships of Success and Expectations to Affect on Goal Seeking Tasks," Journal of Personality and Social Psychology, Vol. 7, No. 2, 125-34.
- Lovelock, Christopher H. (1981), "Why Marketing Management Needs to be Different for Services," in Marketing of Services, J.H. Donnelly and W.R. George, eds., Chicago: American Marketing Association, 5-9.
- \_\_\_\_\_ and R.F. Young (1979), "Look to Consumers to Increase Productivity," in: Harvard Business Review, Vol. 57 (May-June), 168-78.
- Maddox, Raymond Neil (1976), "Measuring Consumer Satisfaction," Unpublished Doctoral Dissertation, Ohio State University.
- \_\_\_\_\_ (1981), "Two-Factor Theory and Consumer Satisfaction: Replication and Extension," Journal of Consumer Research, Vol. 8 (June), 97-102.
- Mandler, George (1975), Mind and Emotion. NY: Wiley.

- \_\_\_\_\_ (1982), "The Structure of Value: Accounting for Taste," in Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition, M.S. Clark and S.T. Fiske, eds., Hillsdale, NJ: Erlbaum, 3-36.
- Mehrabian, A. (1976), Public Spaces and Private Spaces: The Psychology of Work, Play and Living Environments. New York: Basic Books, Inc.
- \_\_\_\_\_, and J.A. Russell (1974), An Approach to Environment Psychology, Cambridge, Mass.: MIT Press
- Messick, David M. and Keith P. Sentis (1979), "Fairness and Preference," Journal of Experimental Social Psychology, Vol. 15 (July), 418-34.
- Meyer, Robert J. (1981), "A Model of Multiattribute Judgment Under Attribute Uncertainty and Informational Constraint," Journal of Marketing Research, Vol. 16 (November), 428-41.
- Miller, John A. (1977), "Studying Satisfaction, Modifying Models, Eliciting Expectations, Posing Problems and Making Meaningful Measurements," in Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction, H. Keith Hunt, ed. Cambridge, MA: Marketing Science Institute, 72-91.
- Miller, Susanne M. (1981), "Predictability and Human Stress: Toward a Clarification of Evidence and Theory," in Experimental Social Psychology, Vol. 14, 205-57.
- Milliman, Ronald E. (1982), "Using Background Music to Affect the Behaviour of Supermarket Shoppers," in Journal of Marketing, Vol. 46 (Summer), 86-91.
- \_\_\_\_\_ (1986), "The Influence of Background Music on the Behaviour of Restaurant Patrons," in Journal of Consumer Research, Vol. 13 (September), 286-9.
- Mills, P.K. and D.J. Moberg (1982), "Perspectives on the Technology of Service Operations," Academy of Management Review, Vol. 7, No. 3, 467-78.
- \_\_\_\_\_ and James H. Morris (1986), "Clients are 'Partial' Employees of Service Organizations: Role Development in Client Participation," in Academy of Management Review, Vol. 11, No. 4, 726-35.
- Moreland, Richard L. and Robert B. Zajonc (1982), "Exposure Effects in Person Perception: Familiarity, Similarity and Attraction," in Journal of Experimental Social Psychology, Vol. 18, 395-415.

- Morris, Earl W. (1976), "A Normative Deficit Approach to Consumer Satisfaction," in Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction, H. Keith Hunt, ed., Cambridge, MA: Marketing Science Institute, 240-74.
- Myers J.G., S.A. Greyser and W.F. Massy (1979), "The Effectiveness of Marketing's R&D for Marketing Management: An Assessment," Journal of Marketing, Vol. 43 (January), 17-29.
- Nelson, Phillip (1974), "Advertising and Information," Journal of Political Economy, Vol. 81, 729-54.
- Nunnally, Jum C. (1981), Psychometric Theory, NY: MacGraw-Hill.
- Oliver, Richard L. (1976), "Hedonic Reactions to the Disconfirmation of Product Performance Expectations: Some Moderating Conditions," Journal of Applied Psychology, Vol. 62 (April), 246-50.
- \_\_\_\_\_ (1980a), "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions," Journal of Marketing Research, Vol. 17 (November), 460-9.
- \_\_\_\_\_ (1980b), "Theoretical Bases of Consumer Satisfaction Research: Review, Critique, and Future Direction," in Theoretical Developments in Marketing, Charles W. Lamb, Jr. and Patrick M. Dunne, eds., Chicago: American Marketing Association, 206-10.
- \_\_\_\_\_ (1985), "An Extended Perspective on Post-Purchase Phenomena: Is Satisfaction a Red Herring?", unpublished paper presented at 1985 Annual Conference of the Association for Consumer Research, Las Vegas (October).
- \_\_\_\_\_ and William O. Bearden (1985), "Disconfirmation Processes and Consumer Evaluation in Product Usage," in Journal of Business Research, Vol. 13, 235-46.
- \_\_\_\_\_ and Wayne S. DeSarbo (1988), "Response Determinants in Satisfaction Judgements," Journal of Consumer Research, Vol. 14 (March), 495-507.
- \_\_\_\_\_ and E. John Swan (1989), "Consumer Perceptions of Interpersonal Equity and Satisfaction in Transactions: A Field Survey Approach," Journal of Marketing, (April), 21-35.
- Olshavsky, Richard W. and John A. Miller (1972), "Consumer Expectation, Product Performance and Perceived Product Quality," Journal of Marketing Research, Vol. 9 (February), 19-21.

- \_\_\_\_\_ and Philip Dover (1979), "Disconfirmation of Consumer Expectations Through Product Trial," Journal of Applied Psychology, Vol. 64 (April), 179-89.
- Olson, Jerry C. and Philip A. Dover (1979), "Disconfirmation of Consumer Expectations Through Product Trial," Journal of Applied Psychology, 64, 179-89.
- Osgood E.E. (1969), "On the Whys and Wherefores of E., P. and A.," Journal of Personality and Social Psychology, Vol. 12, 194-9.
- Parasuraman, A., Valarie A. Zeithaml and Leonard L. Berry (1985), "A Conceptual Model of Service Quality and its Implications for Future Research," Journal of Marketing, Vol. 49 (Fall), 41-50.
- \_\_\_\_\_, \_\_\_\_\_ and \_\_\_\_\_ (1988), "Servqual: A Multi-Item Scale for Measuring Consumer Perceptions of Service Quality," Journal of Retailing, Vol. 64 (Spring), 12-37.
- Peterson Robert A., D. Wayne Hoyer and William R. Wilson (1986), "Reflections on the Role of Affect in Consumer Behaviour," in The Role of Affect in Consumer Behaviour: Emerging Theories and Applications, Robert A. Peterson, D. Wayne Hoyer and William R. Wilson, eds., MA: Lexington Books, 141-60.
- Plutchik, Robert (1980), Emotion: A Psychoevolutionary Synthesis, New York: Harper & Row.
- Pfaff, Martin (1977), "The Index of Consumer Satisfaction: Measurement Problems and Opportunities," in Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction, H. Keith Hunt, ed., Cambridge, MA: Marketing Science Institute (May), 36-71.
- Prakash, Ved (1981), "An Exploratory Investigation of Post Purchase Consumer Satisfaction Within a Cognitive Process Framework," Unpublished Doctoral Dissertation, The University of Tennessee.
- \_\_\_\_\_ (1984), "Validity and Reliability of the Confirmation of Expectations Paradigm As a Determinant of Consumer Satisfaction," in Journal of the Academy of Marketing Science, Vol. 12 (Fall), 63-76.
- \_\_\_\_\_ and John W. Lounsbury (1984), "The Role of Expectations in the Determination of Consumer Satisfaction," in Journal of the Academy of Marketing Science, Vol. 12 (Summer), 1-17.

- Proshansky, H., W. Ittelson and L. Rivlin (1970), "Freedom of Choice and Behavior in a Physical Setting," in Environmental Psychology, H. Proshansky, W. Ittelson and L. Rivlin, eds., NY: Holt, Rinehart and Winston.
- Rich, T. (1977), Personal Communication, NY: Plenum Press.
- Riddle, Dorothy I. (1986), Service-Led Growth, Westpoint, Ct: Praeger.
- Roselius, T. (1971), "Consumer Rankings of Risk Reduction Methods," Journal of Marketing, Vol. 35, 56-61.
- Roseman, I. (1979), "Cognitive Aspects of Emotion and Emotional Behaviour," Paper Presented at the Meeting of the American Psychology Association, NY: September.
- Rumelhart D.E. and A. Ortony (1977), "The Representation of Knowledge in Memory," in Schooling and the Acquisition of Knowledge, R.C. Anderson, R.J. Spiro and W.E. Montague, eds., Hillsdale, NJ: Erlbaum, 99-135.
- Russell, James A. (1978), "Evidence of Convergent Validity on the Dimensions of Affect," Journal of Personality and Social Psychology, Vol. 36, 1152-68.
- \_\_\_\_\_ (1979), "Affective Space is Bipolar," in Journal of Personality and Social Psychology, Vol. 37, No. 3, 345-56.
- \_\_\_\_\_ (1980), "A Circumplex Model of Affect," Journal of Personality and Social Psychology, Vol. 39, No. 6, 1161-78.
- \_\_\_\_\_ and G. A. Pratt (1980), "A Description of the Affective Quality Attributed to Environments," Journal of Personality and Social Psychology, Vol. 38 (August), 311-22.
- \_\_\_\_\_ and Doreen Ridgeway (1983), "Dimensions Underlying Children's Emotion Concepts," in Developmental Psychology, Vol. 19, 795-804.
- \_\_\_\_\_, L.M. Ward and G. Pratt (1981), "Affective Quality Attributed to Environments - A Factor Analysis Study," Environment and Behavior, Vol. 13, 259-88.
- Sarel, Dan (1978), "An Investigation of the Effects of Risky Attributes on Consumer Evaluation Strategies," Unpublished doctoral dissertation, Harvard University.
- Sawyer, Alan G. (1977), "The Role of Role-Playing in Experiments About Consumer Behaviour," in Contemporary Marketing Thought, Barnett A. Greenberg and Danny N. Bellenger, eds., American Marketing Association, 191-4.

- Schachter S. and J. Singer (1962), "Cognitive, Social and Physiological Determinants of Emotional State," Psychological Review, Vol. 69, 379-99.
- Schiffman, Leon G. and Leslie Lazar Kanuk (1987), Consumer Behavior. NJ: Prentice Hall.
- Schneider, Benjamin (1980), "The Service Organization: Climate is Crucial," in Organizational Dynamics, Autumn, 52-65.
- Schutz, W.C. (1966), The Interpersonal Underworld. Palo Alto, Calif: Science and Behavior Books.
- Schwartz, N., F. Strack, D. Kommer and D. Wagner (1987), "Soccer, Rooms, and the Quality of Your Life: Further Evidence on Informative Functions of Affective States," European Journal of Social Psychology, Vol. 17, 69-80.
- Shamir, Boas (1980), "Between Service and Servility: Role Conflict in Subordinate Service Roles," in Human Relations, Vol. 33, No. 10, 741-56.
- Shanteau, J. (1974), "An Information Integration Analysis of Risky Decision Making," in Human Judgement and Decision Processes, M.F. Kaplan and S. Schwartz, eds., NY: Academic Press, 110-34.
- Sherif, Muzafer and Carl J. Hovland (1961), Social Judgments: Assimilation and Contrast Effects in Communication and Attitude Change. New Haven, Connecticut: Yale University Press.
- Shocker, A.D. and V. Srinivasan (1979), "Multiattribute Approaches to Product Concept Evaluation and Generation: A Critical Review," Journal of Marketing Review, Vol. 16 (May), 159-80.
- Shostack, G. Lynn (1984), "Service Design in the Operating Environment," in Developing New Services, William R. George and Claudia Marshall, eds., Chicago: American Marketing Association, 27-43.
- \_\_\_\_\_ (1985), "Planning the Service Encounter," in The Service Encounter, John A. Czepiel, Michael R. Solomon and Carol F. Surprenant, eds. Lexington, MA: Lexington Books, 243-54.
- \_\_\_\_\_ (1987), "Breaking Free from Product Marketing," Journal of Marketing, Vol. 41, 73-80.
- Simon, Herbert A. (1982), "Affect and Cognition: Comments," in Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition, Margaret S. Clark and Susan T. Fiske, eds., Hillsdale, NJ: Erlbaum, 333-42.

- Sirgy, M. Joseph (1984), "A Social Cognition Model of Consumer Satisfaction/Dissatisfaction: An Experiment," Psychology and Marketing, 1 (Summer), 27-44.
- Slocum, J.W. and H.P. Sims (1980), "A Typology for Integrating Technology, Organization, and Job Design," Human Relations, Vol. 33, 193-212.
- Smith, Craig A. and Richard S. Lazarus (1990), "Emotion and Adaptation," in Handbook of Personality: Theory and Research, L. A. Pervin, ed., New York: Guildford, 609-37.
- \_\_\_\_\_ and Phoebe C. Ellsworth (1985), "Patterns of Cognitive Appraisal in Emotion," Journal of Personality and Social Psychology, Vol. 48, 813-38.
- \_\_\_\_\_, Jill Novacek, Richard S. Lazarus and Lois K. Pope (1990), "Antecedents of Emotion: Situations, Dispositions, Attributions and Appraisals," Paper Presented at the 98th Annual Convention of the American Psychological Association, Boston.
- Smith, Ruth A. and Michael J. Houston (1983), "Script-Based Evaluations of Satisfaction with Services," in Emerging Perspectives in Services Marketing, Leonard L. Berry, G. Lynn Shostack and Gregory D. Upah, eds. Chicago, American Marketing Association, 59-62.
- Solomon, Michael R., Carol F. Surprenant, John A. Czepiel and Evelyn G. Gutman (1985), "A Role Theory Perspective on Dyadic Interactions: The Service Encounter," Journal of Marketing, Vol. 49 (Winter), 99-111.
- Sorrentino, Richard M, and E. Tory Higgins (1986), Handbook of Motivation and Cognition: Foundations of Social Behaviour. NY: The Guildford Press.
- Surprenant, Carol F. and Gilbert A. Churchill (1984), "Can Role Playing be Substituted for Actual Consumption?" in Advances in Consumer Research, Vol. 11, 122-6.
- \_\_\_\_\_ and Michael R. Solomon (1987), "Predictability and Personalization in the Service Encounter," Journal of Marketing, Vol. 51 (April), 86-96.
- Swan, John E. (1983), "Consumer Satisfaction Research and Theory: Current Status and Future Directions," in International Fare in Consumer Satisfaction and Complaining Behavior, Ralph L. Day and H. Keith Hunt, eds., Bloomington, Indiana School of Business, Indiana University, 124-9.
- \_\_\_\_\_ and Alice Atkins Mercer (1982), "Consumer Satisfaction as a Function of Equity and Disconfirmation," in Conceptual and Empirical Contributions to Consumer Satisfaction and Complaining Behaviour, eds., H. Keith Hunt and Ralph L. Day, Bloomington, IN: Indiana University School of Business, Indiana University, 2-8.

- \_\_\_\_\_ and Richard L. Oliver (1985), "Automobile Buyer Satisfaction With the Sales Person Related to Equity and Disconfirmation," in Consumer Satisfaction, Dissatisfaction and Complaining Behavior, H. Keith Hunt and Ralph L. Day, eds., Bloomington, IN: Indiana University School of Business, Indiana University, 10-6.
- \_\_\_\_\_ and Frederick I. Trawick 1980, "Satisfaction Related to Predictive vs. Desired Expectations," in H. Keith Hunt and Ralph L. Day, eds., Refining Concepts and Measures of Consumer Satisfaction and Complaining Behavior, Bloomington: School of Business, Indiana University, 7-12.
- Taylor, S.E. and J. Crocker (1981), "Schematic Bases of Social Information Processing," in Social Cognition: The Ontario Symposium, E.T. Higgins, C.P. Herman and M.P. Zanna, eds., Hillsdale, NJ: Erlbaum, 249-88.
- Tesser, A. (1978), "A Self-Generated Attitude Change," in Advances in Experimental Social Psychology, L Berkowitz, ed., NY: Academic Press.
- Thomas, Michael J. (1988), "Getting Closer to the Customer," Marketing Intelligence and Planning, Vol. 6, No. 1, 28-31.
- Tomkins, Sylvan S. (1980), "Affect as Amplification: Some Modifications in Theory," in Emotion: Theory, Research and Experience, R. Plutchik and H. Kellerman, eds., NY: Academic Press.
- Tse David K. and Peter C. Wilton (1988), "Models of Consumer Satisfaction Formation: An Extension," Journal of Marketing Research, Vol. 25 (January), 204-12.
- Tversky A. and D. Kahneman (1981), "The Framing of Decisions and the Psychology of Choice," Science, Vol. 211, 453-8.
- US Department of Defense (1981) (ed), "Human Engineering Design Criteria for Military Systems, Equipment and Facilities," MIL-STD-1472C, 2 May 1981.
- Weiner, Bernhard (1980), Human Motivation, New York: Holt, Rinehart and Winston.
- \_\_\_\_\_ (1980a), "A Cognitive (Attribution)-Emotion-Action Model of Motivated Behaviour: An Analysis of Judgement of Help-Giving," Journal of Personality and Social Psychology, Vol. 39, 186-200.
- \_\_\_\_\_ (1982), "The Emotional Consequences of Causal Attributions," in Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition, Margaret S. Clark and Susan T. Fiske, eds., Hillsdale, NJ: Erlbaum, 185-209.

- \_\_\_\_\_ (1985), "Spontaneous Causal Thinking," Psychological Bulletin, Vol. 97, 74-85.
- Westbrook, Robert A. (1982), "Consumer Satisfaction and the Phenomenology of Emotions During Automobile Ownership Experiences," in Proceedings of the Seventh Annual Conference on Consumer Satisfaction, Dissatisfaction and Complaining Behavior, H. Keith Hunt and Ralph L. Day, eds., Bloomington: School of Business, Indiana University.
- \_\_\_\_\_ (1987), "Product/Consumption-Based Affective Responses and Postpurchase Processes," Journal of Marketing Research, Vol. 24 (August), 258-70.
- \_\_\_\_\_ and Michael D. Reilly (1983), "Value-Precept Disparity: An Alternative to the Disconfirmation of Expectations Theory of Consumer Satisfaction," in Advances in Consumer Research, Vol. 10, Richard P. Bagozzi and Adams M. Tybout, eds., Ann Arbor, MI: Association of Consumer Research.
- \_\_\_\_\_ and Richard L. Oliver (1981), "Developing Better Measures of Consumer Satisfaction: Some Preliminary Results," Advances in Consumer Research, Vol. 8, 94-9.
- \_\_\_\_\_ and \_\_\_\_\_ (1991), "The Dimensionality of Consumption Emotion Patterns and Consumer Satisfaction," Journal of Consumer Research, forthcoming.
- White, R.W. (1959), "Motivation Reconsidered: The Concept of Competence," Psychology Review, Vol. 66, 297-323.
- Wilkie, William L. and Edgar A. Pessemier (1973), "Issues in Marketing's Use of Multiattribute Attitude Models," Journal of Marketing Research, Vol. 10 (November), 428-41.
- Wilkinson, Max (1990), "Now for the Watching Bank," in Financial Times, March 3.
- Wilton, Peter C. and M. Nicosia (1986), "Emerging Paradigms for the Study of Consumer Satisfaction," European Research, Vol. 14 (January), 4-11.
- Wirtz, Jochen and John E.G. Bateson (1990), "The Concept of Affect Applied to Satisfaction Theory with Services," Paper Presented at the TIMS College on Marketing Special Interest Conference on Services Marketing, Nashville.
- Woodruff, Robert B., Ernest R. Cadotte and Roger L. Jenkins (1981), "Understanding Consumer Expectation and Satisfaction Processes Using Experience-Based Norms: A Modelling Approach," Paper presented at the Sixth Annual Conference on Consumer Satisfaction, Dissatisfaction and Complaining Behavior.

- \_\_\_\_\_, \_\_\_\_\_ and \_\_\_\_\_ (1983), "Modelling Consumer Satisfaction Processes Using Experience-Based Norms," Journal of Marketing Research, Vol. 20 (August), 296-304.
- Yalch, Richard F. and Eric Spangenberg (1988), "An Environmental and Psychological Study of Foreground and Background Music as Retail Atmospheric Factors," Educators Proceedings, Efficiency and Effectiveness in Marketing, 106-10.
- Zajonc, Robert B. (1980), "Feeling and Thinking: Preferences Need No Inferences," American Psychologist, Vol. 35, 151-75.
- \_\_\_\_\_ and Hazel Markus (1982), "Affective and Cognitive Factors in Preferences," Journal of Consumer Research, Vol. 9 (September), 123-31.
- Zaltman, G., C.R.A. Pinson and R. Angelmar (1973), Metatheory and Consumer Research. Hinsdale, IL: Dryden Press.
- Zeithaml, Vallery A. (1981), "How Consumer Evaluation Processes Differ between Goods and Services," in Marketing of Services, J.H. Donnelly and W.R. George, eds., Chicago: American Marketing Association.
- \_\_\_\_\_, A. Parasuraman and Leonard L. Berry (1985), "Problems and Strategies in Services Marketing," in Journal of Marketing, Vol. 49 (Spring), 33-46