2008

The Influence of the Servicescape on Post-Consumption Processes

Daire Hooper
Technological University Dublin, daire.hooper@tudublin.ie

The Influence of the Servicescape on Post-Consumption Processes: A Structural Equation Modelling Approach

Daire Hooper, B.Sc.

November 2008

This thesis is submitted to the Dublin Institute of Technology for the degree of Ph.D. in the Faculty of Business, School of Marketing.

Supervisors: Dr. Joseph Patrick Coughlan
Dr. Michael R. Mullen

Abstract

The physical and atmospheric cues in service environments have long been accepted as potent contributors to consumers' overall evaluations of their service experiences. Theoretical frameworks conceptualising how these environmental cues impact on both emotional and cognitive processes have been put forward, yet the empirical work supporting these theories has become highly disjointed with a great deal of confusion regarding what should be classified as an environmental stimulus. By integrating the extant literature, this dissertation attempts to overcome theoretical ambiguities and proposes a second order factor model of service environments, also known as the servicescape, which is subsequently tested using a multi-item servicescape scale that is developed herein. The service chosen to test the applicability of this model was a low involvement, utilitarian service. The rationale for this choice is that there is a dearth of research in services which are frequently used by consumers, with a greater emphasis placed on services which are more experiential in nature. The analysis demonstrates that the servicescape adheres to a multi-dimensional structure, wherein, each of the sub-constructs are related to one another through a higher level of abstraction.

This dissertation also highlights the theoretical confusion that exists between the service quality and servicescapes literature. Rather than subsume the servicescape as an additional dimension within the service quality construct, it is proposed that the servicescape is theoretically distinct whereby it implicitly communicates to consumers the level of service quality they should expect to receive. This view regards the servicescape as an antecedent to service quality with service quality depicted as a parsimonious, two dimensional structure. This relationship is tested using structural equation modelling and it is shown how various aspects of the servicescape impact upon service quality.

Emotional reactions are widely regarded as direct consequences of servicescape evaluations; similarly, consumption emotions are thought to be synonymous with customer satisfaction. This premise is tested through a comprehensive model which tests whether consumers adhere to a more emotive processing system, or to one where cognitive processes dominate. In addition, the model also elucidates how modelling the servicescape as a multi-dimensional structure allows one to determine which aspects of the servicescape are more salient than others. Finally, the dissertation concludes with a discussion on the implications of the findings as well as providing recommendations for further research in this area.

Declaration

I certify that this thesis which I now submit for examination for the award of Doctor of Philosophy, is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

This thesis was prepared according to the regulations for postgraduate study by research of the Dublin Institute of Technology and has not been submitted in whole or in part for an award in any other Institute or University.

The work reported on in this thesis conforms to the principles and requirements of the Institute's guidelines for ethics in research.

Signature _____ Date: _____
Candidate

Material published from this work:

Journal Article(s)
Hooper, D., Coughlan, J. and Mullen, M.R. (2008), 'Structural Equation Modelling: Guidelines for Determining Model Fit', *Electronic Journal of Business Research Methods*, Vol. 6, No. 1, 53-60.

Conference Papers
Hooper, D., Coughlan, J. and Mullen, M.R. (2008), 'Assessing the Dimensionality of the Servicescape: Integrating the Extant Literature', presented at the *2008 American Marketing Association SERVSIG Services Research Conference*, Liverpool, United Kingdom.

Hooper, D., Coughlan, J. and Mullen, M.R. (2008), 'Evaluating Model Fit: A Synthesis of the Structural Equation Modelling Literature' presented at the *7th European Conference on Research Methodology for Business and Management Studies*, Regent's College, London, United Kingdom.

Acknowledgements

Writing this dissertation has been no easy task and I owe a great deal to those who have supported, inspired and helped me along the way. Studying at the DIT has personally and professionally enriched my life and I am hugely thankful to all the staff and colleagues who have provided me with so many opportunities and to those who have made this a unique and memorable experience.

I am especially grateful to my supervisor Dr. Joseph Coughlan for the help and many suggestions he has given that have been instrumental to the development of this dissertation. Without Joe's ideas, continued feedback and support, this thesis would never have gotten off the ground, let alone finished. Now that this stage of my research career is completed, hopefully we can continue to work together as a team long into the future. I am also indebted to my second supervisor, Dr. Michael Mullen, who not only devoted an enormous amount of his time to teach me the intricacies of structural equation modelling, but whose hospitality was second to none during my visit to Fort Lauderdale last October. As well as all of this, Mike has been hugely encouraging and I am proud to say that I had him as my second supervisor.

I would also like to thank the many staff of the DIT Faculty of Business who have offered advice, encouragement and who have had faith in my abilities. In particular, Dr. Katrina Lawlor, Kate Uí Ghallachoir and Paul O'Sullivan, Eoghan O'Grady, Conor Horan, Laura Cuddihy, Roger Sherlock, Mary Faulkner, Tom Fennell, Dr. Margaret-Ann Lawlor, Paul O'Reilly and Patrick Doran. Other staff within the DIT who are worthy of a mention are the IT technicians Wes and Tom who have gotten me out of many a sticky situation. This research would not have materialised if I had not been given access to the three sites under study in this dissertation. Thank you to Shane Flynn for allowing me to conduct the interviews and for rallying the staff to embrace the project. I also appreciate the undergraduates who cheerfully assisted me at the arduous data collection stage.

I also want to thank my friends Fiona, Kevina, Norah and Sara (in alphabetical order so they don't fight!) who have shared this experience with me. Without these four girls, writing the magnum opus would have been nearly impossible. We have shared laughs, wise-words, sympathy, understanding and most of all memories that I will have forever. My friends, Clare, Emma, Sabrina and Tara (among others!) have been a constant source of friendship and I thank each one of them for talking to me about non-PhD things.

My family have also provided me with extensive support for which I am immensely appreciative. My parents, Paddy and Catherine, have instilled in me a love of learning and have continually encouraged me to pursue this goal. I am extremely lucky to have such wonderful parents. I would also like to give a special thank you to my three beautiful sisters, Saibh, Ailbhe and Aoife and to my little nephew Fionn who came along in the middle of this process.

Finally I would like to thank my boyfriend James who has been my rock and best-friend throughout all of this. The love, kindness, patience and humour that he has given me have lightened the load of this mammoth task. The success and completion of this dissertation is as much yours as it is mine.

Quid quid latine dictum sit, altum videtur
UNKNOWN

Table of Contents

List of Figures

List of Tables

INTRODUCTION

The significance of both the service and retail sector to the Irish economy cannot be denied as the service industry alone accounts for approximately 49% of our GDP. Indeed, the retail sector accounts for a significant proportion of this having an estimated market value of €31,785 million and employing approximately 230,000 people (Datamonitor, 2007). As Ireland and the rest of the world face an economic downturn, the service sector, and in particular, retailers are faced with an increasingly competitive marketplace. Due to this, maintaining a competitive advantage and understanding how this can be achieved has become a top priority for both service providers and retailers.

Marketing and selling services in itself is a challenge, as services are inherently different to products in that they possess distinct characteristics which requires a different approach to when one engages in traditional product marketing (Zeithaml et al, 1985). With the advent of a service dominant logic for marketing (Vargo and Lusch, 2004) the paradigmatic gap between these two areas has reduced. Despite this, the four characteristics of services, namely, intangibility, inseparability, heterogeneity and perishability remain both central and pertinent to our understanding of services marketing. Of these four service characteristics, it has been argued that intangibility is the single most important and fundamental difference between products and services (Zeithaml et al, 1985; Flipo, 1988; McDougall and Snetsinger, 1990; Bebko, 2000; Zeithaml and Bitner, 2003), with this characteristic denoting that services are often physical performances that cannot be seen tasted or touched (Zeithaml et al, 1985;

Laroche et al, 2003; Edvardsson et al, 2005). Although few service products are wholly intangible, there exists varying levels of intangibility which Shostack (1977) describes as the tangibility 'spectrum'. Recommendations on how to overcome intangibility generally encourage service providers to 'tangiblise' the service through tangible cues within the service environment or to develop a strong corporate brand (Day, 1992; Grace and O'Cass, 2004; Grace and O'Cass, 2005). The focus of the current research delves into the former of these two strategies whereby it endeavours to explicate the intricacies of how the tangible cues in service environments interact with both non-tangible cues and how they subsequently impact on overall evaluative measures.

The tangible cues associated with services encompass a broad range of physical and atmospheric factors that are collectively known as 'the servicescape' (Bitner, 1992). Comprehensive frameworks and typologies of the servicescape and store environments are well recognised in the literature, yet empirical research in this area has become disjointed and haphazard leading one confused as to where research in this area is going. This dissertation attempts to resolve some of these issues by asking questions which will clarify our understanding of the role the servicescape plays in how we experience and evaluate services. Therefore, this dissertation asks questions such as: How are the interior of services perceived by consumers? Do consumers make distinctions between various servicescape elements or do they view the servicescape at a more general level? Do consumers regard the servicescape as part of their service quality evaluations or is it a unique construct in itself? If consumers view the servicescape as a separate construct, does it precede service quality evaluations and in doing so contribute towards consumers

perceptions of service quality? How do perceptions of the servicescape impact on customer satisfaction evaluations as distinct from how consumers view their perceptions of service quality? In a low-involvement utilitarian service, which has the greatest effect on whether an individual will return to a service; the servicescape, perceived level of service quality or customer satisfaction with the service? Does the servicescape elicit emotional responses? And are these emotional responses linked to customer satisfaction?

The answers to these questions are achieved through a large-scale field study of a retail service environment. The retail service chosen for this study were service stations and the rationale for this choice was that thus far, the research on servicescapes has been overly dependent on highly experiential services that consumers only rarely use or frequent. Given that retail services comprise the largest services sector it is therefore essential that we elucidate how the tangible cues of the physical environment influence consumer behaviour. Not only do service stations satisfy the requirement for needing more retail studies, they can also be classified as utilitarian in nature in that consumers' level of involvement and time spent in the service in generally very low. Intuitively one would expect that tangible and atmospheric cues would perhaps dominate the service experience in extended and experiential service transactions, however, whether this would translate to highly standardised, everyday services is not known.

As evidenced by both Bitner (1992) and Baker's (1987) frameworks, multi-faceted conceptualisations of the servicescape and store environment are generally accepted. Yet, despite this, numerous studies have chosen to measure the servicescape as a one

dimensional construct which obfuscates how different aspects of the servicescape influence post-consumption processes such as service quality perceptions, customer satisfaction and behavioural intentions. The limited few who have made attempts at measuring the servicescape in terms of a number of distinct dimensions have unfortunately chosen servicescape factors that are very specific to the service under study (Wakefield and Blodgett, 1996). This approach adds little in the way of clarifying our understanding of the servicescape as our ability to generalise and compare the findings across contexts is severely limited. Through a review of the extant servicescape literature, this dissertation integrated the various different theoretical and empirical approaches which culminated in a more comprehensive view of the servicescape. While models presented by Bitner (1992) and Baker (1987) both propose that correlations exist between different elements in the servicescape, up until this point no empirical research has suggested the possible existence of a second-order relationship among the different servicescape elements. Second-order factor models are a type of structural equation model which allowed the researcher to specify a theoretical structure and test whether it can be upheld as a valid theoretical model (Rindskopf and Rose, 1988).

To meet this end, a measurement instrument containing multiple items was required so that the second order model could be operationalised. A further thorough review of the servicescape scales to date revealed that each of those available fell short in some way. Due to the absence of a valid and reliable measure of the servicescape it became necessary to develop a measurement instrument that could be used in the current

research. A scale development process resulted in a scale which was subsequently used in all further analysis.

This dissertation also seeks to illuminate our understanding of service quality and proposes that the servicescape serves as a unique construct in itself. Heretofore, service quality models have consistently subsumed tangible cues as part of the service quality domain. Yet, in the servicescape literature very little attention has been given to how it relates to service quality, with almost all studies overlooking how the physical environment fits into our understanding of service quality. Indeed, it seems that the service quality and the servicescape literature have evolved entirely independently of one another with little consideration given to how these constructs may be aligned. Therefore, it is proposed herein that the servicescape does not constitute an additional service quality dimension, but as a construct in itself which helps consumers form beliefs about service quality. As the servicescape is an extrinsic cue which consumers use to infer quality it is logical that servicescape evaluations precede perceptions of service quality. For instance, if we consider a person on a fairly standard shopping trip, before the individual even enters a service environment they are beginning to infer service quality based on how the exterior of the store is presented and designed. Then, once in the store their perceptions are further compounded by the surrogate service quality indicator that is the servicescape.

As well as attempting to resolve the theoretical ambiguities that exist between the servicescape and service quality, this dissertation also tests the applicability of Mehrabian

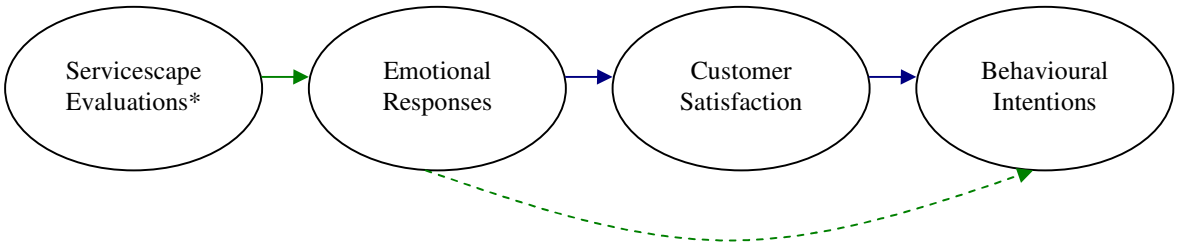
and Russell's (1974) theory of environmental psychology. Unlike the service quality literature, Mehrabian and Russell's (1974) model is grounded in a much more emotive processing system. This model proposes that the design of an environment has a significant effect on individuals which causes emotional reactions to the environment which subsequently manifests in responses in the form of approach and avoidance behaviours. This theory has been fundamental to servicescape theory with both atmospherics studies and Bitner's (1992) servicescape model based on its premise. The research documented within this dissertation adds to the Mehrabian and Russell (1974) theory by moving beyond the author's original description of the environment and asserts that the service environment can be viewed in terms of discrete sub-constructs. The reasoning behind this approach is that if we can meaningfully categorise the servicescape into multiple sub-dimensions it will improve our understanding of how difference aspects of the servicescape impact on overall service evaluations.

Through the emotions scale used in this dissertation, further contributions are made. Mehrabian and Russell's (1974) model claims that emotions elicited as a result of environments adhere to a three-dimensional structure of Pleasure-Arousal-Dominance (PAD). While multiple studies have employed the PAD measure and various emotion derivatives of it, here it is argued that this view of emotions offers little in the way of managerial implications. For instance, how can a manger or service employee differentiate between pleasure, arousal and dominance? In knowing that service environments have the ability to elicit pleasure, arousal and or dominance, how can this be maximised? To overcome the vagueness of the PAD measure, the current study

employs Watson et al's (1988) PANAS. This scale is based upon the Pleasure and Arousal circumplex proposed by Russell (1981) and so theoretically they are akin to one another. However, PANAS is advantageous in that it contains individual emotions which are broadly classified as being either positive or negative in valence, unlike the rather ambiguous Pleasure and Arousal measure put forward by Russell (Mehrabian and Russell, 1974; Russell, 1980).

Customer satisfaction is widely regarded as a 'strategic imperative' for firms (Mittal and Kamakura, 2001, p. 131) as it is thought to lead to greater profitability and consumer loyalty (Oliver, 1997). Understanding how the servicescape impacts upon customer satisfaction is therefore an area which requires further investigation. As mentioned above, emotions are generally believed to be a direct outcome of environmental evaluations, but on the other hand, emotions are also believed to be a direct antecedent of satisfaction with it widely accepted in the literature that satisfaction contains both cognitive and affective elements. Interestingly, no research to date has sought to integrate both the servicescape and customer satisfaction literature whereby emotion is hypothesised to moderate the servicescape and customer satisfaction. Unlike Mehrabian and Russell's (1974) model, it is expected that emotions will not have a direct effect on behavioural intentions but will be moderated by customer satisfaction. The relationships between each of these constructs and their theoretical underpinnings are described below with the green arrows representing Mehrabian and Russell's (1974) model and the blue representing the stance of customer satisfaction theorists (Westbrook, 1987; Westbrook

and Oliver, 1991). As it is not expected that a direct path between emotions and behavioural intentions exists, this relationship is shown with a dashed arrow.



*The servicescape has been depicted here as a single construct for ease of discussion

Figure 1: Integrating the Servicescape and Customer Satisfaction Literatures

The Organisation of the Thesis

The dissertation begins by introducing the servicescape as this is the focal construct of the thesis. Within this chapter, Mehrabian and Russell's (1974) theory of environmental psychology is presented and how various researchers have applied it to consumer behaviour situations is analysed. The chapter then turns to what is known as the atmospherics literature, a stream of research that has generally focussed on one isolated variable at a time in retail contexts. Integrative theories which build upon the single stimulus studies by conceptualising service environments as multi-dimensional structures are then discussed. The extant servicescapes literature is then reviewed and through this it is highlighted that very few multi-dimensional studies have been conducted. To provide further evidence for this, a summary of the various dimensions which are frequently cited in the literature is provided. From this it can be seen that quite a

disordered body of knowledge exists with each study taking a different approach to the previous one. Cohesion and clarity is obviously needed and to meet this objective a second-order factor structure of the servicescape is proposed.

Chapter Two continues the discussion on the servicescape however, the emphasis here is wholly methodological. The chapter begins by discussing the various servicescape scales that have been used in the past and it is from this that it becomes evident that a thorough servicescape scale is needed which contains multiple items so that each of the servicescape sub-dimensions can be adequately captured. The remainder of this chapter recounts each of the stages involved in the scale development process. In Chapter Three the dissertation temporarily moves away from the servicescape so that structural equation modelling can be introduced. From this point on in the thesis, structural equation modelling is the method of choice, with each analytical stage drawing on structural equation modelling in some way, therefore providing an overview of this technique was required.

Chapter Four reverts back to the servicescape and it is in this chapter that the model presented in Chapter One is tested using structural equation modelling. This chapter can broadly be broken into two sections: the first half of this chapter reiterates the proposed second-order factor structure as well as providing a rationale for the choice of service. This section also describes the methodological decisions that were made in relation to the format and structure of the questionnaire. The second half of the chapter focuses solely on testing the second-order factor structure which was earlier hypothesised. This model

is also compared with prominent theories of the store environment and the theoretical implications are discussed.

In Chapter Five our attention is then turned to service quality and current thinking on this construct is reviewed. The chapter also highlights how a grey area between the service quality and servicescape literatures exists and from this an alternative perspective on how these constructs relate to one another is presented. This is subsequently resolved using structural equation modelling.

Further literature is introduced and described in Chapter Six. The chapter commences with a discussion on the theoretical viewpoints which underpin the various emotion theories put forward by cognitive psychologists. Consumer behaviour studies in which the servicescape has been shown to elicit emotional responses are also discussed. Not only are emotions believed to be strongly linked to how we perceive environments, but they have also been shown to be associated with customer satisfaction evaluations. The chapter then presents an overview of this literature along with the empirical research in this area. The third construct introduced in this chapter are behavioural intentions and as with all the constructs previously discussed, the literature and theoretical assumptions relevant to behavioural intentions are presented. The final section of this chapter presents a theoretical research model which assimilates all literature covered in the dissertation thereby forming a comprehensive model which captures the inter-relationships between the servicescape, service quality, emotions, customer satisfaction and behavioural intentions.

This model is fully tested in Chapter Seven which is wholly devoted to testing the model's hypothesised paths and examines in detail the implications of the findings. The final chapter, Chapter Eight concludes the thesis by summarising the contributions made herein. This chapter also underscores the limitations of the study and from these limitations further avenues for research are recommended. Finally, the dissertation considers the managerial implications of the study and suggests how the findings can be beneficial to practitioners.

A graphical depiction of each of these chapters is provided in Figure 2 overleaf which further clarifies the linkages between the chapters and how they logically flow from one to the next.

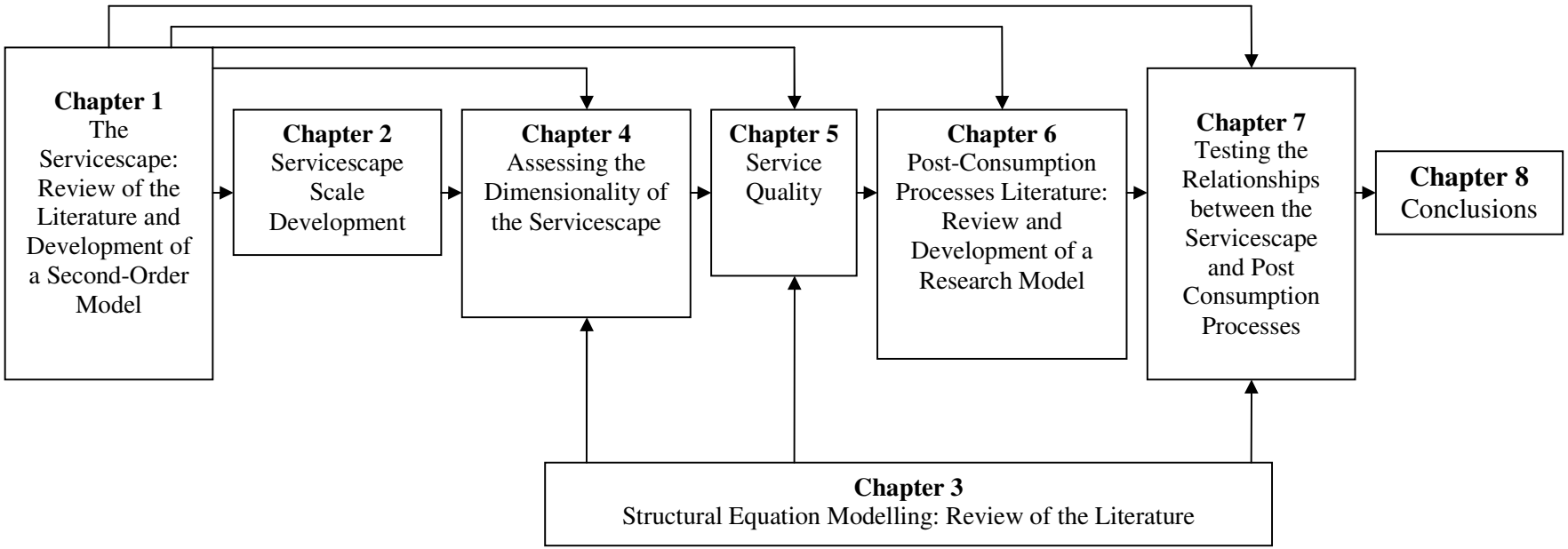


Figure 2: Structure of the Thesis

1.0 SERVICESCAPES

1.1 Introduction

A primary objective of this dissertation is to extend the empirical evidence on the structural interrelationships between the servicescape and postconsumption constructs such as, emotion, service evaluation, customer satisfaction and behavioural intentions. However, before the correspondence between these constructs can be examined, there is a need to more fully understand the servicescape and its theoretical underpinnings. In order to do so, the chapter explores the various streams of literature within this area. The chapter begins by introducing Mehrabian and Russell's (1974) theory of environmental psychology as it is this theory that has laid the foundation for much of the research on servicescapes. Although Mehrabian and Russell's (1974) model forms the basis of the work in this area, a number of research streams seem to have evolved from this theory. The chapter examines each of these bodies of knowledge in turn. From this review it becomes evident that there is little agreement as to what the servicescape comprises. Within this chapter, an attempt is made to bring together these theories from which a servicescape conceptualisation is presented. This model brings together the various different research streams to date and is conceptualised as a six factor second order structural equation model.

This chapter is structured as follows. It begins by introducing the various terms that are used to describe this important concept. Following this, Mehrabian and Russell's (1974) seminal theory is presented and consumer behaviour studies that have directly employed

this theory are reviewed. In Section 1.4, studies which have isolated specific environmental stimuli and determined their impact on behavioural responses are presented. The chapter then explores the theoretical contributions made by authors within the services marketing domain. Within this section, the conflicting conceptualisations are presented and it is from this that a proposed servicescape framework is devised.

1.2 Four Terms: One Over-Arching Theory

Upon reviewing the literature, it was noted that four terms are commonly used to describe the same concept, these being; environmental psychology, atmospherics, store environment and servicescapes. The literature surrounding each of these is derived from Mehrabian and Russell's (1974) work on environmental psychology however, upon further inspection it became evident that this quadripartite terminology appears to be attributed to the literary body from which they are derived.

The first term popularly used is 'Environmental Psychology', and those that coined this expression are not surprisingly, from a psychology background (Mehrabian and Russell, 1974; Russell and Mehrabian, 1976). Mehrabian and Russell's (1974) theory lies at the core of this work with researchers directly applying their theory albeit in a consumer context. Mehrabian and Russell's (1974) influential theory is discussed foremost along with those studies that have made a direct application of it in a consumer context (Donovan and Rossiter, 1982). The second term used in the literature, 'atmospherics', can be traced back to Kotler's (1973-1974, p. 50) seminal article in which he argued that

atmospherics is 'the conscious designing of space to create certain effects in buyers'. He contended that environmental elements within the service environment such as noises, shapes, scents, music and colour can be modified so that particular emotional effects will be elicited which in turn enhance purchase probability. This conceptualisation is similar to Mehrabian and Russell's (1974) in that they both infer that environmental stimuli can be manipulated so that certain behaviours manifest. Although some see Kotler (1973-1974) as the first to advocate the manipulation of environmental cues in order to effect consumer behaviour (Bitner, 1992; Yalch and Spangenberg, 2000), a number of studies previous to this had manipulated environmental elements (Smith and Curnow, 1966; Kotzan and Evanson, 1969; Chevalier, 1975). In general, researchers who use the term atmospherics tend to focus on one or two environmental stimuli and measure how they impact upon internal responses (such as emotions) and behavioural responses (such as spending more time in-store). There are a wealth of studies within this area which has provided interesting insights into the effects that individual stimuli can have on consumers. However, one major problem with this research is that the influence of stimuli which are excluded from the study cannot be gauged. Despite this flaw, the literature in this area is reviewed following a discussion on Mehrabian and Russell's (1974) theory.

The third stream of study has emerged from the retail literature and aptly refers to this area of research as the 'Store environment' (Baker, 1987; Baker et al, 1994; Baker et al, 2002). Unlike the work of atmospherics authors, those who have taken this view have attempted to take an all-inclusive approach to the measurement of the environment by

including multiple environmental stimuli (Baker et al, 1992; Baker et al, 1994; Baker et al, 2002). The final term, 'servicescapes', is a portmanteaux of 'services' and 'landscape' and was first introduced by Bitner (1992) and is used predominantly by academics within the services marketing realm. Bitner's (1992) conceptual framework expanded the earlier work of environmental psychologists (Mehrabian and Russell, 1974; Donovan and Rossiter, 1982) to include cognitive and physiological responses. In many ways Bitner's (1992) work is very similar to that of Baker et al (1992; 1994; 2002) as they both conceptualise the environment in terms of a number of discrete dimensions which work together to give a consumer an overall impression of the service environment. A number of differences do exist between their theoretical frameworks and this will be discussed in more detail later in the chapter.

While business academics use these four terms interchangeably (environmental psychology, store atmospherics, store environment, servicescapes), they are inextricably linked to Mehrabian and Russell's (1974) stimulus-organism-response theory upon which each are based. The next section discusses this theory in more detail and then provides an insight into those studies that have implemented this theory in a consumer context. Following this, a review of the studies on atmospherics studies is provided. Finally, multi-dimensional models of service environments are presented. This section first introduces Bitner's (1992; 2000) work on the area and the various proponents of her model. Following this, the various servicescape dimensions that have appeared in the literature are discussed. Within this section, theoretical advances made by Baker (1992; 1994; 2002) are presented. Finally a theoretical conceptualisation is developed. This

conceptualisation clarifies the earlier conceptualisations of the model by bringing together the various research streams that exist in this area.

1.3 The Mehrabian and Russell Model of Environmental Psychology (M-R Model)

As earlier highlighted, Mehrabian and Russell (1974) are widely recognised as the founding fathers of environmental psychology (Foxall, 1984). They developed a model which draws from the cognitive psychology school of thought wherein it uses the stimulus-organism-response (SOR) paradigm (see Figure 1.1). In this model, the stimulus refers to the physical environment and it includes physical features such as colour, lighting, décor and temperature. Only a general measure for stimulus is provided by Mehrabian and Russell (1974) and is measured in terms of how novel or complex the environment is. This is referred to as 'information rate'. Novelty involves the unexpected, the surprising and new, while complexity refers to the number of features and the extent of the motion within the environment. These stimuli are then thought to trigger internal responses in the individual which are referred to as the organism. The organism consists of three affective responses; pleasure, arousal and dominance (PAD). Pleasure is indicated by respondents' verbal assessments of the environment: happy rather than unhappy, pleased rather than annoyed, satisfied rather than dissatisfied, content rather than melancholy, hopeful rather than despairing and finally, relaxed rather than bored. Arousal is determined on the basis of how respondents rate the following extremes; stimulated to relaxed, excited to calm, frenzied to sluggish, jittery to dull, wide-awake to sleepy and aroused to unaroused. The final dimension, dominance is then measured using the following polar opposites; controlling to controlled, influential to

influenced, in control to cared-for, important to awed, dominant to submissive and autonomous to guided. These emotional responses are then followed by either one of two behavioural responses; approach or avoidance. Approach behaviours are positive responses to the environment such as a desire to stay, to explore and to communicate with others in the environment. On the other hand, avoidance behaviours include not wanting to stay, not wanting to explore, or to avoid communicating with others present. From a consumer behaviour perspective, wanting to explore an environment relates to the decision to shop or not to shop in a store and may also relate to the length of time spent in the store, while communication involves the willingness to communicate with sales staff (Donovan et al, 1994; Yalch and Spangenberg, 2000).

Mehrabian and Russell (1974) found the PAD dimensions to be factorially orthogonal, however the model does specify an interaction between pleasure and arousal. For example, if the environment is pleasant, it will result in higher levels of arousal which in turn leads to higher approach behaviours. However, if the environment is unpleasant, the greater the arousal and the propensity to avoid the environment becomes greater. From this it seems that arousal attenuates behavioural responses and it is likely that when arousal does not exist then behavioural responses will be less exaggerated. Using the measurement scales devised by the authors, the model was used to examine the interaction between the physical environment and human behaviour in a number of situations including, housing, offices, schools and prisons and all results generally supported the model.

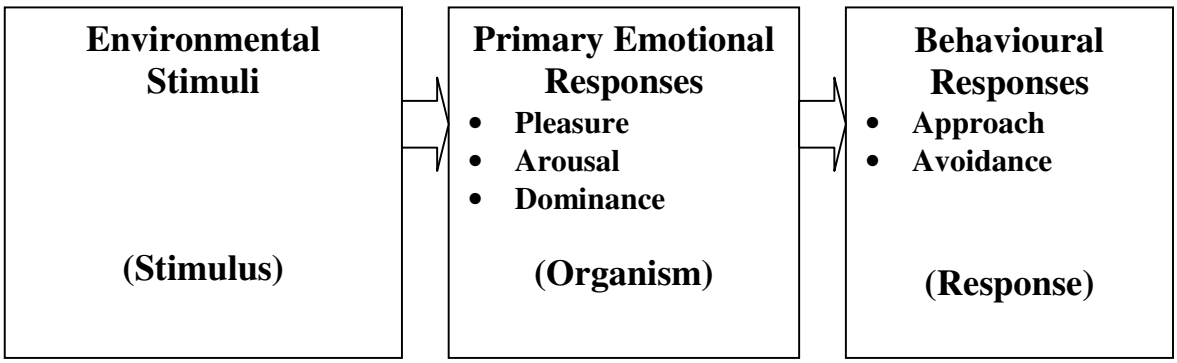


Figure 1.1: Mehrabian and Russell Model (1974)

1.3.1 Consumer Behaviour Applications of Mehrabian and Russell's (1974) Model

This model was introduced with the primary intention of being applied to organisational settings such as hospitals, schools and work environments. However, since its inception numerous marketing scholars have applied it to consumer behaviour situations (e.g. Donovan and Rossiter, 1982; Donovan et al, 1994; Sherman et al, 1997). The first of these is Donovan and Rossiter's (1982) seminal article on retail atmospherics. This study demonstrated that the Mehrabian and Russell (M-R) model could be successfully applied to consumer behaviour. In measuring the environment, no particular environmental stimuli was isolated, instead, Mehrabian and Russell's (1974) information rate scale was used which measures the environmental 'load' factor. It was found that in-store induced pleasure was a very powerful determinant of approach behaviours. Arousal was measured on the basis of in-store alertness and excitement and similar to the results on pleasure, high levels of arousal can lead to greater time spent in the store and willingness to interact with the sales staff. However, findings for dominance, the third emotional

dimension, were quite weak and did not seem to relate well to in-store behaviours. As a result of this, some studies have not included dominance as an emotional influence on approach/avoidance behaviours (Donovan et al, 1994; Sherman et al, 1997).

In summary, the results from this study indicated that perceptions of the interior influence approach and avoidance behaviours, in particular, time spent in the environment and sales. However, Donovan and Rossiter (1982) employed quite a small sample size (n = 30), which consisted entirely of student subjects. In addition, attitudes and intentions were measured rather than actual shopping behaviour. Despite these methodological limitations, Donovan and Rossiter's (1982) work is well regarded by academics and prompted a stream of research in this area. A follow-up study by Donovan et al's (1994) attempted to overcome the limitations of their earlier study. Findings from this indicated that in-store pleasure was a significant predictor of the length of time respondents remained in store with higher levels of pleasure associated with longer periods of time respondents stayed in-store. In addition, pleasure was also found to be a predictor of actual money spent in-store. These findings are very encouraging and add credence to the theory that emotional responses influence behaviour. However, in this study respondents did not rate the environment itself, so it becomes difficult to determine whether the emotional states reported by respondents can be attributed to the store environment.

Building on the work of earlier authors, Sherman et al (1994) not only examined the general ambience of the environment but also explored the influence of social, image and

general design aspects on emotions (pleasure and arousal) on subsequent overt behavioural responses. A methodological advancement was also made by the authors through the use of structural equation modelling which was used to test the interrelationships between the variables. Dependent variables included; time spent in the store, liking of the store, number of items purchased and money spent. It was found that social factors and the design of the store had a positive impact on pleasure, while ambience affected arousal. Subsequently, pleasure had a positive impact on money spent in the store, time spent in the store and the number of items purchase. These results are not directly comparable to Donovan et al's (1982; 1994) as additional dependent and independent variables were included in the study. Despite this, the results generally support that environments influence emotions which in turn influence behaviour. A slightly more recent application of the M-R Model was completed by Tai and Fung (1997) also lends some credence to the usefulness of the Mehrabian and Russell (1974) Model. It was found that the majority of their hypotheses were accepted, however, their environment measure (Mehrabian and Russell's the information rate) was found to only partly predict emotional states.

1.4 Atmospherics Literature: Empirical Research in Which Individual Environmental Stimuli Were Examined

As discussed above, since Mehrabian and Russell (1974) first introduced their S-O-R model many have sought to test its applicability in a variety of service environments. Others have isolated key environmental stimuli such as music (Bruner, 1990; Kellaris and Kent, 1992; Areni and Kim, 1993; Kellaris and Kent, 1993; Herrington and Capella,

1994; Dubé et al, 1995; Herrington and Capella, 1996; Oakes, 2000; Yalch and Spangenberg, 2000; Chebat et al, 2001; Dubé and Morin, 2001; Garlin and Owen, 2006; Jacob, 2006), odours (Hirsch, 1995; Mitchell et al, 1995; Bone and Ellen, 1999; Chebat and Michon, 2003; Davies et al, 2003; Ward et al, 2003), colour (Bellizzi et al, 1983; Bellizzi and Hite, 1992), and lighting (Areni and Kim, 1994; Summers and Hebert, 2001). In the sections that follow each of these variables are discussed in turn.

1.4.1 Music

While the studies in Section 1.3 have holistically analysed the impact of environmental stimuli on emotions and subsequent behaviours, others have isolated key environmental variables and examined their influence on behaviours. Music is the most studied environmental variable, is considered an important component of retail atmosphere (Smith and Curnow, 1966; Herrington and Capella, 1996) and is one of the easiest environmental variables to manipulate, therefore, it will be discussed foremost. This will be followed by a discussion of the remaining environmental stimuli; olfactory cues, lighting and colour.

As indicated above, the influence of music on consumer behaviour has been extensively studied. Over the years manipulations have included; volume (Smith and Curnow, 1966; Herrington and Capella, 1996), tempo (Milliman, 1982; Milliman, 1986; Areni and Kim, 1993; Kellaris and Kent, 1993; Herrington and Capella, 1996; Chebat et al, 2001; Sweeney and Wyber, 2002), foreground versus background music (Yalch and Spangenberg, 1990; Yalch and Spangenberg, 1993), musical genre (Areni and Kim,

1993; Yalch and Spangenberg, 1993; Sweeney and Wyber, 2002), familiarity (Areni and Kim, 1993; Yalch and Spangenberg, 2000), preference (Herrington and Capella, 1996) and emotions elicited as a result of the music played (Dubé et al, 1995; Chebat et al, 2001; Mattila and Wirtz, 2001).

Volume and Tempo

One of the first experimental studies was that of Smith and Curnow (1966) who used music loudness as the independent variable. In this study music was varied from loud to soft in eight supermarkets. Results showed that in the loud sessions significantly less time was spent in the stores and the rate of spending was greater (Smith and Curnow, 1966). However, a more recent study found no support for this thesis (Herrington and Capella, 1996) which questions the legitimacy of Smith and Curnow's (1966) earlier results. Milliman (1982) conducted two separate experiments (1982; 1986) where he manipulated music tempo, one in a supermarket (1982) and the other in a restaurant (1986). In the supermarket study customers spent more time and money when slow music was played than when foreground music was played. In the second study customers took longer to eat their meals compared to those in the fast music condition. It was also found that people spent more money on drinks in the slow-music condition. Herrington and Capella (1996) are critical of Milliman's (1982; 1986) methodology as it utilised two different musical compositions (one for the slow tempo and another for the fast tempo). They suggest that this does not control for other characteristics of the music such as musical preference, musical style, key, mode, harmonies and timbre. In order to overcome this limitation, Herrington and Capella (1996) used the same musical

compositions but played at alternative tempos. Interestingly it was found that tempo did not influence the shopping time or money spent as previously found by Milliman (1982; 1986). While Herrington and Capella (1996) and Milliman (1982; 1986) used two treatment levels (fast and slow), a later study by Kellaris and Kent (1994) added to this by including music played at a moderate tempo. While Milliman (1982; 1986) examined the effects of tempo on overt behaviours such as levels of expenditure and time spent in the environment, Kellaris and Kent (1994) extended this line of research by measuring the effects of tempo on pleasure and arousal. It was found that faster tempi classical music led to more pleasure for respondents. However pleasure did not increase arousal for those exposed to classical music. Quite different results were found for those exposed to pop music, as faster tempo led to increased arousal. While Kellaris and Kent (1994) examined the emotional effects of tempo, Chebat et al (2001) contributed further to the body of knowledge by examining the effects of tempo on cognitive activity. Cognitive activity was described as the depth of information processing and the number of thoughts held by the respondent and it was found that soothing music (low arousal) leads to higher levels of cognitive activity.

Variations in Musical Style

Building on the work of Milliman (1982; 1986), Yalch and Spangenberg (1990) compared the effects of foreground (Top 40 music with lyrics) and background music (instrumental easy listening music). It was found that when background music was played, levels of arousal increased, as did unplanned buying and unplanned time spent in the store increased. As noted above, it was found by Chebat et al (2001) that low arousal

music leads to higher levels of cognitive activity. This may explain Yalch and Spangenberg's (1990) findings, where background music, similar to slow, soothing music, does not interfere with cognitive processes therefore allowing the consumer to concentrate on the task at hand. Yalch and Spangenberg (1990) also explored differences between age groups and contrary to the researchers expectations, it was found that younger customers did more unplanned spending and felt they spent more time in store when background music was played. In this study shopping times were self-reported, therefore it could not be determined if the shoppers actually shopped for longer (Yalch and Spangenberg, 1990). This was rectified in their follow-up study (Yalch and Spangenberg, 1993) which examined the effect of playing different types of music in different parts of a store. It was found that actual shopping time was longer in the less familiar background music condition, but perceived shopping time was longer in the more familiar foreground music condition. It was also found that playing appropriate music for a department enhances the likelihood of purchases and spending more money. Contrary to their earlier work, it was found that younger shoppers bought more when exposed to foreground music. With regards to gender, it was found that women preferred background music and were more likely to make a purchase when this kind of music was played. On the other hand, men were more likely to make a purchase and spend more when foreground music was played.

Yalch and Spangenberg's (1993) results raised the possibility that musical styles (they examined top-forty versus instrumental easy listening music) can influence consumer behaviour. Which then led a number of authors to extend this line of research and vary

the genre of music to gauge its impact on consumer responses (Areni and Kim, 1993; Herrington and Capella, 1994; Herrington and Capella, 1996). In Bruner's (1990) synthesis of the literature on music in marketing, he suggested that genre of background music may have an influence on consumer preferences and perceptions. This hypothesis was later tested by Areni and Kim (1993) who examined the influence of two musical styles (Classical versus Top-Forty) in a wine store. Drawing on the earlier work of MacInnis and Park (1991) who proposed the notion of congruency theory, they hypothesised that when music is appropriate, or 'fits' the context, it becomes more persuasive. To test this proposition, Areni and Kim (1993) speculated that classical music would elicit more positive behaviours than in the Top 40 condition. They believed that because drinking wine is often thought of as sophisticated, almost snobbish that it would correspond with peoples' attitudes towards classical music. Their results overwhelmingly agreed with MacInnis and Park's (1991) contention as it was found that in the classical music, congruent condition, shoppers spent more. Interestingly, it was found that consumers did not increase the quantity of wine purchased, but chose higher priced wine when classical music was played. As a study by Herrington and Capella (1996) found that the amount of money spent was positively related to shoppers' level of preference for the background music, it is possible that the customers sampled in Areni and Kim's (1993) wine-cellar study had an a priori preference for classical music and it was this that led them to increase the amount they spent, not the musical 'fit' or appropriateness of the music to the environment. Adding further credence to this is Yalch and Spangenberg's (2000) study which found that familiar music impacts positively on product evaluations. The authors measured product evaluations using seven

semantic differential scales which included items such as lowest performance/highest performance and not at all stylish/very stylish. It was found that shoppers evaluated products more favourably when exposed to familiar music compared with unfamiliar music. This study cannot be directly compared to the work of Areni and Kim (1993) as it used music familiarity as the independent variable, rather than music preference, however, it does suggest that partiality towards the music positively influences both overt and intrapersonal behaviours.

1.4.2 Olfactory Cues

For thousands of years it has been believed that particular scents and aromas can have significant effects on both our physical and mental states (Maxwell-Hudson, 1999). The use of fragrances can be traced as far back as the Ancient Egyptians who both used them to embalm the dead and also for medicinal purposes (Maxwell-Hudson, 1999). In recent years business people have begun to recognise their potent effects on overt behaviours with a number of articles highlighting its effects on consumer responses. In her thorough servicescape framework, Bitner (1992) included many environmental dimensions one of which was ambient scent. Despite this, business academics have not been as enthusiastic about this environmental variable as the other dimensions in her model with this area not getting as much attention as it deserves (Turley and Milliman, 2000). Authors may be less enamoured by the notion that olfaction has an impact on consumers as the evidence to support its influence is both contradictory and scant (Bone and Ellen, 1999). Of all the senses, olfaction is thought to be closest linked to emotional responses (Michon et al, 2005), however the experimental evidence to support this is quite limited (Ellen and

Bone, 1998; Bone and Ellen, 1999). Much of the research to date on odours have taken place in simulated experimental settings (Mitchell et al, 1995; Spangenberg et al, 1996). Others have attempted to overcome the limitations of external validity associated with experimental designs and have conducted field research on odours and scents (Hirsch, 1995; Mattila and Wirtz, 2001; Chebat and Michon, 2003; Michon et al, 2005).

One of the first studies to appear in the business literature on odours was that of Mitchell et al's (1995). Ambient odour was chosen as the stimulus variable, this is a scent that is not emanating from any particular object but is present in the environment (Spangenberg et al, 1996). As with Areni and Kim (1993), the authors used MacInnis and Park's (1991) congruency theory as a basis for their research, and it was found that when the ambient odour was congruent with the product class the respondents spent more time processing data than when the subjects were exposed to an incongruent odour. They posit that the reason for this is that when congruency exists cognitive enrichment is occurring. Contradicting this, it was found by Blackwell (1995) that when visual cues did not match odours, cognitive analysis became more difficult for the subjects. However, the convergent results may be explained by methodological differences. Spangenberg (1996) extended the work of Mitchell et al (1995) as they tested the impact of varying scents and their intensity on consumer responses in a simulated store environment rather than two scented rooms. These authors applied the M-R model including the following approach behaviours; intentions to visit the store, purchase intentions for specific products, actual and perceived time spent in the environment and products examined in the store (Spangenberg et al, 1996). Cognitive states were also

measured and these were; evaluations of the store, evaluations of the store environment, merchandise in general and specific products. It was found that scent does have a profound effect on consumer behaviour as those exposed to the scented environment rated the store more favourably and rated both the store and its environment more favourably. Similarly in the scented environment, evaluations of the merchandise were more favourable, prices were thought to be lower and quality ratings were also more favourable. With regards to approach and avoidance behaviours it was found that those exposed to the scented environment felt they spent less time in the environment (however, they actually did not remain any longer) and examined more products. In addition, it was also found that the intensity or type of scent did not matter, even neutral scents were able to generate enhanced perceptions. It was also found that scent had no significant effect on mood.

While Spangenberg's (1996) study examined the influence of odours in a controlled environment, Hirsch (1995) added to the body of knowledge by testing his hypothesis in a field setting which has the advantage of measuring actual purchase behaviours, rather than purchase intentions. Hirsch (1995) chose a slot-machine setting in a Las Vegas Casino, using two odourant types which were placed in two different areas, with a third, unodourised area acting as a control. It was found that sales significantly increased in the area where Odourant 1 was placed. Moreover, when this scent was removed sales significantly decreased. Interestingly, in both the unodourised control area and the Odourant 2 area sales did not increase significantly. Contrary to Spangenberg et al's (1996) findings, this suggests that only certain scents can have an impact on consumer

behaviour. However, it is not possible to make any inferences with regards to the properties that these scents possess as Hirsch (1995) does not disclose the details of the two scents. A more recent study which was also conducted in a real-life situation was that of Mattila and Wirtz (2001) who manipulated both scent and music. They postulated that adding a pleasant scent and pleasant background music would enhance the perceived positive state of the store environment, the amount of pleasure associated with it, impulse buying and customer satisfaction of the service encounter. The results showed that congruency between the arousal levels of the scent and the background music led to enhanced evaluations of the shopping experience. A methodological advancement was made by Michon et al (2005) who used structural equation modelling to test the relationships between odours, shoppers emotions and perceptions of the retail environment, product quality under various levels of retail density. As with Spangenberg et al's (1996) findings, no main or interactive effects were found regarding scent on mood.

1.4.3 Lighting

Despite Kotler's (1973-1974) early call for research into lighting as an environmental variable, little research since then has been conducted on this area (Summers and Hebert, 2001). The first to publish research on the effect of lighting on consumer behaviour was Areni and Kim (1994). Using the Mehrabian and Russell (1974) model, the authors attempted to gauge the impact of in-store lighting on shopping behaviours (Areni and Kim, 1994). Similar to their work on in-store music, the study was conducted in a wine store where two manipulations were utilised; the first using soft lighting and the second

brighter lighting. It was found that consumers examined and handled significantly more items under the bright lighting condition than the soft lighting condition. An extension of this study was conducted by Summers and Hebert (2001) in two independently owned stores, one a hardware store (displaying tools) and the other a western apparel/feed store (displaying belts). These authors examined the influence of two supplemental lighting treatments on three behaviours; the length of time spent at the display, the number of items touched and the number of items picked up. In agreement with their hypothesis it was found that consumers examined and handled more items under the bright lighting condition. It was inferred from these results that the increased brightness caused higher levels or arousal and pleasure, yet no measure was included in the study to confirm this postulation.

1.4.4 Colour

Two simulated experimental studies were found to test the influence of colour on consumers (Bellizzi et al, 1983; Bellizzi and Hite, 1992). Drawing on the physiology and psychology of colour literature, Bellizzi et al (1983) examined the influence of colour on approach orientation, physical attraction and consumer perceptions of the store environment and merchandise. Experimental colours were; red, yellow, green, blue and white. Approach orientation was measured on the basis of how closely the subjects sat to the experimental colour wall of the laboratory. It was found that the respondents were physically drawn to warm colours (red and yellow) in comparison to the cooler colours (green and blue), despite the fact that the respondents reported that the warmer colour environments are generally unpleasant. An extension of this work was later conducted

which tested Kotler's (1973-74) contention that atmospheric variables may serve as an affect-creating medium which would in turn contribute favourably to purchase probability (Bellizzi and Hite, 1992). In the earlier colour study (Bellizzi et al, 1983), a variety of colours were tested, however, in this study only blue and red were selected as these are considered to be polar-opposites of one another. Red is thought to have negative and tense properties, but physically arousing, whereas blue is considered calm, cool and positive (Bellizzi and Hite, 1992). Bellizzi and Hite (1992) conducted two experiments, the first of which measured behavioural intentions and length of shopping time. In this first test, subjects in the blue environment were found to purchase significantly more than in the red environment. In addition, in the blue environment, half of the respondents selected the most expensive item of the choice criteria, while only 19% selected the most expensive item in the red environment. With regards to shopping time, no significant effects were found as a result of the colour manipulations. In the second experiment, blue and red colour manipulations were also used. Dependent measures included emotional states while in the store, in order to measure this, Mehrabian and Russell's (1974) semantic differential pleasure-arousal-dominance (PAD) was utilised. They also measured approach/avoidance behaviours, again using Mehrabian and Russell's (1974) scales. Similar to their first experiments' results, it was found that respondents in the blue group expressed a greater intention to shop, browse and buy in the simulated store. From both of these studies it is evident that colour conditions can have a profound influence on consumer behaviour with respondents in both studies reacting positively to both of the blue environmental conditions.

1.5 Dimensional Conceptualisations of the Servicescape

The positive results achieved from both the direct applications of Mehrabian and Russell's (1974) theory and the individual atmospheric studies have lent credence to the theory. Others have sought to bring together both of these streams of research which has led to a number of dimensional frameworks of services environments. Although conceptual uniformity within this area does not appear to exist with many variations on Mehrabian and Russell's (1974) model appearing in the literature over the years, the most frequently cited dimensional typology is Bitner's (1992) Servicescape Model. Other conceptualisations are indeed very similar to Bitner's (1992) model however important differences do exist. The following section first introduces Bitner's (1992) model and is followed by a discussion on the work of Baker (1987; 1992; 1994; 1998; 2002).

1.5.1 Bitner's (1992) Servicescape Model

The most widely regarded environmental framework is that of Bitner (1992). It was in this seminal work that Bitner (1992) modified Mehrabian and Russell's (1974) theory to develop a model explaining the processes in and around what she termed the 'servicescape'. In her wholly theoretical article, she extended Mehrabian and Russell's (1974) work to include both cognitive and physiological internal responses (see Fig. 1.2). Furthermore, she advanced the notion that both employees' and customers' behaviours were influenced by the service environment. On the left-hand side of the model Bitner (1992) describes environmental dimensions that can influence both consumer and employee behaviour. This litany of environmental stimuli is quite comprehensive,

however it is by no means complete and this she recognises herself. Unlike many atmospheric studies, Bitner (1992) makes the important point that consumers perceive their environments holistically rather than processing each discrete stimulus. She categorises environmental stimuli into three distinct groups; Ambient Conditions, Spatial Layout and Functionality and Signs, Symbols and Artifacts. As discussed earlier in this chapter, much of the work on store atmospherics has focused on ambient conditions such as music, scent, lighting and colour. The second category, Spatial Layout and Functionality, refers to the ways the furniture and equipment are arranged, in particular it relates to how these items facilitate customers and employees to accomplish their goals. The third group are those environmental variables that fall under the Signs, Symbols and Artifacts category. These range from signage that communicates the nature of the service to basic things like price lists and contact information. Also included in this dimension are more indirect tangible cues that convey to the consumer what level of quality they should expect to receive.

As with the Mehrabian and Russell (1974) S-O-R model, Bitner (1992) proposes that these environmental stimuli then lead to internal effects (see Figure 1.2). These internal effects take the form of cognitive processing, emotional reactions and physiological responses which affect both the consumer and the employee. Although Bitner's model is based on the S-O-R framework, two notable distinctions exist between the two. Firstly, the Mehrabian and Russell (1974) model focuses only on emotional reactions to the environment while Bitner (1992) has extended the theory to include cognitive information processing and physiological responses to the environment. Secondly, Bitner

also moves beyond Mehrabian and Russell's (1974) information rate measure of environments by describing the servicescape in terms of three discrete dimensions. On the whole, Bitner (1992) approaches environmental psychology theory from a more pragmatic perspective than earlier authors, calling for a coherent theory that service managers can apply so that the 'factory' in which the service is delivered will have a strong impact on customers' service experience.

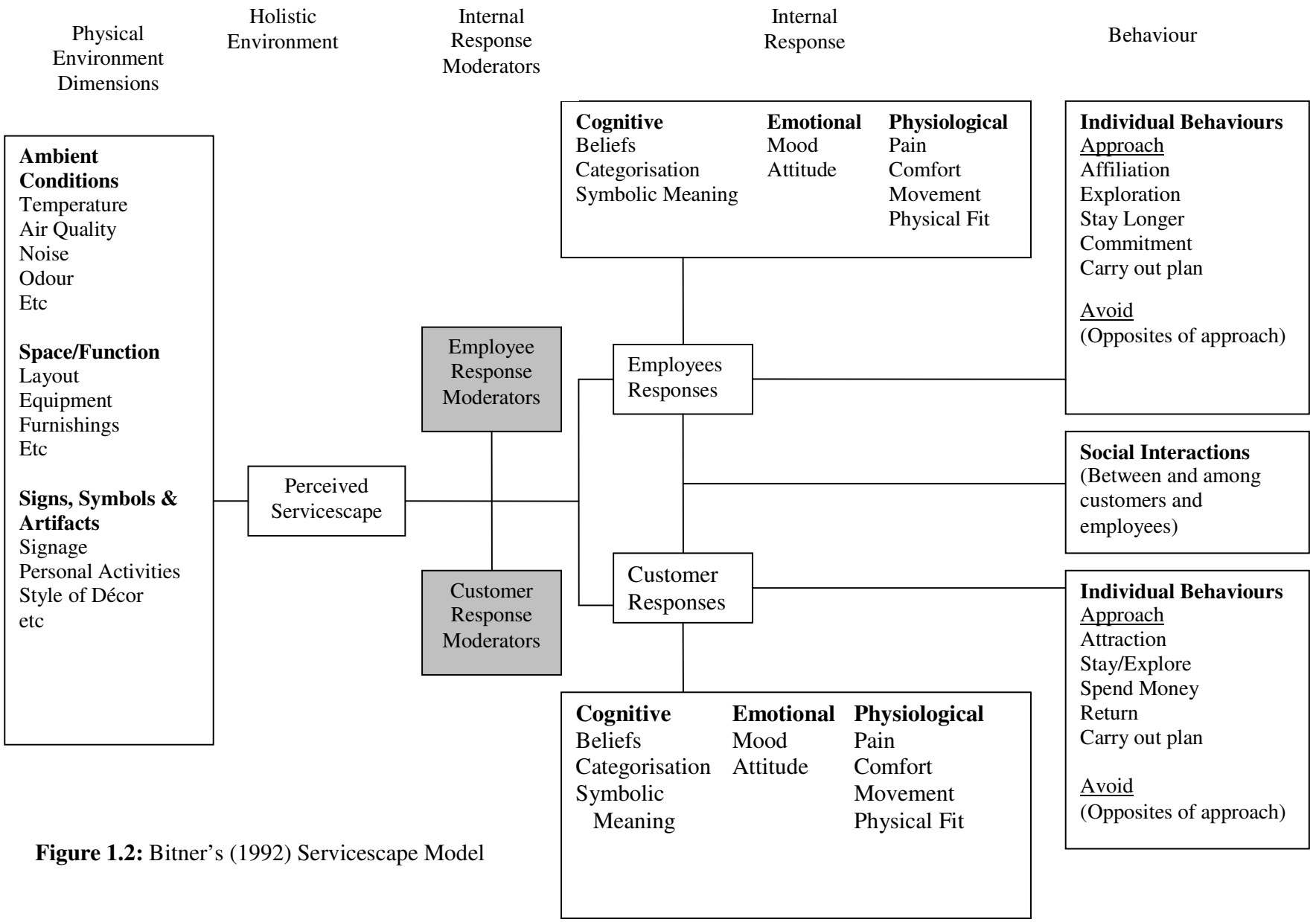


Figure 1.2: Bitner's (1992) Servicescape Model

1.5.2 Baker's Model of Store Environments

While Bitner's (1992) typology is based directly on M-R's theory of environmental psychology, a second model which has also received a notable amount of attention in the literature is Baker's (1987) framework. While Bitner (1992) and Baker's (1987) models bear some similarities, the major point of difference between the two is that Baker does not draw on the stimulus-organism-response theory as she entirely omits the mediating role of emotions. Baker's (1987) model also takes a multi-dimensional view of service environments where she describes a three factor structure which encompasses social cues, ambient conditions and design factors. These dimensions are different to those proposed by Bitner (1992), however, they both agree that consumers process service environments on a general level while also processing distinct stimuli within. They also both concur that environments influence behaviour (for Bitner this occurs indirectly through emotions) with Baker stating that environments have a direct impact on overt consumer responses.

Table 1.1: Baker's Servicescape Dimensions

Ambient Conditions	**Design**	**Social Factors**
Temperature	Architecture	(Other Customers)
Air Quality	Colour	Number
Noise	Scale	Appearance
Scent	Materials	Behaviour
Cleanliness	Shape Style	(Service Personnel)
	Layout	Number
	Comfort	Appearance
	Signage	Behaviour

Unlike Bitner, Baker has empirically tested the applicability of her model (Baker et al, 1994; Baker et al, 2002). Both of these studies chose experimental designs whereby

manipulations within the environment were made. The dependent variables chosen in her first study were: merchandise quality, service quality and store image (Baker et al, 1994). Her second study built upon her earlier findings by including additional dependent variables which were: store patronage, merchandise quality, merchandise value and time/effort costs. Findings from both studies supported the notion that environmental stimuli impact on consumers' inferences about the various outcome variables chosen. Unfortunately, whether emotions mediated this relationship was not tested. Further to this, the experimental method of these studies limits their generalisability.

1.5.3 Empirical Studies on the Servicescape

The remainder of the research on servicescapes has tended not to develop new theories or conceptualisations, but have instead sought to test both Bitner (1992) and Baker's (1987) models. These studies can broadly be categorised as those who have taken a single dimensional view or those who have taken a multi-dimensional view. For those who have taken the multi-dimensional perspective, these can be further classified into those who have used either multiple regression or structural equation modelling as their analytical tool.

Two of Wakefield and Blodgett's (1994; 1999) studies took a unidimensional view of the servicescape whereby the mediating role of excitement was examined in relation to approach behaviours and in both it was found that the servicescape was a potent predictor of behaviour. Hightower et al (2002) extended this line of work through the development of a conceptual model which incorporated a number of constructs not previously

modelled as both determinants and consequences of the servicescape. These included service quality evaluations, perceived waiting time, involvement and perceived value. Multiple items were used to measure the servicescape, however since regression analysis was the method of analysis here, these items were averaged and a single indicator was used. Findings indicated that the servicescape has a positive effect on service quality and customer satisfaction. Reimer and Kuehn (2005) took both a hedonic service (restaurant) and a utilitarian service (banking) and used these services to determine whether tangible services dimensions (the servicescape) indirectly effected service quality evaluations through intangible aspects of the service. Given that structural equation modelling produced a poor model fit, path analysis, using one servicescape indicator, was used as an alternative form of analysis to test the hypotheses. Interestingly, it was found that in the restaurant (hedonic service) sample the servicescape had a greater influence on service quality than in the bank (utilitarian service).

All of the above studies used either an overall servicescape evaluation (Wakefield and Blodgett, 1994) or aggregated servicescape items to form a single indicator (Wakefield and Blodgett, 1999; Hightower et al, 2002; Reimer and Kuehn, 2005). In line with both Baker and Bitner's models, each of the authors demonstrated how the servicescape serves as an antecedent to cognitive and behavioural responses. However, what these authors have failed to show, is whether the servicescape is perceived by consumers as a multi-dimensional structure. In recent years there has been a move towards multi-dimensional modelling of the servicescape structure whereby various aspects of the servicescape are described as discrete dimensions. These studies can be broadly classified into those who

have used regression-based analyses (Lucas, 2003; Johnson et al, 2004; Harris and Ezeh, 2008) and those who have used structural equation modelling (Wakefield and Blodgett, 1996; Ryu and Jang, 2007; Lee et al, 2008). Both Lucas (2003) and Johnson et al (2004) examined the impact of the servicescape in Las Vegas casinos. These authors did not explore whether emotions mediated service evaluations, but instead investigated whether a series of servicescape dimensions impacted satisfaction with the service. In both studies it was found that the servicescape is a significant determinant of customer satisfaction. Also taking a hedonic service, Ezeh and Harris (2008) chose six servicescape dimensions (music, aroma, cleanliness, implicit communications, furnishings, and physical attractiveness). All of the servicescape dimensions, apart from music were found to significantly affect loyalty intentions.

Of those who have employed structural equation modelling, the first to appear in the literature was Wakefield and Blodgett's (1996) study which was conducted in sports stadiums and a casino. The servicescape was divided into a five dimensional structure which consisted of, layout accessibility, facility aesthetics, seating comfort, electronic equipment displays and facility cleanliness which were modelled as antecedents to an overall servicescape measure which subsequently impacted upon satisfaction with the servicescape and repatronage intentions. El Sayed et al (2004) also took a dimensional approach to conceptualising the servicescape where an experimental design was employed in which various different manipulations of the mall environment were presented. Dimensions used here were: music, interior design, crowding, location and lighting. The results were analysed using ANOVA and it was found that all constructs,

apart from interior design, had a significant effect on behavioural intentions. More recently, Ryu and Jang (2007) deconstructed the servicescape into six dimensions (facility aesthetics, lighting, ambience, layout, dining equipment and employees) and examined each constructs' effect on emotional responses and behavioural intentions in a restaurant setting. Many of the structural paths between the environmental constructs and emotional responses were insignificant or achieved path coefficients that were less than the standard cut-off of .41. Why this occurred is unclear, however it may be due to the measures used. The authors do not present how their servicescape measures were devised and it is possible that they were inappropriate for the situation. Lee et al (2008) conducted an analogous study in a festival environment using a seven-dimensional structure. As with Ryu and Jang (2007), many of the paths hypothesised in the structural model were not supported. Upon examination of the constructs used to represent the servicescape, it becomes clear that perhaps Lee et al's (2008) servicescape dimensions do not represent service environments. For instance, consumers' perceptions of souvenirs, program content, food, information and convenience seem to be more in line with consumers' service evaluations than servicescape evaluations.

The first problem with these studies is that there has been a heavy emphasis on hedonic service contexts such as casinos, (Wakefield and Blodgett, 1994; Wakefield and Blodgett, 1996; Wakefield and Blodgett, 1999; Lucas, 2003; Johnson et al, 2004) sports stadiums (Wakefield and Blodgett, 1994; Wakefield and Blodgett, 1996; Wakefield and Blodgett, 1999), recreation centres (Wakefield and Blodgett, 1999) and restaurants (Reimer and Kuehn, 2005; Ryu and Jang, 2007; Harris and Ezeh, 2008). The second problem that

becomes evident is that each of the servicescape studies that have appeared in the literature have been highly inconsistent with each new paper conceptualising the servicescape in a different way. While there is no uniformity in the literature, the overall conclusion that can be taken from these studies is highly encouraging as each study could statistically support the potential power of the servicescape on postconsumption processes. The analytical tools employed by authors have varied greatly from regression analysis (Hightower et al, 2002; Lucas, 2003; Johnson et al, 2004), ANOVA (El Sayed et al, 2004) to the use of more sophisticated techniques such as structural equation modelling (Wakefield and Blodgett, 1996; Tang et al, 2001). More recently, further developments have been made through the conceptualisation of the servicescape as a multi-dimensional structure. For these studies there seems to be quite an arbitrary approach in deciding what elements should be included as part of the servicescape structure. As mentioned above, Lee et al (2008) included items that were clearly non-environmental variables in their conceptualisation of the servicescape. The items used in other studies are less problematic in that they quite definitely represent the service environment (Wakefield and Blodgett, 1996; Lucas, 2003; Johnson et al, 2004; Ryu and Jang, 2007). However these studies seem to tailor their environmental items according to the service under scrutiny (e.g. casinos, sports stadiums, restaurants) thus making the findings more difficult to generalise. In addition, this context specific approach to the development of items also makes replication of these studies into different industries more problematic (e.g. Wakefield and Blodgett, 1996; Ryu and Jang, 2007). In the next section, the dimensional ambiguities in the literature are examined. From this a new

servicescape conceptualisation is presented which will be tested in a utilitarian retail setting.

1.5.4 Servicescape Dimensions that Frequently Occur in the Literature

The work of Bitner (1992) has quite definitely had an immense impact upon subsequent research and it is from her article that a healthy research stream has developed. However, as briefly discussed above, consensus does not exist as to what the servicescape comprises (Wakefield and Blodgett, 1996; Baker et al, 2002; Lee et al, 2008). The various models appearing in the literature have not caused any debates between these authors per se, but it has led to a disordered and disjointed body of knowledge. Given the complexity of the servicescape concept, it is unlikely that an ideal service composition actually exists that can be applied to any given service industry (Kotler, 1973-1974; Ezeh and Harris, 2007). However, it is possible that certain generic dimensions can be identified, isolated and subsequently applied to a wide variety of contexts. In this next section various different dimensions that have appeared in the literature are examined.

Ambience

Ambient conditions can generally be regarded as those that effect the five senses (Bitner, 1992). When ambient factors are maintained at a satisfactory level they do not directly motivate consumer purchases (Ezeh and Harris, 2007), however if any of these ambient cues fall outside their acceptable limits, or are absent from the environment, they have negative effects on consumer behaviour (Baker, 1987; Hightower, 2003).

It can be seen from the atmospherics literature that ambient cues are one of the most widely studied aspects of the servicescape (Turley and Milliman, 2000). Ambience factors that have been extensively studied include music (Bruner, 1990; Areni and Kim, 1993; Kellaris and Kent, 1993; Herrington and Capella, 1994; Herrington and Capella, 1996; Oakes, 2000; Yalch and Spangenberg, 2000; Chebat et al, 2001; Dubé and Morin, 2001; Mattila and Wirtz, 2001; Sweeney and Wyber, 2002), lighting (Areni and Kim, 1994; Summers and Hebert, 2001) and scent. A number of dimensional studies have also included ambience in their studies however, there is no consistency in the literature with regards to what constitutes ambient factors. For instance, Wakefield and Blodgett (1999) categorise cleanliness and employee appearance as ambient factors and it is doubtful whether these were correctly defined as ambient factors. Other ambient factor classifications are less obscure, although no two authors include the same factors in their study. Table 1.2 below highlights some of the inconsistencies in the literature:

Table 1.2: Ambience Items Used in the Literature

Authors	**Items Used**
Ryu and Jang (2007)*	Relaxing music Pleasing music Comfortable temperature Enticing aroma
Wakefield and Blodgett (1999)	Kept clean Temperature is comfortable Employees are neat appearing
Wakefield and Baker (1998)	Music x 2 Lighting Temperature
Baker et al (1994)**	Music Lighting

* A separate dimension measuring lighting was also included in this study
** Experimental study

From the above literature it is expected that music, lighting, temperature and olfactory cues will have an impact on consumer responses. By incorporating each of these dimensions a more thorough overview of ambient cues will be presented.

Design

Unlike ambient cues, design elements are more visual in nature (Baker, 1987; Hightower, 2003) with the overall interior design of a facility communicating a multitude of messages to the consumer from how to behave, to the standard of the service they should expect to receive (Bitner, 1992). Of all the servicescape dimensions it is this dimension that is likely to have the most significant impact upon consumers' impressions of the service (Wakefield and Blodgett, 1996; Aubert-Gamet, 1999) as it is this dimension that helps consumers form first impressions of the service provider. Unlike Baker (1987), Bitner (1992) does not propose a separate dimension for design, but instead places style of décor under 'Signs, Symbols and Artifacts'. She also includes in this dimension communication objects that serve as explicit cues, examples of which are signage which aids wayfinding and company logos and branding.

Others take Baker's (Baker, 1987; Baker et al, 1994) viewpoint, whereby design elements are categorised as being either functional or aesthetic (Baker, 1987; Baker et al, 1994; Hightower, 2003; Ezeh and Harris, 2007). Comfort, privacy and layout are subsumed under functional aspects of design while architecture, materials, colour and style are incorporated into the aesthetic elements of this dimension. Other authors regard the layout and space of an environment as a discrete dimension in itself (Wakefield and

Blodgett, 1996; Wakefield and Baker, 1998) and this dissertation will follow this line of research by describing layout as a separate construct.

Space/Layout

Within a service context, the layout and space dimension relates to the way furnishings and equipment are arranged and the spatial relationships among these objects (Bitner, 1992; Wakefield and Blodgett, 1996; Ryu and Jang, 2007). The positioning of objects and furnishings within the environment should facilitate the free flow of movement throughout the service facility and ought to allow customers to accomplish their tasks during the service encounter (Bitner, 1992; Nguyen, 2006). For example, aisle space must be sufficiently large so as to allow shoppers to browse comfortably without feeling crowded or cramped (Harrell et al, 1980; Eroglu and Harrell, 1986; Hui and Bateson, 1991; Machleit et al, 2000). The issue of space becomes much more pertinent for self-service environments as customers must navigate through the structure of the facility to serve themselves (Aylott and Mitchell, 1998). Employing more 'user-friendly' layout designs should improve customer search time and service reliability thus facilitating service delivery (Sulek and Lind, 1995). If the layout is such that it inhibits movement throughout the environment it may lead to frustration on the part of the customer which can subsequently lead to a more dissatisfying experience (Bitner, 2000). More specifically, spatial crowding in terms of cluttered shelves, narrow, irregular aisles can negatively effect customer satisfaction with the service (Machleit et al, 2000). The ability with which individuals are able to move easily around an environment has also been found to have a positive effect on perceived quality (Wakefield and Blodgett, 1996),

levels of excitement and desire to stay at the facility (Wakefield and Baker, 1998). Ideally, the service environment should be structured so as to assist individuals' movement rather than impede it (Newman, 2007).

Equipment

In self-service situations this aspect of the service environment is also likely to have a significant impact upon consumer evaluations of the service as it too determines how well the consumer can perform various tasks (Bitner, 1992). This dimension is critical in communicating the image of the store to customers (Ryu and Jang, 2007). For example, outdated tills and technology implicitly states that the quality of service that the customer should expect is below average. In a critical incident study of service failures by Hoffman et al (2003), it was found that a significant proportion of servicescape failures were attributed to mechanical problems such as inoperative computers and other equipment which subsequently inconvenienced the customer. Hoffman et al (2003) recommends implementing quality controls that assess whether mechanical or equipment failures have occurred thus eliminating the likelihood of future equipment failures.

Hygiene

The cleanliness of a service environment is without doubt one of the key aspects of the servicescape that customers equate with quality perceptions (Wakefield and Blodgett, 1996). Of all servicescape failures reported by Hoffman et al (2003) cleanliness issues were reported the most frequently. In fact, cleanliness issues were regarded by respondents as being the most serious service failures experienced. In reality maintaining

a clean and hygienic environment should be one of the easiest and cheapest of the servicescape dimensions to control.

Employees/Social Factors

With regards to employees two alternate approaches seem to exist; one which incorporates employees and other customers as stimuli within the servicescape and one which does not. Bitner's (1992) work takes the latter view, regarding the servicescape as a stimulus which impacts upon employees and customers rather than an environmental variable that influences respondents' perception of the servicescape. While, Bitner considers the servicescape as merely a material stimulus, Baker (1987; 1994; 1993; 2002) moves beyond this by regarding the servicescape as both a material stimulus and a social construct in which the actions and interactions between people shape individuals servicescape evaluations. Baker (1987; 1994; 1993; 2002) draws on the work of Russell and Snodgrass (1987) who claimed that the presence of other individuals are an integral part of any environment. From a services or retail perspective, this includes the 'number, type, and behavio[u]r of other customers and sales personnel', (Baker, 1994, p. 331). In a similar vein to Baker (1987), Tombs and McColl-Kennedy (2003) present a conceptual model in which they argue that the social environment in which a service takes place has a significant effect on emotional and behavioural responses. Drawing on Social Facilitation Theory (Zajonc, 1965), they propose that 'the mere presence of others will elicit monitoring of other social behavio[u]rs' (Tombs and McColl-Kennedy, 2003; p. 454). No direct operationalisation of Tombs and McColl-Kennedy's (2005) model has appeared in the literature however, for their study, Hightower et al (2002) devised items

which measured the helpfulness, appearance and friendliness of other customers. Unfortunately these items did not meet reliability criteria and were thus dropped from further analyses. However, it has been found by Chebat et al (1995) that customers evaluate the service delivery process in terms of the service they received but also in terms of how they observe other customers to be treated.

1.6 Assessing the Dimensionality and Measurement Validity of the Servicescape

One of the aims of this dissertation is to conceptualise and develop valid measures of key dimensions of the servicescape construct. Most existing measures of the servicescape use either single item scales, have questionable measurement properties, or their measurement properties have not been clearly assessed in terms of discriminant and convergent validity. In addition, dimensional representations of the servicescape have been wholly inconsistent (see Appendix 1A) thus highlighting the need for a meaningful conceptualisation of the servicescape to be presented. Previous measurement efforts are also plagued with items that attempt to measure constructs that should not be subsumed under the servicescape domain (Wakefield and Blodgett, 1999; Baker et al, 2002; Lee et al, 2008). From this it is obvious that there is a need to overcome these problems by 1) developing a context-free conceptualisation of the servicescape that can be applied across a broad spectrum of industries, 2) developing a set of objective servicescape indicators that facilitate empirical measurement and thus guide future research into this area.

While Bitner's (1992) model made advances in terms of spurring such a healthy body of research, it falls short in terms of how to operationalise the construct. Finer scrutiny of

her model reveals that she is relatively vague with regards to what should be incorporated under each dimension. For example, layout and equipment are categorised under the Space/Function dimension yet theoretically these seem quite distinct. Similarly, furnishings are also placed under the Space/Function dimension yet it seems more sensible for items relating to furnishings to be placed alongside style of décor which is placed within the Signs, Symbols and Artifacts dimension. Baker's contributions are clearer in their propositions in that items were developed for each dimension and subsequently tested in both an experimental (Baker et al, 1994) and a field study (Baker et al, 2002). Despite the significant additions to the literature made by Baker (1987; 1994; 1998; 2002) and others (Wakefield and Blodgett, 1994; Wakefield and Blodgett, 1996; Wakefield and Blodgett, 1999; Hightower et al, 2002), theoretical definitions of the servicescape remain quite ambiguous and unclear. It is proposed that the servicescape can be conceptualised as a six dimensional construct which includes the following dimensions: ambience, design, space/layout, equipment, hygiene and employees. These constructs will be measured as a second-order construct. Second order constructs are typically used in contexts where there are several related constructs measured by multiple items (Rindskopf and Rose, 1988; Chen et al, 2005). As such, it is hypothesised that each of these dimensions will be highly correlated and can thus be accounted for by a common higher order factor; namely the servicescape.

It is hypothesised that the ambience dimension will reflect those items included by previous authors; namely, music, temperature, lighting and odours. As mentioned above, heretofore work on ambient cues tends to take quite a haphazard approach to item

selection. It is hoped that items will be chosen to represent this sensory construct in a more thorough manner than previous research in this area. The design element of the model will capture environmental aesthetics such as the interior design and decorative aspects of the servicescape along with communication cues and signage materials. Space and layout will be modelled as a discrete construct which will measure issues such as the ease of movement throughout the environment and consumers ability to find products in store. In line with the work of both Wakefield and Blodgett (1996; 1999) and Ryu and Jang (2007), it is proposed that consumers will describe equipment as distinct from other constructs in the model. Items relating to the cleanliness and hygiene of the environment have been included in both multi-dimensional (Wakefield and Blodgett, 1996) and single construct conceptualisations (Reimer and Kuehn, 2005; Hightower, 2002) of the servicescape. This dissertation follows the work of Wakefield and Blodgett (1996) who have successfully demonstrated that consumers' perceptions of hygiene can be modelled as a construct in itself. The employees dimension is in keeping with previous authors' propositions, this dimension will include items relating to both the appearance and friendliness of staff.

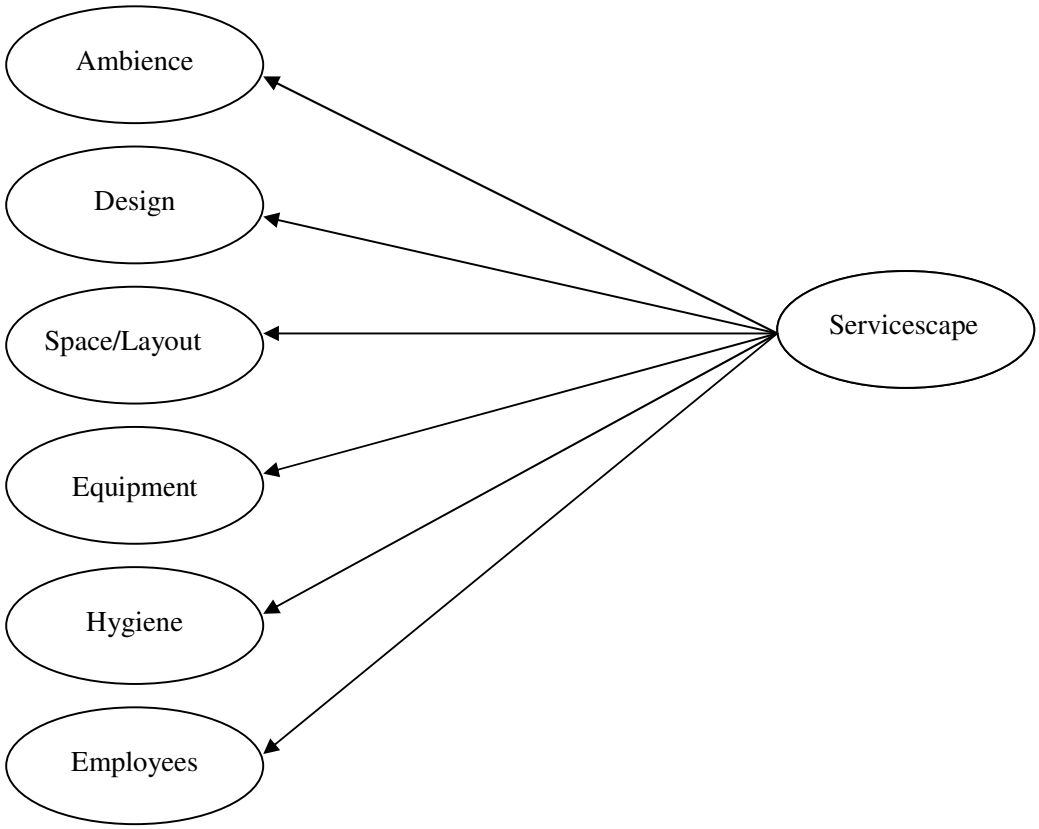


Figure 1.3: Proposed Theoretical Framework: A Second Order Factor Model of the Servicescape

1.7 Measuring the Servicescape: A Call for Research into Non-Hedonic Service Situations

The servicescapes literature seems to heavily emphasise hedonic service situations over more traditional retail contexts (Arnould and Price, 1993; Wakefield and Blodgett, 1996; Arnould et al, 1998; Wakefield and Blodgett, 1999; Ryu and Jang, 2007; Lee et al, 2008). This line of research has more than likely grown from an implicit acceptance that in more extended service transactions the servicescape will play quite a significant role in

customers' service experience. Indeed it seems quite logical that when an individual spends a considerable amount of time in a service environment perceptions of the servicescape are likely to come to the fore when making post-consumption evaluations. This has led to quite a paucity of work focussing on more traditional or utilitarian services, however Reimer and Kuehn's (2005) study is the exception to this. This study compared responses in a utilitarian service to those from a hedonic service to determine whether the role the servicescape played in service quality evaluations differed across the two contexts. It was found in the bank sample (utilitarian service) that the servicescape had a significant effect on service quality evaluations, however, in comparison to the restaurant sample (hedonic service), this relationship was significantly weaker. Reimer and Kuehn's (2005) findings demonstrate that for utilitarian services, the impact of the servicescape on quality evaluations are indeed significant, however the strength of this relationship is slightly weaker than that found in the restaurant. As Reimer and Kuehn (2005) measured the servicescape at the aggregate level, we cannot discern if certain aspects of the servicescape play a greater role than others. Unlike much of the servicescapes literature, the one-stimulus, atmospherics literature has also focussed heavily on utilitarian services such as traditional retail environments and authors within this domain have routinely demonstrated the ability of specific environmental stimuli to affect post-consumption processes.

While results from both Reimer and Kuehn's (2005) work and the atmospherics studies are promising, they suffer from a number of key disadvantages. Firstly, although Reimer and Kuehn's (2005) results are encouraging, further research needs to explore the

applicability of a dimensional structure of the servicescape in a retail setting. Unfortunately for Reimer and Kuehn (2005) measurement problems with their structural model led them to aggregate their data and use path modelling. As mentioned previously, the atmospherics literature suffers from some severe limitations with regards to the heavy emphasis on single stimulus studies. This research hopes to add to the services literature by examining a comprehensive servicescape structure in a retail service environment. As highlighted in the introduction, the retail market plays a significant role in the Irish economy. Therefore, to fully understand the variables which impact upon consumer behaviour in this important sector is of utmost importance to practitioners working in this industry.

1.8 Conclusions

It is evident from the above review that the literature in this area is highly disjointed which has resulted in quite a diverse literature base. The main cause of this is most likely to be due to conceptual ambiguities associated with Bitner and Baker's models. Therefore, given the confused nature of the literature to date, research needs to address this limitation by clearly defining what is subsumed within the servicescape and then holistically examining how the various stimuli interact with one another in a field environment. Previous to this, authors have attempted to investigate the influence of the servicescape on various post consumption process such as service quality, satisfaction and behavioural intentions (Wakefield and Blodgett, 1994; Lucas, 2003; Johnson et al, 2004) but these studies have been confined to more unusual services such as casinos and sports stadiums. Clearly research needs to address this limitation and explore the

servicescapes influence on consumer behaviour in more traditional services settings such as retail environments. In addition to the emphasis on more unusual service contexts, there has been a strong focus on studies which have modelled the servicescape as a single construct. Given the multi-faceted nature of the servicescape it makes psychometric sense to model this construct consisting of a number of distinct but conceptually similar dimensions. Thus far, attempts to model the servicescape as a multi-dimensional construct have lacked consistency with regards to the selection of dimensions with most authors in the area taking quite an arbitrary approach to factor choice decisions.

This chapter presented a servicescape model that includes dimensions from the extant literature that provides a more holistic overview of this concept than previously considered. As well as bringing together the literature in a cohesive and structured manner, a second purpose of this dissertation is to extend the theoretical evidence on the relationships among the servicescape and other aspects of the postconsumption experiences, such as: service quality, emotions, satisfaction and behavioural intentions. Prior to testing these relationships, there is a need for a servicescape scale that is both psychometrically and dimensionally sound. The next chapter reviews the various servicescape scales that have appeared in the literature and it is from this that it becomes obvious that there was a need for a servicescape scale to be developed. It then outlines the scale development process which led to the construction of a twenty-four item instrument.

2.0 SERVICESCAPE SCALE DEVELOPMENT

2.1 Introduction

As noted in Chapter One, the term 'servicescapes' only came into being after Bitner's (1992) seminal article on the topic. Since then, various attempts to measure the servicescape have been proffered (Sherman et al, 1997; Hightower et al, 2002; Lucas, 2003). Unfortunately many of these efforts have fallen short, primarily due to poorly conceived representations of the construct. These conceptual and methodological irregularities have led to a disordered body of knowledge that lacks any coherent understanding of what is subsumed within the servicescape construct. The previous chapter highlighted these limitations and provided a conceptual framework that overcomes these shortcomings. So that this framework can be tested empirically, a servicescape scale needed to be chosen as a measurement instrument. Choosing an appropriate scale requires that each scale previously published needs to be appraised to determine whether it meets the requirements of the current research problem. Therefore, the purpose of this chapter is to present the various servicescape scales that have been published to date and in doing so highlighting the limitations of these scales. The chapter will begin by tracing the development of these measures, beginning with the core work on this area, namely Mehrabian and Russell's (1974) information rate. The discussion will then turn to consumer research studies and the methods used to quantify consumers' reactions to environment. Following this, a review of servicescape scales will be conducted and it is from this review that it becomes evident that a servicescape scale needed to be developed to comprehensively measure consumers' reactions to the service

environment. The chapter then turns to the steps which were taken to arrive at a full servicescape scale. The section will conclude with a discussion on the servicescape scale that will be used for this study, along with a rationale and justification for its choice.

2.2 Mehrabian and Russell's (1974) Information Rate

The foremost scale available to researchers was Mehrabian and Russell's (1974) information rate. This scale consists of a number of semantic differentials and was developed in conjunction with an emotion measure and an approach-avoidance scale. Relatively few consumer behaviour studies have used Mehrabian and Russell's (1974) information rate, however, Donovan and Rossiter's (1982) original study directly applied each aspect of the Mehrabian and Russell (1974) model which included the information rate, pleasure-arousal-dominance and their approach-avoidance measure. Despite Mehrabian and Russell's (1974) model specifying that arousal is a direct response to the information rate in the environment, Donovan and Rossiter (1982) could not find support for this as it was found that information rate was a poor predictor of an individual's level of arousal. In addition to this, questionable reliabilities were also found with Cronbach alpha coefficients as low as .40 and high cross-loadings occurring in an exploratory factor analysis. A later application of this scale by Tai and Fung (1997) did however find some support for this scales' applicability. Firstly, it was found that information rate was the only significant predictor of arousal yet had no significant effect on pleasure. Further analyses in which the information rate was regressed against behavioural intentions found a significant relationship. Overall, this shows that the predictive power of the information rate measure is fairly limited with only partial support for environment →

emotional response relationship was found. The limited predictability of the measure, in conjunction with the tenuous reliability issues found by Donovan and Rossiter (1982), seems to have led to the relative obsolescence of this measure.

Despite these problems, Mehrabian and Russell's (1974) information rate does offer the advantage of being a very parsimonious measure. However, because this scale measures the environment on a very general level, it does not isolate for managers the aspects of the service environment that may be causing problems. By not informing practitioners on which components of the environment they need to modify to optimise consumer behaviour, this scale become becomes redundant from a practical perspective. This failing, coupled with tenuous measurement issues, has resulted in a measure of the environment that offers little in terms of contributing to theory and was thus disqualified on that basis.

2.3 Indirect Measures of the Service Environment

While Donovan and Rossiter's (1982) first study employed Mehrabian and Russell's (1974) information rate, their later study utilised an indirect measure of the environment. Building on their previous work, Donovan et al (1994) used a field study to examine the role of the environment on shopping behaviours. In this study they did not use any measure of the environment per se, instead inferences were made based upon pleasure levels experienced by individuals. For example, it was assumed individuals who provided neutral or negative scores on the pleasure scale must have done so because they found the environment unpleasant. By equating negative responses on the pleasure scale

to unpleasant environmental evaluations is a theoretical leap that is somewhat unfounded. Other researchers have not made such suppositions and have tended to implement experimental designs. By employing experimental designs, it allows the authors to side-step the need to devise or utilise an environmental measure as these studies manipulate the stimuli under study and therefore infer how the environment effects behaviour based on how the respondents react to the different conditions.

Two of Baker et al's (1992; 1994) studies have exercised this technique. In both studies the authors did not directly question respondents on their perceptions of the environment but instead manipulated the stimuli to represent various service settings. In their first study they used a 2 x 2 factorial design to simulate store ambience levels (low and high) and two store social levels (low and high). In their second study (Baker et al, 1994) a 2 (Prestige vs. Discount Ambient) x 2 (Prestige vs. Discount Design) x 2 (Prestige vs. Discount Social) factorial design was applied. Others have also used this technique, such as Hansen (2002) who implemented this form of design, whereby they constructed two situations; one represented elegant physical surroundings while the other represented less elegant surroundings. While indirect, experimental measures of the servicescape offer the advantage of ease of use and convenience, they potential reduce external validity so as to increase internal validity (Winer, 1999). In addition, there is the possibility that demand artefacts may have impacted upon the validity of such studies (Shimp et al, 1991; Perrien, 1997).

2.4 Multi-Item Measures of the Servicescape

Sherman et al (1997) reinterpreted Mehrabian and Russell's (1974) stimulus-organism-approach model using structural equation modelling. Unlike many of the atmospherics studies, Sherman et al (1997) measured respondents in store rather than in a laboratory setting. Environmental stimuli included in the study were: lighting, colour schemes and evaluations of the general interior. The scale was based upon items suggested by Dickson and Albaum (1977) along with a number of items which were developed to measure in-store environment. This consisted of thirty-one semantic differential items which were reliability tested using exploratory factor analysis. This revealed seven factors with many low loading and double loading items. Several items were thus dropped leaving twenty items in total. These twenty items represented four dimensions: Social factor (4 items), overall image (2 items), design factors (10 items) and an ambience factor (4 items).

Wakefield and Blodgett (1994; 1996; 1999) have written a number of articles on the influence of the servicescape on responses such as satisfaction and quality in a number of leisure services. While the first of Wakefield and Blodgett's (1994) papers included a one-item 'satisfaction with the servicescape item', subsequent papers overcame this limitation through the use of multi-item scales (Wakefield and Blodgett, 1996; Wakefield and Blodgett, 1999). The first of these studies employed a scale that was specifically designed for the service at hand (Wakefield and Blodgett, 1996) and given the leisure service context of this study (casinos and sports stadiums) it is difficult to successfully

translate these items to other contexts. For example, the layout items relate specifically to gaming areas (4 items) and the equipment items refer only to the quality of the visual displays (5-items). However, within this scale not all items are context specific (e.g. facility aesthetics items) and could be used in other situations.

A more recent study by Wakefield and Blodgett (1999) utilised a shorter scale which is less context specific and could possibly be applied to a variety of services. This scale included four items on design elements, three on equipment and three on ambient influences. This measure is more superior to other Wakefield and Blodgett (1994; 1996) scales as it is less context specific, however it omitted a number of key items relating the layout, employees and the cleanliness of the environment. In addition to this, the environmental factors chosen by the authors for this study are considerably different to those chosen for their earlier study (Wakefield and Blodgett, 1996). In the former study the dimensions chosen to represent the servicescape were: layout accessibility, facility aesthetics, seating comfort, electronic equipment and displays and cleanliness. Whereas for their more recent study (Wakefield and Blodgett, 1999) the items chosen were: design, equipment and ambience. In both of these studies no clear rationale was provided for choosing these dimensions and it seems as though quite an ad hoc approach to choosing variables was taken.

Both Sherman et al (1997) and Wakefield and Blodgett's (1994; 1996; 1999) scales cover a variety of servicescape stimuli, however many environmental features which were included in the servicescape conceptualisation presented in Chapter One are absent. The

most thorough servicescape scale that has appeared in the literature to date is that of Reimer and Kuehn (2005). This scale is comprised of eleven questions which addressed almost all key stimuli outlined by previous authors in this area (Baker, 1987; Wakefield and Blodgett, 1996; Wakefield and Baker, 1998; Wakefield and Blodgett, 1999; Baker et al, 2002). Using structural equation modelling, Reimer and Kuehn (2005) attempted to model all thirteen of these items onto one construct with little success. It is unclear why this occurred but there is a possibility that these items were not uni-dimensional but instead consisted of a series of separate subdimensions. If a thorough overview of the servicescape is sought, modelling the servicescape in this way will provide only limited results. Ideally, the servicescape construct should be divided into a series of inter-related constructs, allowing for inter-correlations between the dimensions. Second-order factor models are a parsimonious and statistically stringent method of conceptualising a construct (Rindskopf and Rose, 1988) and this is the approach that is to be taken for the current research problem. This will allow each of the dimensions to be analysed in relation to each other and also to all other constructs in the full structural model. Given the wide variety of servicescape items used by Reimer and Kuehn (2005), all items from their scale were included in the scale development item pool which is discussed later in this chapter.

An attempt to develop a comprehensive servicescape battery was made by Hightower et al (2002). Unlike previous efforts to construct a servicescape measure, this instrument is based directly on the conceptual work of Bitner (1992) and Baker (1987). A further methodological advancement was made by using exploratory factor analysis to test the

reliability of the scale. The twenty-eight items produced a six-factor solution which resulted in a number of items being deleted as they were found to have high cross-loadings. Once these unreliable items were deleted, 14 items remained which all loaded onto one factor capturing 52.74% of the variance. Items excluded from further analyses which did not meet reliability criteria related to: music, noise levels, temperature, the influence of other customers, architecture, signage and parking facilities and restrooms. Since such a large number of items were deleted in this scale, unfortunately, the final instrument was not a complete representation of Bitner (1992) or Baker's (1987) frameworks.

While the above research streams are direct interpretations of Bitner's (1992) servicescape model, a second body of research has evolved from Baker's (1987) store environment literature. As discussed in Chapter One, Baker's (1987) conceptualisation of store environments is in many respects quite similar to Bitner's (1992) model. Tantamount to Bitner's (1992) space and function elements is Baker's (1987) design dimension. However a key difference lies in their placing of social influences. Bitner (1992) sees environmental stimuli as impacting upon both consumers and employees, while Baker (1987) regards the employees themselves as environmental factors which effect consumer responses. Both Hightower et al (2002) and Reimer and Kuehn (2005) have incorporated this view into their instruments through the inclusion of items such as, 'The employees are neat and well dressed', 'The employees are friendly', and 'The employees are helpful'. In Baker's (1987) first attempt to explore the magnitude of the environment's effect on consumer responses she took an experimental approach which

did not require the use of a store environment scale. Instead, high and low arousal environments were artificially simulated using a 2 x 2 factorial design. A more recent study by Baker et al (2002) extended this line of research by assessing the impact of environmental elements on customers' perceived merchandise value, monetary prices, interpersonal service and store patronage intentions. As with her earlier work, Baker (2002) chose an experimental design here also. However, for this study, an instrument was designed to measure environmental constructs. Previous work on store atmospherics (Donovan and Rossiter, 1982) and environmental psychology (Mehrabian and Russell, 1974; Russell and Pratt, 1980) provided the basis for store environment perception items. Three dimensions, loosely based on her store environment conceptualisation (Baker, 1987) were assessed. These were: design elements, employee elements and music perceptions. Design elements consisted of three questions related to; the colour scheme, attractiveness of facilities and merchandise organisation. With the three employee perception items taken from the SERVQUAL battery (Parasuraman et al, 1988). Music perceptions were the only ambient cue included and referred to its pleasantness, the appropriateness of the music and whether or not consumers found it to be bothersome. Baker et al's (2002) work is indeed a step closer to understanding the dimensional structure of the servicescape and how these sub-constructs impact on other post-consumption processes. However, this work falls short in relation to the range of servicescape dimensions included in her studies. For instance, important aspects of the environment are excluded relating to ambient cues, layout and the cleanliness of the environment.

Table 2.1: Servicescape Scales

Scale Description	Author(s)	Limitations to the Scale
Information Rate Measure a general measure of environments	Mehrabian and Russell (1974)	An abstract measure of the environment that does not provide any indication on how various aspects of the servicescape perform Poor predictive ability and low reliabilities found
Single Item Measures	Wakefield and Blodgett (1994)	- as with the above measure, single item approaches do not highlight which dimensions are more salient than others
Multiple Item Measures	Sherman et al (1997)	- Semantic Differential used which does not lend itself well to SEM
	Wakefield and Blodgett (1996; 1999) Hightower et al (2002)	- Highly context-specific - Incomplete representation of the servicescape with many dimensions measured using one item only
	Reimer and Kuehn (2005)	- All servicescape dimensions measured with one item only
	Baker et al (2002)	- Incomplete representation of the servicescape. Many key issues not addressed.

It can be seen from the above discussion that approaches to measuring the servicescape environment vary widely. Choosing an appropriate scale is an arduous task as each scale previously used needs to be assessed in isolation. Bearing the current research problem in mind, all of the scales available were considered and it became obvious that none of the above scales would meet the research needs of this dissertation. Table 2.1 above has summarised the scales used by a number of key studies, noting their deficiencies and thus highlighting the need for a multi-dimensional servicescape scale to be developed.

2.5 Servicescape Scale Development

The above discussion has highlighted that the servicescape literature is deficient of a comprehensive, reliable servicescape battery. Having addressed many of the servicescape stimuli proposed by earlier authors (Baker, 1987; Bitner, 1992), Reimer and Kuehn's (2005) scale is by far the most superior scale that has appeared in the literature to date. However, it too suffers from the problem of having too few indicators which would cause identification problems at the structural equation modelling stage of this research. Given the problems associated with both Reimer and Kuehn's scale (2005) and each of the above scales, it became apparent that a servicescape scale needed to be constructed which addressed these limitations. To achieve this, a scale development process was embarked upon which ultimately resulted in a twenty-four item servicescape scale. It is proposed that this scale will adequately measure each of the proposed servicescape dimensions outlined in Chapter One wherein multiple indicators can be used to measure each construct. The next section details the procedures used in the scale development process and the analytical techniques used to do so.

To fill the need for a reliable and valid servicescape scale, a twenty four item servicescape scale was developed. Numerous servicescape scales have been developed and studied in the past. These scales have generally fallen short, with none of them comprehensively or systematically developed. Indeed, some servicescape scales are simply better than others in that they cover many of the servicescape stimuli proposed in the literature (e.g. Hightower et al, 2002; Reimer and Kuehn, 2005). However, the

reliability and validity of many of these measures must be questioned. Some servicescape scales have been developed through factor analysis (Hightower et al, 2002), but others have been constructed on a purely ad hoc basis (Wakefield and Blodgett, 1996; Wakefield and Blodgett, 1999; Reimer and Kuehn, 2005). Clearly, there is a need for a reliable and valid servicescape scale that is also brief and easy to administer. Adhering to recommendations for scale development (Churchill, 1979; Spector, 1992; DeVellis, 2003), a servicescape scale was developed.

Although there is little difference between scale development models (Worthington and Whittaker, 2006), for the development of the servicescape scale, a combination of DeVellis's (2003), Spector's (1992) and Netemeyer et al's (2003) steps were followed. These steps are: (a) Construct definition (b) Generation of an item pool, (c) Determine the format of the measure, (d) Review the initial item pool, (e) Administer items to a development sample, (f) Evaluate the items and (g) Optimise scale length. The sections that follow deals with each of these stages in succession.

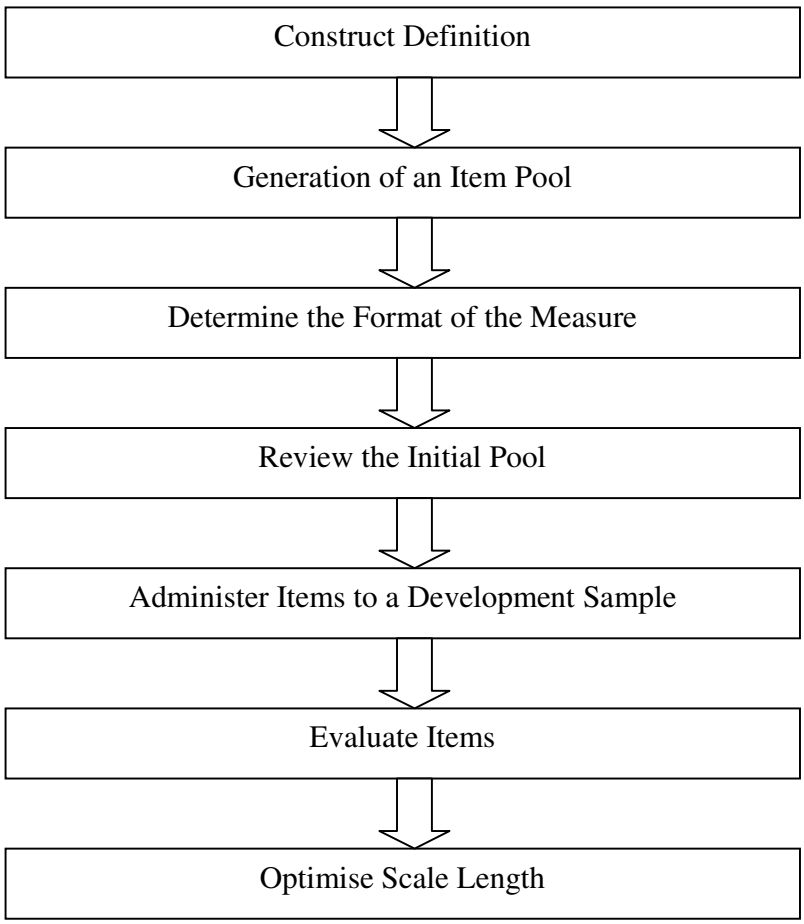


Figure 2.1: Scale Development Process

2.5.1 Construct Definition

Clarifying exactly what a scale intended to measure is 'deceptively obvious' (DeVellis, 2003, p. 60) but as Spector (1992, p. 7) notes, it is in this very step that 'scale efforts go astray'. If a construct is ill-defined it is possible that peripheral items are included or items of interest are omitted (Worthington and Whittaker, 2006). For this reason, conceptualising the construct is one of the most important steps in scale development (Netemeyer et al, 2003). To aid in defining the servicescape a comprehensive literature

review was conducted. Tracing the literature from Mehrabian and Russell's (1974) and Kotler's (1973-1974) work to the present day work of Bitner (1992; 2000), Wakefield and Blodgett (1994; 1999), Hightower (2002) and Reimer and Kuehn (2005) and others. Drawing on the literature reviewed in Chapter One and for the purposes of instrument design, the servicescape was defined here as '*the environment in which the service is delivered which encompasses both tangible and intangible atmospheric cues*' (Bitner, 1992; Bitner, 2002). As this research intends to portray the servicescape as a second order factor structure in which the various dimensions of the servicescape can be explained by six first-order factors there is also a need to define each of the sub-dimensions of the model. Each of these factors will be measured using their own sub-scales which are modelled as distinct constructs. As outlined in Chapter One these dimensions are: Ambient conditions, Design Elements, Space and Layout, Equipment, Hygiene and Social Factors. Each of these are defined in turn in Table 2.2 below.

Table 2.2: Servicescape Dimensions

Dimension	**Definition**
Ambient Conditions	Ambient conditions are temporal variables which generally impact upon the five senses. Items to be included under this construct are: music, lighting, scents and temperature.
Design	Design elements relate to all those aspects of the servicescape that contribute to the aesthetic appeal of the service environment. Evaluative items pertaining to interior design and décor fall under this category. This dimension also includes all those stimuli that facilitate consumers with the task at hand, such as signage materials.
Space & Layout	This pertains to the arrangement of furnishings within the environment and whether their arrangement facilitates the free-flow of customer traffic through said environment.
Equipment	This dimension encompasses both the appearance and functionality of electronic or mechanical equipment that the customer may use or observe during the service experience.
Hygiene	Hygiene factors refer to the appearance of cleanliness both within and outside the store.
Social	Social factors can be defined as including the appearance and friendliness of both other customers and service delivery staff.

2.5.2 Generation of an Item Pool

Based upon the above dimensional definitions, the literature was further reviewed so that items could be generated. From this review a complete list of all servicescape items were compiled and categorised according to their similarities with one another (these can be viewed in Appendix 2A).

One of the key objectives of this research was to measure the servicescape as comprehensively as possible. To meet this end, Bitner's (1992) servicescape dimensions were scrutinised so that each aspect of the servicescape proposed by her was included in the item pool. In addition to this, authors such as Wakefield and Blodgett (1994; 1996; 1999), Baker (1987; 2002) and Hightower (2002) were also analysed. From these studies a relatively large item pool was generated. It must be noted that items pertaining to the friendliness of employees was also included. In Bitner's (1992) original work on servicescapes she proposed that environmental stimuli influence the behaviours of customers and employees alike. However, Baker et al's (2002) and Reimer and Kuehn's (2005) work include the employees as environmental stimuli. Given this discrepancy in the literature, it was decided to include employees as a servicescape element as the majority of studies have taken this approach.

It is recommended that the item pool includes as many items as possible so as to achieve redundancy (DeVellis, 2003). Redundancy allows the phenomenon of interest to be revealed in many different ways thus increasing reliability, while any irrelevant redundancies will cancel out when the items are subjected to factor analysis (DeVellis,

2003). In addition to the items taken from the literature, a number of items were also written by the author and added to the item pool to ensure the servicescape concept was adequately captured. Using the researcher as a source of items is common practice in scale development studies and should not cause validity problems (Netemeyer et al, 2003). Using DeVellis's (2003) recommendations, the approach that was taken was to generate additional items but tap into the construct in an alternative way. Care was taken not to generate exceptionally lengthy or difficult items. A number of negatively worded items were written to avoid acquiescence and agreement bias (Spector, 1992).

The final item pool covers each of the above environmental stimuli with multiple items from each. The complete list of these items and their origin is listed in Table 2.3. Upon examination of the items listed below it becomes obvious that the scale development items are slightly different to those used in the final service station study. The reason for this is that the original context of this dissertation was to be dental services. Unfortunately gaining access to a dental surgery was not possible. A number of dentists were contacted and asked to participate in the study, however, all dentists contacted declined the offer. This called for an unexpected re-examination of the chosen industry which ultimately led to an entire change of service context. The scale developed below was altered slightly to suit the final service context which was service stations. The rationalisation for this change in context is discussed in greater detail in Chapter Four as it is in this chapter that the dimensionality of the servicescape is tested using structural equation modelling.

Table 2.3: Servicescape Item Pool

Servicescape Items	Authors
Ambience	
Music	
I found the music at the dental surgery bothersome	Baker et al, 2002
The background music was pleasant	Reimer and Kuehn, 2005
The background music was appropriate	Hightower et al, 2002
Odour	
The odour was unpleasant	Self-developed
The surgery had a pleasant smell	Hightower et al, 2002
Lighting	
I found the lighting comfortable	Reimer and Kuehn, 2005
I found the lighting excellent	Reimer and Kuehn, 2005
Temperature	
The room temperature was pleasant	Reimer and Kuehn, 2005
Employees	
The employees were neat and tidy in appearance	Reimer and Kuehn, 2005
I found the employees at the dental surgery friendly	Baker et al, 2002
I found the employees at the dental surgery unhelpful	Self-developed
The employees were helpful	Baker et al, 2002
Hygiene	
I found the dental surgery very clean	Reimer and Kuehn, 2005
The surgery was very unhygienic	Self developed
The bathroom facilities were clean and tidy	Self developed
I found the physical facilities very unclean	Self developed
The other patients at the surgery were neat and well dressed	Hightower et al 2002,
Communication and Signage and Brochures	
The brochures and other communications material were visually appealing	Reimer and Kuehn, 2005
There were appropriate reading materials to read in the waiting room	Self developed
I found the signs used (e.g. bathroom, enter, exit,) helpful to me	Hightower et al, 2002
Design & Décor	
The physical facilities were visually unappealing	Self developed
I found the interior design visually appealing	Reimer and Kuehn, 2005
The waiting room was attractive	Self developed
I found the physical facilities of this dentist surgery to be superior to others	Hightower et al, 2002

The colour schemes were pleasant	Baker et al, 2002
The style of the interior was fashionable	Hightower et al, 2002
The materials used were of high quality	Self developed
The exterior of the dentist's was appropriate	Self developed
The architecture of the surgery was attractive	Reimer and Kuehn, 2005
The parking facilities were satisfactory	Hightower et al, 2002
The waiting room was comfortable	Self developed
The seating in the surgery was comfortable	Wakefield and Blodgett, 1996
I found the physical facilities comfortable	Self developed
Equipment	
The dental equipment was modern looking	Wakefield and Blodgett, 1999
The electronic equipment was of high quality	Wakefield and Blodgett, 1999
The electronic equipment was excellent	Wakefield and Blodgett, 1999
The dental equipment was of poor quality	Self developed
Layout and Flooring	
The surgery had more than enough space to be comfortable	Hightower et al, 2002
I found the dental surgery easy to find	Self developed
The flooring was appropriate	Self developed
I found my way around the surgery quite easily	Wakefield and Blodgett, 1996
The interior layout was pleasing	Hightower et al, 2002
The surgery had a poorly organised layout	Self developed

2.5.3 Format of the Measure

The predominant literature on servicescapes has used Likert scales to measure the phenomenon. Likert scales employ declarative statements to which the subjects' level of agreement with each statement used as the response choice. This scale format is widely used in the marketing literature to measure opinions, beliefs and attitudes and are thus a suitable measure for the current context. Due to the widespread use of Likert scales, most respondents are quite familiar with this format and thus find these questions relatively easy to fill out. For these reasons, Likert-type scales are the most commonly used scale format in the marketing literature (Cox, 1980). Likert scales are also preferable to other scaling formats such as semantic differentials, because the resulting data from Likert

scales is amenable to both Confirmatory and Exploratory Factor Analysis. The next problem pertained to the number of levels the scale would contain. Five-point scales anchored with 'strongly disagree' and 'strongly agree' with a neutral mid-point were used here. Simulation tests evaluating the optimal number of scale points have shown that reliability increases up to about five categories (Bandalos and Enders, 1996) thus justifying the choice of a five-point scale.

2.5.4 Item Review

The item list was discussed with colleagues and supervisors to determine whether any key aspects of the service environment were omitted. This generated items, 'There were appropriate reading materials in the waiting room' and 'the waiting room was attractive'. This stage of scale development ensured that each item measured what it intended to measure. It also evaluated items' in terms of face validity, salience, clarity, readability and conciseness with regards to their wording and grammatical structure.

2.5.5 Administration of Items to a Development Sample

The respondents chosen for the development of this scale were students enrolled on a number of undergraduate business degree courses at the Dublin Institute of Technology. Using student samples is generally not advised (Spector, 1992) as they may provide mean values that are either higher or lower than one would expect from the population (DeVellis, 2003). However, DeVellis (2003) states that having a mean value that is not representative of the population is not necessarily a problem as an accurate picture of internal consistency can still be achieved and thus indicate which items are most strongly

related. In addition, students as a category of consumers are well suited to the purpose of this study seeing that they are familiar with the service and this was ensured by the inclusion of a question which related to the last time they had used this particular service. For these reasons using a student sample was deemed both appropriate and acceptable.

The questionnaire was administered online using an internet survey service (www.surveymonkey.com). Respondents were contacted via email and invited to participate in the survey. The email briefly explained the purpose of the research and also contained a link to the questionnaire. Participants were contacted three times in total, the first inviting them to participate, the second as a reminder and the third, last call to participate. Data collection took place between January and February 2007.

Recommendations relating to sufficient sample size vary widely, with each author providing a threshold that conflicts with many others in this area. Nunnally (1978) recommendations that a sample of 300 is adequate for scale development while others suggest an item to response ratio of 1:4 (Floyd and Widaman, 1995; Hinkin, 1995). In fact, research has also found that a sample size as low as 150 is sufficient for exploratory analysis (Guadagnoli and Velicer, 1988). A sample size of 205 was achieved here which meets the criteria of a 1:4 ratio and also of Gualagnoli and Velicer's (1988) minimum sample size of 150. Of the 205 student respondents, of which, 35% were male, 45% were female and 20% did not complete this question. Finally, a response rate of 59.7% was achieved which was considered highly satisfactory.

A scenario approach was used where individuals were asked to reflect on their last experience of the service. This approach is widely used in the literature and has been proven to be a reliable technique (Diamantopoulos and Schoefer, 2008). This involved asking respondents to reflect on their last experience with the service and to use their memory to picture the atmosphere and the interior design of the service setting. Full details of this questionnaire can be found in Appendix 2C.

2.5.6 Psychometric Evaluation of Items

Once all respondents completed the questionnaires it was necessary to evaluate the items so that the most suitable items could be identified and thus be incorporated into the final scale. The goal here was to create a measure of the servicescape that was both precise and theoretically sound. The most widely used item selection criteria has been based on some form of internal consistency analysis (Clark and Watson, 1995). Although Churchill's (1979) highly cited article on scale development advises that internal consistency analysis (such as Cronbach's alpha) should be conducted prior to exploratory factor analysis, this is no longer considered the correct sequence of analyses as internal consistency analysis suffers from some serious disadvantages. Many authors misinterpret the true meaning of internal consistency analysis and calculate Cronbach's alpha in an attempt to achieve a unidimensional scale (Gerbing and Anderson, 1988). However, Cronbach's alpha does not inform us on the unidimensionality of a scale (Cortina, 1993) but 'refers to the overall degree to which the items that make up a scale are intercorrelated' (Clark and Watson, p. 315). Cortina (1993) provides an example of a situation where a scale can provide a reasonable level of internal consistency when in

actual fact the scale is multi-dimensional in nature. He artificially constructed a scale which contained 18 items and calculated the alpha of these items. This provided a coefficient alpha of .85 which on first inspection this seems satisfactory, however, there are in fact two orthogonal sub-scales within this 18 item scale containing nine items each. This anomaly occurs when many high coefficients are averaged with many low ones thus hiding any multi-dimensionality that may exist. In addition to alpha coefficients disguising multi-dimensionality, the statistic is also dependent on the number of items in the scale (Gerbing and Anderson, 1988) with the statistic artificially inflating coefficients when a large number of items are contained in the scale (Cortina, 1993; Clark and Watson, 1995). Finally, there are also no clear standards regarding an acceptable level of reliability for Cronbach's alpha. Originally, Nunnally (1978) recommended thresholds of .80 or .90, however in a more recent text (Nunnally and Bernstein, 2006) no such recommendations for an adequate cut off point are given.

In light of the disadvantages associated with internal consistency analysis, psychometricians now recommend that scale developers should begin item analysis by conducting factor analysis (Comrey, 1988; Floyd and Widaman, 1995). This is especially pertinent when it is expected that a number of dimensions or subscales exist in the data (Netemeyer et al, 2003). It is often the case in both marketing and other social sciences that the construct being measured is conceptualised in terms of a single higher order dimension under which a number of inter-related dimensions are subsumed. As the servicescape is hypothesised to contain a number of discrete sub-dimensions, it is

necessary to use exploratory factor analysis so that the underlying factors can be revealed (Netemeyer et al, 2003).

Factor Analysis

Given the above shift in the measurement literature, exploratory factor analysis was used to evaluate scale items. Exploratory factor analysis belongs to a family of factor analytic techniques which are used to define the underlying structure between variables in an analysis (Stewart, 1981). By examining intercorrelations among variables, factor analysis reduces the data to a more simplified structure thus representing the data in terms of distinct dimensions which are subsequently interpreted. There has been significant debate surrounding the appropriate role of factor analysis. Some researchers regard this technique as a purely exploratory tool that is useful only as a data reduction method, others take a more confirmatory stance where substantive theory is used to hypothesise the expected underlying structure of the data. As this research hopes to examine the underlying structure of the data, however, this stage of the research is regarded as exploratory in nature whereby the underlying structure of the data is to be detected.

Common Factor Analysis (Exploratory Factor Analysis) vs. Principal Components Analysis (PCA)

There are two general approaches to factor analysis, namely common factor analysis and principal components analysis (PCA). Principal components analysis is widely used and this can most likely be attributed to it being the default extraction method in both SPSS and SAS (Costello and Osborne, 2005). Holding this default position has more than

likely led to principal components analysis being mistakenly used when exploratory factor analysis is more suitable (Park et al, 2002). The goal of PCA is to reduce the measured variables to a smaller set of composite components that capture as much information as possible in as few components as possible. On the other hand, the goal of exploratory factor analysis (EFA) is to find the latent structure of the dataset by uncovering common factors. Therefore, exploratory factor analysis accounts for shared variance which is an important distinction from PCA which accounts for total variance as it fundamentally means that EFA is more suitable when exploring underlying theoretical constructs. There has been much debate over which of these techniques is the true method of factor analysis, with some arguing in favour of exploratory factor analysis (Snook and Gorsuch, 1989; Gorsuch, 1990; Floyd and Widaman, 1995) while others argue that there is little difference between the two (Velicer and Jackson, 1990). Principal axis factoring, a type of EFA, is superior to principal components analysis as it analyses common variance only which is a requirement when developing new scales. In addition to this, it is a useful technique for identifying items that do not measure an intended factor or that simultaneously measure multiple factors (Worthington and Whittaker, 2006). For these reasons and also because a proposed theoretical structure is expected to underlie the dataset, exploratory techniques are the most important for scale development procedures.

Principal Axis Factoring: Testing the Assumptions

Prior to conducting any form of factor analysis it is necessary to assess the factorabilty of the data. Issues that need to be considered are: whether the sample size is large enough to

permit factor analysis, whether the inter-correlations between variables are sufficiently correlated and to determine whether the data can be considered multivariate normal (Tabachnick and Fidell, 2007). As discussed above, varying opinions seem to exist on the recommended sample size for factor analysis. For instance, Tabachnik and Fidell (2007, p. 613) state that 'it is comforting to have at least 300 cases for factor analysis' however they do acknowledge that a sample size of 150 cases is sufficient when there are a number of very high loading items. A sample size of 205 was achieved here which meets the minimum sample size of 150 criterion proposed by Guadagnoli and Velicer (1988).

Inter-correlations among the items can firstly be evaluated using the correlation matrix. Given the large number of items (47), the correlation matrix was too large to include here. However, upon inspection of the items it became clear that a considerable number of inter-correlations greater than .3 exist which indicates that factor analysis is appropriate (Tabachnick and Fidell, 2007). Both Bartlett's test of sphericity and the Kaiser-Meyer-Olkin measure of sampling adequacy were also examined to determine the factorability of the data. Bartlett's test of sphericity examines the correlation matrix and uses the chi square statistic to test the significance of the inter-correlations in the matrix. The null hypothesis is that all variables are noncollinear, therefore a p value less than .05 indicates that factor analysis is appropriate. This criterion was met here. The Kaiser-Meyer-Olkin ranges from 0 to 1 with higher values indicating the suitability of the data to factor analysis. A KMO statistic of 0.875 was achieved here which is considered a

substantial figure and indicates that factor analysis is again appropriate here (Tabachnick and Fidell, 2007).

Deviations from univariate normality were also assessed by examining the skewness and kurtosis of the data. Although a number of items violated assumptions associated with distributional normality (Tabachnick and Fidell, 2007), no decisions were made at this point with regards to these items. All statistics for skewness and kurtosis can be found in Appendix 2D.

Principal Axis Factoring: Method of Extraction

Upon deciding that exploratory factor analysis was the most appropriate form of factor analysis for the current situation, the next decision pertained to which of the exploratory factor analytic techniques is the most suitable. The two most commonly used exploratory techniques are Maximum Likelihood and Principal Axis Factoring (often referred to as Principal Factor Analysis). Maximum Likelihood is advantageous in that it provides fit indices which allow for significance testing of the factor loadings (Fabrigar et al, 1999). However, this method of extraction has strict distributional assumptions and is not recommended for use when multivariate normality is violated which is often the case. Principal Axis Factoring is an alternative to maximum likelihood, to which less stringent assumptions are attached. Adjunct to having more lenient distributional assumptions, it is also less likely to generate Heywood cases or convergence problems (Fabrigar et al, 1999). Principal axis factoring explores relationships among items and is used to identify underlying factors and dimensions. This technique is appropriate when the underlying

dimensions are unknown and the primary objective is to identify the latent dimensions or constructs represented in the original variables (Stewart, 1981). It is similar to other exploratory factor analysis techniques in that it is used to identify underlying dimensions of the measurement instrument (Floyd and Widaman, 1995). Unlike Maximum Likelihood, Principal Axis Factoring provides a limited number of fit statistics and confidence tests (Fabrigar et al, 1999). This is not considered a major disadvantage here, as scale development is functionally an exploratory technique for which there is no great need for fit statistics at this stage of the analysis.

Factor Rotation

The purpose of factor rotation is to clarify and simplify the data structure to achieve a more theoretically meaningful pattern (Tabachnick and Fidell, 2007) that facilitates interpretation of the factors (Lawley and Maxwell, 1962) The simplest and most common form of rotation is orthogonal factor rotation where the axes are maintained at 90 degrees (Stewart, 1981; Costello and Osborne, 2005). Varimax rotation or other such orthogonal techniques should only be used when the dimensions are expected to be independent of one another. In marketing and other social sciences it is often the case that some correlation is expected between the factors. In such situations oblique methods are advisable as they assume there is a central construct that links the dimensions to one another. Examples of oblique methods are: direct oblimin, quartimin, and promax. According to Fabrigar et al (1999) there is no widely preferred method of oblique rotation as all tend to produce similar results.

It is proposed here that the servicescape consists of a number of discrete factors that are inextricably linked to one another. On a more general level, given these interconnections between the dimensions it was proposed that oblique rotation methods would be most suitable and therefore Promax with Kaiser Normalization was chosen here (Netemeyer et al, 2003). A factor analysis was run using Varimax rotation (orthogonal rotation method) to determine whether there were any major differences between the two and interestingly the results were almost identical. This finding adds credence to the Promax results as it indicates that the solution is stable regardless of the technique used.

2.5.7 Optimise Scale Length

Communalities

The variance of an item consists of common, unique and error variance (Hair et al, 2006). Common variance is the variance that an item shares with other items. Unique variance is item specific whereas, error variance is random variance associated with a particular item. Communalities refer to the variance that is shared, or common with other items in the analysis. According to Fabrigar (1999, p. 274) 'variables with low reliability will have low communalities and thus should be avoided'. Using Fabrigar et al (1999) and Comrey's (1988) recommendations, communalities were inspected to determine which items were unreliable. Communalities less than .4 are deemed low and this was subsequently used as a cut off point here. The first run of factor analysis indicated that 8 items had communalities less than .40. These were then removed from the data-set and EFA was re-run. This iterative process was repeated three times until all variables with communalities over .40 remained. These stages can be seen in Table 2.4 below:

Table 2.4: Communality Assessment

Items Deleted in Stage One (Communalities less than 0.40)
The odour was unpleasant
I found the lighting excellent
There were appropriate reading materials to read in the waiting room
I found the employees at the dental surgery unhelpful
The room temperature was pleasant
The parking facilities were satisfactory
The surgery had a poorly organised layout
The other patients in the surgery were neat and well dressed
Items Deleted in Stage Two (Communalities less than 0.40)
I found the signs used (e.g. bathroom, enter, exit, helpful to me
I found the music at the surgery bothersome
I found the dental surgery easy to find
The surgery had more than enough space to be comfortable
Items Deleted in Stage Three (Communalities less than 0.40)
The brochures and the communications material were visually appealing
The exterior of the dentist's was appropriate
The surgery was very unhygienic
Items Deleted in Stage Four (Communalities less than 0.40)
The dental equipment was of poor quality
I found the physical facilities very unclean

When stage four of the analysis was completed, twenty-four items remained. Interestingly, all items that were negatively worded were deleted from the item pool. Also, many of the items which were found to be unstable have been found by previous authors to lack validity (cf. Hightower et al, 2002). The next decision related to the number of factors to extract.

Deciding on the Number of Factors to Extract

Once items with low communalities were eliminated from the analysis the next stage was to determine the number of factors to retain. How one should decide upon the number of factors to extract has created extensive controversy and misunderstanding in the literature (Stewart, 1981). According to Tabachnik and Fidell (2007), this stage should take quite an exploratory approach by experimenting with different numbers of factors until a satisfactory solution is found. The most well known and most popular method for the retention of factors is Kaiser's eigenvalue greater than 1 criterion (Stewart, 1981; Fabrigar et al, 1999). This rule specifies that all factors greater than one are retained for interpretation. This method offers the advantage of being easy to understand and is also the default method on most programs. Some argue that this method oversimplifies the situation and also has a tendency to overestimate the number of factors to retain (Zwick and Velicer, 1986). In fact, this method may lead to arbitrary decisions, for example it does not make sense to retain a factor with an eigenvalue of 1.01 and then to regard a factor with an eigenvalue of .99 as irrelevant (Ledesma and Pedro, 2007). A technique which overcomes some of the deficiencies inherent in Kaiser's approach is Cattell's Scree Test (Cattell and Vogelmann, 1977). The Scree Test presents the eigenvalues in descending order which are linked with a line. This graph is then scrutinised to determine the point at which the last significant break takes places. The point at which a large break, or elbow in the plot of roots is taken to indicate the point where factoring should stop.

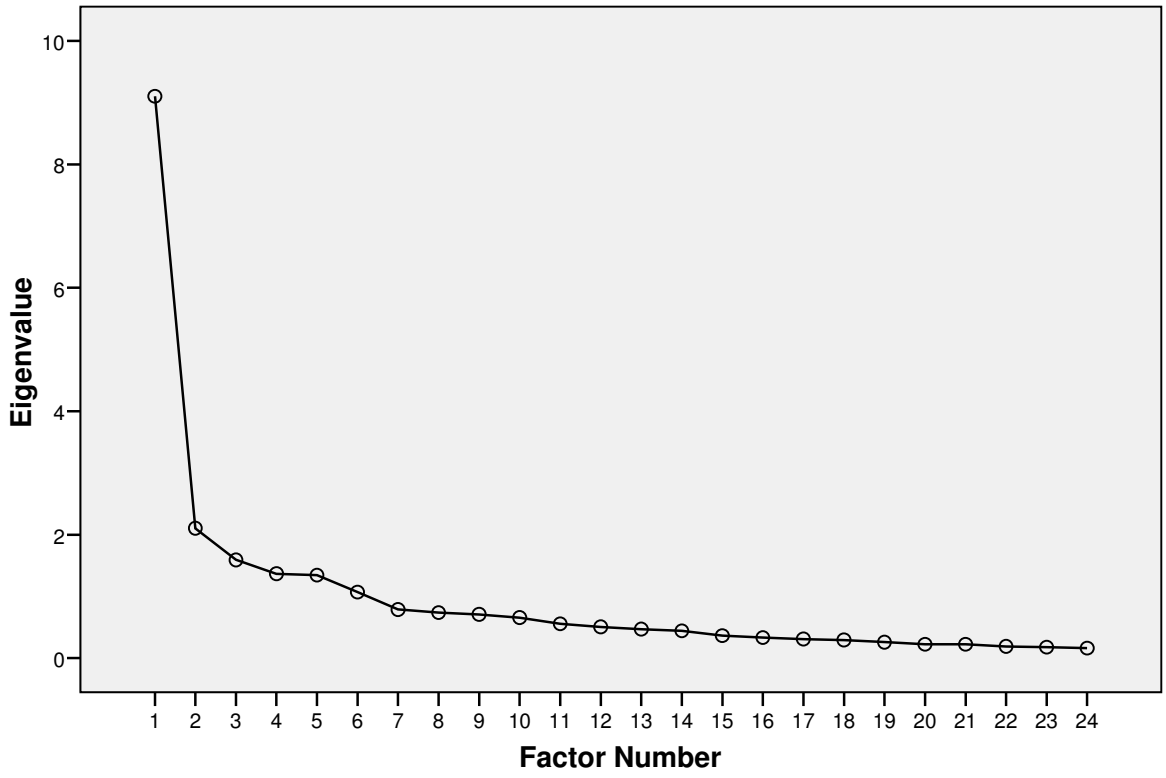


Figure 2.2: Eigenvalue Scree Plot

Figure 2.2 plots the eigenvalues for the current data and it can be seen that two distinct breaks exist in the eigenvalues. The first of these seems to occur between factor two and three while the second drop in eigenvalues seems to occur between factor six and seven. Based on theoretical assumptions proposed earlier in this chapter it was suspected that the six factor solution would be more in keeping with theory than a two factor solution. Both a two factor and six factor solution were produced and upon analysis of the two solutions it became immediately evident that a six factor solution was much more appropriate. The six factor solution is more advantageous, because firstly, it captures more variance (60%) than the two factor solution (42%). Secondly, the six factor solution is more or less in line with what was hypothesised theoretically. In addition, it is also interesting to note

that a six factor solution in also in keeping with the Kaiser's eigenvalue >1 criterion. The two factor solution can be found in Appendix 2E.

Further analysis using parallel analysis (Horn, 1965) and Velicer's (1976) MAP test were conducted to ensure the correct number of factors had been extracted. Using syntax provided by O'Connor (2000) these analyses were run and it was found that the parallel analysis suggested a three factor solution while Velicer's (1976) MAP indicated a two factor solution. As the two factor solution is not in keeping with theoretical expectations it was decided that the Scree Plot and eigenvalue greater than 1 criterion is more appropriate here, thus retaining the 6 factor solution. Interested readers can consult Appendix 2F for the statistical outputs associated with the parallel analysis and Velicer's (1976) MAP.

Factor Inspection and Interpretation

Following the decision to retain a six factor solution the next stage related to inspecting the factor loadings to discern whether there were any serious crossloadings existed. Fortunately no items crossloaded and an extremely tidy factor solution was produced. This final solution consisted of six factors which captured 60.1% of the variance which is considered a sizeable amount. The Promax Rotated Pattern Matrix can be viewed in Table 2.5 overleaf.

Table 2.5: Pattern Matrix

	Factors and Percentage of Variance					
	Décor & Design	Equipment & Materials	Space & Functionality	Cleanliness & Hygiene	Employees	Music
	36.3 %	7.2%	5.3%	4.5%	3.8%	3%
	1	2	3	4	5	6
The style of the interior was fashionable	**.836**	.123	-.301	-.118	-.006	.021
The interior layout was pleasing	**.780**	-.086	.098	.112	-.007	-.006
I found the physical facilities comfortable	**.772**	.102	.089	-.123	.113	-.116
The colour schemes were pleasant	**.723**	-.184	.138	.113	-.081	.168
The surgery had a pleasant smell	**.707**	.029	-.197	.122	-.016	.020
I found the interior design visually appealing	**.639**	-.001	-.069	-.066	-.026	.290
The waiting room was attractive	**.623**	-.059	.108	.088	-.008	.074
The architecture of the surgery was attractive	**.598**	.150	.032	-.140	.034	-.214
I found the lighting comfortable	**.474**	.030	.372	.034	-.128	-.130
The waiting room was comfortable	**.414**	-.021	.341	.046	.068	-.056
The electronic equipment was excellent	-.061	**.739**	.244	.039	-.059	.066
The electronic equipment was of high quality	-.037	**.722**	.010	.157	-.076	.060
The materials used were of high quality	.193	**.643**	.112	-.133	.036	.022
The dental equipment was modern looking	.054	**.488**	-.219	.402	-.089	.054
The bathroom facilities were clean and tidy	.135	**.479**	.049	.082	.177	-.047
The seating in the surgery was comfortable	.024	.052	**.761**	-.074	-.055	.056
The flooring was appropriate	-.150	.122	**.721**	.197	-.093	.013
I found my way around the surgery quite easily	-.087	.021	**.634**	-.182	.228	.045
I found the dental surgery very clean	-.031	.118	-.035	**.886**	.035	-.046
The employees were neat and tidy in appearance	.008	.056	.008	**.635**	.204	-.088
The employees were helpful	.075	.022	-.071	.052	**.844**	.048
I found the staff at the dental surgery friendly	-.089	-.088	.098	.137	**.789**	.040
The background music was pleasant	-.054	.136	.008	-.142	.061	**.968**
The background music was appropriate	.150	-.080	146	.132	.027	**.494**
Cronbach's alpha (α)	**.90**	**.85**	**.74**	**.79**	**.85**	**.73**

The ultimate goal of factor analysis is to identify the underlying constructs that summarise the variables in the data. Interpretation of the individual factors is a task through which this end is met. Interpreting rotated factor solutions is a subjective task (Stewart, 1981) and has been described by some as a 'black art', however one rule which reduces the subjectivity of the task is to analyse only factor loadings greater than .40 (Ford et al, 1986). Items were sorted according to their size to highlight the 'significant' items which can be thus used to define the factor. Significant items, or marker variables (Tabachnick and Fidell, 2007) are those with high factor loadings and should be used to guide the labelling process. This does not mean that weaker loading items are irrelevant, but that these stronger items should take precedence over lower loading items on the factor. Bearing both the stronger and weaker items in mind, the factors should then be interpreted 'in a way that will give a reasonable summary of the data' (Lawley and Maxwell, 1962, p. 219).

From Table 2.5 interpretation of the solution proved theoretically straightforward. While each item did not load exactly where expected, the overall solution was not dissimilar to the proposed dimensions of the physical environment. The first factor was named 'Décor and Design' as the aesthetics of the service seemed to dominate here. Factor two was named 'Equipment and Materials' given the presence of two equipment and one material item. The third factor was labelled 'Space and Functionality' given that items such as the flooring and layout fell on this factor. Naming the fourth factor was relatively straightforward as the two items both relate to the cleanliness of the environment and was therefore named 'cleanliness and hygiene'. The fifth factor also contained two items both

of which referred to the employees thus leading this factor to be named 'Employees'. Finally, the sixth factor contained two items and was simply named 'Music'.

The inter-factor correlation matrix was also examined to determine whether each of the factors were correlated with one another. The results from this can be seen in Table 2.6 below. Moderate to high inter-correlations seem to exist across the factors demonstrating that oblique rotation was indeed necessary in the current situation as many of the factors appear to be related to one another.

Table 2.6: Inter-Factor Correlation Matrix

Factor	1	2	3	4	5	6
1	1.000					
2	.514	1.000				
3	.616	.538	1.000			
4	.436	.484	.479	1.000		
5	.277	.294	.297	.200	1.000	
6	.403	.243	.316	.305	.092	1.000

In recent years, hierarchical factor analysis has been recommended to determine whether multiple higher order factors exist in a dataset (Comrey, 1988; Wolff and Preising, 2005). So as to ensure that only one higher order factor exists in the data hierarchical factor analysis was conducted (Wolff and Preising, 2005). Using the inter-factor correlation matrix in Table 2.6 as input it was found that only one higher order factor exists, thus corroborating that the servicescape is the sole construct that exists at a higher level of abstraction. The results from this can be found in Appendix 2G. Further to this, Schmid-Leiman's solution (SLS) was also run to determine the relative impact of the lower order servicescape sub-dimensions on the higher order construct (Schmid and

Leiman, 1957). The results found that the variance between the two levels was split almost equally between the two with the general factor explaining 49.5% of the variance and the first order factors accounting for 50.5% (results are shown in Appendix 2H). This highlights the importance of including all items in any further analysis as by disregarding those items that are not captured by one of the levels of generality may jeopardise our understanding of the servicescape. The SLS also found that no items had loadings less than .3 therefore providing further confidence in the final servicescape items.

2.6 Conclusions

From the above factor analytic study, a twenty-four item instrument of the servicescape was developed. It was enlightening to find that the dimensions underlying this scale generally support the structure proposed. The next stage in the research process is to use this scale to confirm and perhaps refine its structure in a field environment. In accordance with Netemeyer et al's (2003) recommendations, structural equation modelling in the form of confirmatory factor analysis and second-order factor modelling will be used. In the chapter that follows, an overview of structural equation modelling is provided as this is the technique of choice for all further analyses in this thesis.

3.0 STRUCTURAL EQUATION MODELLING

3.1 Introduction

One of the functions of this dissertation is to ascertain whether the servicescape conceptualisation proposed in Chapter One is an accurate depiction of a complicated construct. In addition to assessing the dimensionality of the servicescape, its relation to a number of post-consumption processes is also to be determined. In order to meet these objectives, structural equation modelling was chosen as the analytical tool. Structural equation modelling was chosen over more basic techniques such as multiple regression as it is a multivariate technique that permits multiple constructs to be examined simultaneously. From this point on, structural equation modelling forms the foundation of much of the analyses in this dissertation. This chapter serves as an introduction to structural equation modelling and the theory surrounding it. In addition to this, techniques for dealing with missing data in structural equation modelling are also discussed.

3.2 Theoretical Underpinnings of Structural Equation Modelling

Structural Equation Modelling (SEM), also known as path analysis with latent variables (McDonald and Ho, 2002), or covariance structure modelling (Bollen, 1989) is a result of the convergence of psychometrics, econometrics and biometrics (Bentler, 1980). In the early twentieth century, Spearman (1904) was the first to develop latent variables and errors in variables, but it was not until thirty years ago that Jöreskog, Keesling and Wiley provided a covariance-based general model that integrated regression and factor analysis

(Bentler, 1980; Chin, 1998). This breakthrough thus allowed researchers to deal with more complex social science models (Bentler, 1980; MacKenzie, 2001) that were previously not possible to examine.

SEM achieves this by assessing whether a sample covariance matrix is consistent with a hypothesised model (Rigdon, 1998). This model consists of a series of structural equations which are simultaneously examined as a series of highly restricted linear regression equations that depict all the relationships between constructs. A construct, or latent variable is not directly observable, but is inferred indirectly through observed/manifest variables which are often called indicators (Kline, 2005). Relationships between indicators and latent variables and between latent variables are visually represented as path diagrams. Although path diagrams are not essential to structural equation modelling, they provide an overall view of the model's structure which clarifies the system of hypotheses to the researcher (Diamantopoulos and Siguaw, 2000). Within path diagrams, latent variables are represented as a circle or an oval, and indicators are enclosed in rectangles or squares (Byrne, 1998). Latent variables that are determined by constructs outside the model are referred to as exogenous constructs (Anderson and Gerbing, 1988; McDonald and Ho, 2002) and these are linked to other exogenous constructs with double-ended arrows as they are assumed to covary (Kline, 2005). Exogenous constructs are linked to endogenous constructs (multi-item equivalent of dependent variables) by straight arrows (→). These arrows illustrate which construct is an antecedent and which is an outcome and is referred to as a dependence relationship. Greek characters are the dominant symbolic language in SEM and have been added to the

path diagram in Figure 3.1 below. A complete discussion of the meaning of each of these characters in provided in Appendix 3A.

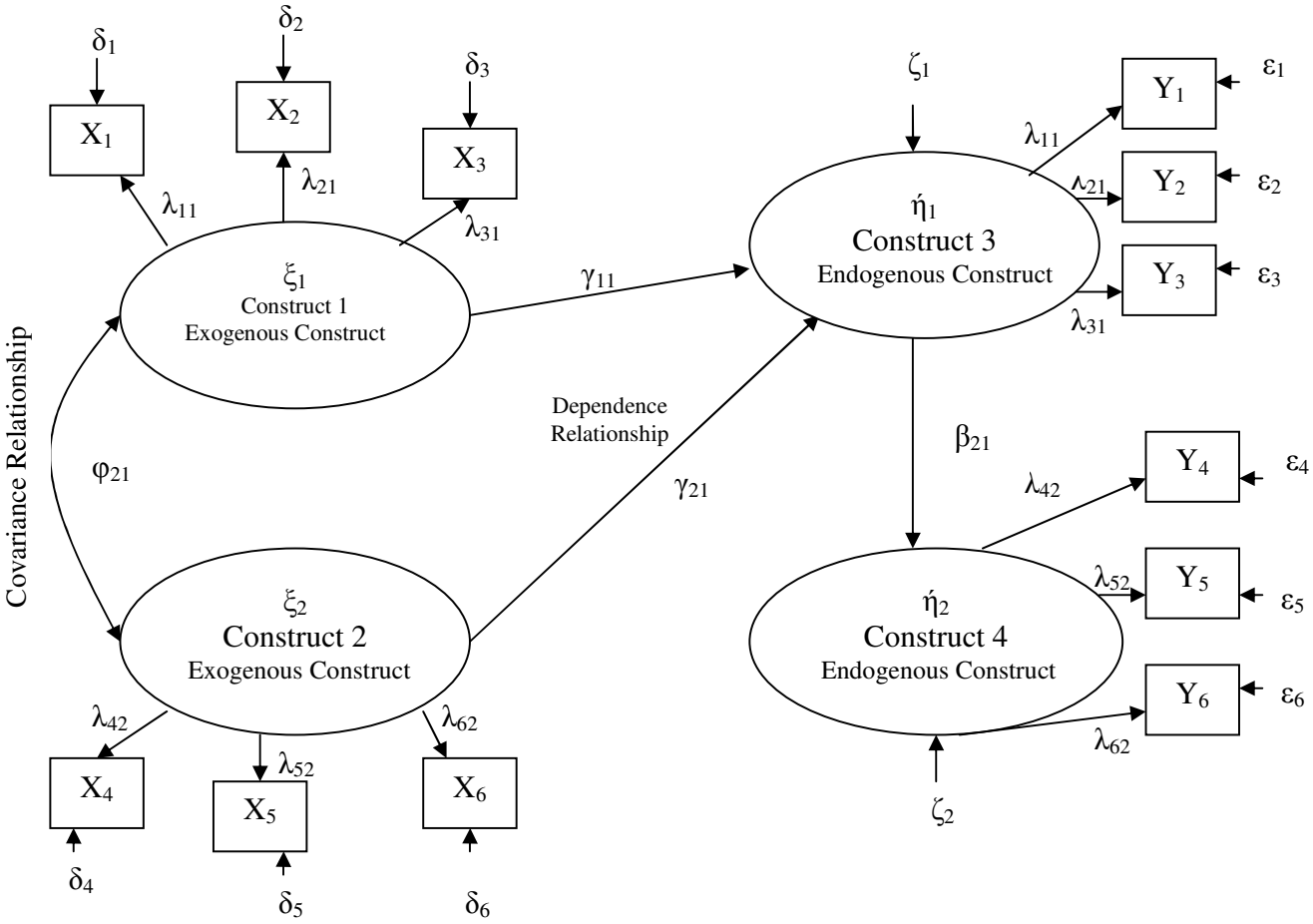


Figure 3.1: Structural Equation Model Example

3.3 SEM vs. Alternative Techniques

An alternative analysis technique for the current research problem would be to utilise multiple regression analysis, however multiple regression was not chosen for a number of reasons. Firstly, in SEM the relationships in the model are free from error as the error has been estimated and removed entirely from the analysis (Ullman, 2007). Additionally,

SEM is a highly flexible technique as it allows researchers to estimate multiple and interrelated dependence relationships such as those specified in Figure 3.1 (Chin, 1998). It also permits comparison of a given model across different populations and is referred to as multigroup structural equation modelling (Rigdon, 1998). In multiple regression, multicollinearity of independent variables causes problems when the matrix of correlations is inverted (Ullman, 2007). Quite the opposite is the case for SEM as it requires a high degree of multicollinearity in order to improve the stability of the measurement model (Rigdon, 1998). A final advantage of SEM is the convenient and powerful graphical language that has been developed which allows for complex relationships to be presented. This system allows each relationship in the model to be succinctly presented in equation format (Rigdon, 1998; Diamantopoulos and Siguaw, 2000). The structural model in Figure 3.1 can thus be represented as fourteen equations:

Measurement Equations for Endogenous Variables:

$$y_1 = \lambda_{11}\,\acute{\eta}_1 + \varepsilon_1$$
$$y_2 = \lambda_{21}\,\acute{\eta}_1 + \varepsilon_2$$
$$y_3 = \lambda_{31}\,\acute{\eta}_1 + \varepsilon_3$$
$$y_4 = \lambda_{42}\,\acute{\eta}_2 + \varepsilon_4$$
$$y_5 = \lambda_{52}\,\acute{\eta}_2 + \varepsilon_5$$
$$y_6 = \lambda_{62}\,\acute{\eta}_2 + \varepsilon_6$$

Measurement Equations for Exogenous Variables:

$$x_1 = \lambda_{11}\,\xi_1 + \delta_1$$
$$x_2 = \lambda_{21}\,\xi_1 + \delta_2$$
$$x_3 = \lambda_{31}\,\xi_1 + \delta_3$$
$$x_4 = \lambda_{42}\,\xi_2 + \delta_4$$
$$x_5 = \lambda_{52}\,\xi_2 + \delta_5$$
$$x_6 = \lambda_{62}\,\xi_2 + \delta_6$$

Structural Equations:

$$\acute{\eta}_1 = \gamma_{11}\xi_1 + \gamma_{21}\xi_2 + \zeta_1$$
$$\acute{\eta}_2 = \beta_{21}\acute{\eta}_1 + \zeta_2$$

An additional alternative to structural equation modelling is path analysis. Path analysis has been referred to as 'less interesting' or facile as it only allows the researcher to use a single indicator per construct (Kline, 2005). The major problem with this is that the

exogenous variables are measured without error, an assumption which is unlikely to be true (Grapentine, 2000) as all measures contain at some error (Tabachnick and Fidell, 2007). While path analysis and structural equation modelling will almost always provide very similar results, structural equation modelling is preferred for a number of reasons. Firstly, structural equation modelling is thought to be more accurate as it takes into account measurement error. Secondly, as latent variables with multiple indicators are used, it is possible to determine which of the items are the best indicators of the unobserved latent variable (Grapentine, 2000).

3.4 Structural Equation Modelling Software

To date there are eight SEM computer programs: CALIS, EQS, Mplus, Mx Graph, RAMONA, SEPATH, AMOS and LISREL (Kline, 2005). The first six of these are less popular than AMOS and LISREL. A possible reason for this could be that they are generally less user-friendly than AMOS and LISREL. AMOS is distributed by SPSS and offers the distinct advantage of having a graphical interface that allows the users to draw their specified model. The disadvantage of this however, is that for large, complex models this becomes quite a cumbersome task. The final application is LISREL (Linear Structural Relationships), a program which has become synonymous with structural equation modelling in many circles. The reason for this being that the forefathers of structural equation modelling, Jöreskog and Sörbom were the developers of this program. LISREL is an integrated suite of programs that provides a researcher with tools for data entry and exploratory analysis and a wide range of structural equation model analyses. Included in this is PRELIS which allows researchers to conduct a number of pre-

processing analyses such as bootstrapping, transformations of variables and multilevel analyses to name a few. For conducting structural equation modelling analysis, LISREL provides users with a choice between the original LISREL syntax or their more user-friendly SIMPLIS syntax. The SIMPLIS syntax is not based on Greek characters, rather it requires 'naming the observed and latent variables and specifying paths with equation-type statements' (Kline, 2005, p. 84) and is the type of programming to be used in this dissertation. On the whole, the features and capabilities of LISREL are comprehensive and has thus remained an industry standard.

3.5 Estimation Methods

Early applications of SEM used ordinary least squares (OLS) regression for model estimation. In later years OLS was replaced with Maximum Likelihood Estimation (MLE) as it was found to be more efficient and unbiased in situations were multivariate normality was met. MLE was developed by Jöreskog, Keesling and Wiley (JKW) and is the most widely used application of structural equation modelling (Bentler, 1980). As with other estimation procedures, it is based upon covariance matrix algebra (Chin, 1998) and uses an iterative process whereby it attempts to minimise the difference between the sample covariances and those specified by the researcher in the theoretical model. Maximum Likelihood Estimation has been criticised for its oversensitivity to nonnormality which thus spurred a number of alternative techniques such as weighted least squares (WLS, generalised least squares (GLS) and asymptotically distribution free (ADF) to be developed. In a study comparing ADF to MLE, Benson and Fleishman (1994) found that MLE showed greater chi square bias than ADF in situations of non-

normality. On the other hand, it was found that MLE performed better in large sample situations. This is particularly problematic since ADF typically requires very large samples. For instance, relatively simple models need sample sizes of 200-500 which is quite an impractical requirement for many social scientists (Kline, 2005). Similar simulation studies have been conducted by Olsson et al (2000; 2004) in which MLE was compared with GLS and WLS it was found that MLE provided more realistic indexes of overall fit and less biased parameter values. These studies have held MLE as the default estimator in most SEM programs and have led to it becoming the standard method of estimating free parameters in structural equation models. For these reasons, MLE was the estimation method chosen here.

3.6 Assumptions Associated with SEM

As with all multivariate analyses, there are a number of assumptions that must be met prior to conducting any analysis. The first of these is that the data is multivariate normally distributed. A second basic assumption in SEM is that all relations are linear (Kline, 2005). It is recommended that scatterplots are inspected for obviously curvilinearity in the data. SEM also has a number of assumptions associated with sample size. Similar to exploratory factor analysis, structural equation modelling is thought to be a large sample technique. Due to this, one of the key questions researchers are faced with is how large a sample size is needed and has been regarded as one of the most deceptively difficult questions to answer (Jackson, 2003). One suggested approach is based upon the number of parameters, with higher sample size to parameter ratio required. Recommendations have ranged from as high as 10:1 to as low as 5:1 (Kline,

2005). Minimum sample size estimates have also been proffered and it is thought that samples under 100 are unsound. A final assumption that must be met relates to missing data. There are an extensive number of issues a researcher must consider when dealing with missing data and for this reason missing data is discussed in detail below.

3.6.1 Missing Data

A key assumption of SEM is that it requires a complete data set. Unfortunately social science data is often hampered with this problem. To cope with missing data, statisticians have provided a number of methods for handling non-response. These include, listwise deletion, pairwise deletion and maximum likelihood approaches. Up until recently, listwise deletion, pairwise deletion and mean substitution were the most popular, however maximum likelihood methods have gained popularity given that many SEM packages provide more advanced imputation methods in their programs. Before deciding on the most appropriate way of handling missing data, the researcher needs to determine the nature of the missing data. Thus, missing data is classified as either missing at random (MAR) or missing at completely at random (MCAR). If an individual does not respond to a particular question, and the lack of response is thought to be related to another question in the study then the data are considered MAR. An example of MAR would be where older people will not provide answers to a salary question. On the other hand, MCAR assumes that a subjects' non-response is not dependent on any other question in the study or the question itself. It is important that a researcher explores their data before deciding on the most appropriate method of handling missing responses.

Next, the various missing data methods are discussed along with their assumptions. Finally the missing data method used here will be presented.

If a case is missing data on some variables in the analysis, then listwise deletion involves discarding the entire response set for that case. The main virtue of this method is its simplicity (Schafer and Graham, 2002). However, the prevalence of listwise deletions has lessened in recent years due to a number of serious inefficiencies. Firstly, listwise deletion can lead to bias as the researcher has assumed that those cases that are deleted are the same as those that contain complete data (Malhotra, 1987; Schafer and Olsen, 1998). As well as this problem, it can also lead to large amounts of data being deleted which can thus result in inefficient parameter estimates (Enders and Bandalos, 2001). When data is Missing Completely At Random (MCAR), it has been found that listwise deletion will lead to unbiased parameter estimates and can thus be used for SEM.

With pairwise deletion, cases are omitted from any calculations involving variables for which they have missing data (Enders and Bandalos, 2001). Pairwise deletion is no longer a popular method as it can produce a covariance matrix that cannot be inverted which is an essential step in structural equation modelling (Acock, 2005). The final less sophisticated missing data method is mean substitution. Mean substitution involves calculating the average response from the available data and replacing missing values with this mean value (Tabachnick and Fidell, 2007). Research has found that this method flattens distributions, increases variance (Schafer and Graham, 2002), underestimates correlations and β weights (Acock, 2005) and redefines the scale (Schafer and Graham,

2002). Due to these serious drawbacks, authors have strongly objected to mean substitution. In fact, Pallant (2004, p. 119) is so adamant about this that she states that mean substitution 'should NEVER be used as it can severely distort the results of your analysis' (original emphasis).

More recent missing data procedures have become much more advanced and are available in many software programs such as SPSS, SAS and PRELIS (Acock, 2005). These methods are based on maximum-likelihood procedures and are thus more sophisticated. These methods are advantageous in that they only assume the less demanding assumption of MAR. The general premise of these methods is that they examine the pattern of responses and thus substitute values based on the pattern of responses obtained from complete responses.

One such maximum likelihood method available in LISREL is the Expectation-Maximisation (EM) Algorithm. This procedure was first introduced by Dempster et al (1977) and heralded a fundamental shift away from the way statisticians viewed data. Through the EM algorithm, researchers can use observed values as a basis for making assumptions about unobserved ones. EM forms a covariance matrix by assuming the distribution of the missing data cases and is an iterative procedure with two steps: expectation and maximisation (Tabachnick and Fidell, 2007). The first step is the expectation stage where the conditional expectation of the missing data is found. These expectations are then replaced with the missing data. The second stage, maximisation

performs the maximum likelihood estimation and then when convergence is achieved the EM variance-covariance matrix can be saved as a separate data set.

For this research, the dataset was first explored to determine whether a pattern existed in the missing data. This is done using Little's MCAR test (Tabachnick and Fidell, 2007). Little's test examines the pattern of the missing data and provides a statistic, which if significant indicates that the data is missing completely at random. For the current dataset a Little's statistic of 0.232 was found, signifying that the missing data pattern was indeed MCAR. Having a dataset where the missing data falls into the MCAR category is quite advantageous as it allows you to choose whichever missing data treatment you wish. The EM algorithm was chosen over more elementary techniques such as listwise and pairwise deletion, given its sophistication and ease of use in SPSS.

3.5 Model Fit

Once all missing data is dealt with, the next stage in the research problem is to run the specified model and to assess its fit. Assessing whether a specified model 'fits' the data is one of the most important steps in structural equation modelling (Yuan, 2005) and is ultimately one of its primary goals (Hu and Bentler, 1999) as it determines whether the model being tested should be accepted or rejected. Due to this, an abundance of indices have been developed as measures of the goodness, or badness of fit (McDonald and Ho, 2002). These fit indices can be classified as either absolute fit indices or incremental fit indices (Hu and Bentler, 1999). Absolute fit indices determine how well an *a priori* model fits the sample data (McDonald and Ho, 2002) whereas incremental fit indices

compare the model tested to a null model. Beginning with absolute fit indices, commonly used indices from each of these index families will be discussed next.

3.5.1 Absolute Fit Indices

Absolute fit indices determine how well an *a priori* model fits the sample data (McDonald and Ho, 2002) and demonstrates which proposed model has the most superior fit. These measures provide the most fundamental indication of how well the proposed theory fits the data. Unlike incremental fit indices, their calculation does not rely on comparison with a baseline model but is instead a measure of how well the model fits in comparison to no model at all (Jöreskog and Sorböm, 1993). Included in this category are the Chi-Squared test, Root Mean Squared Error of Approximation (RMSEA), the Goodness of Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), the Root Mean Square Residual (RMR) and the Standardised Root Mean Square Residual (SRMR).

Model Chi-Square (χ^2)

The chi square value is the traditional measure for evaluating overall model fit and, 'assesses the magnitude of discrepancy between the sample and fitted covariances matrices' (Hu and Bentler, 1999, p. 2). A good model fit would provide an insignificant result at a 0.05 threshold (Barrett, 2007), thus the chi-square statistic is often referred to as either a 'badness of fit' (Kline, 2005) or a 'lack of fit' (Mulaik et al, 1989) measure. While the Chi-Squared test retains its popularity as a fit statistic, there exist a number of severe limitations in its use. Firstly, this test assumes multivariate normality and severe deviations from normality may result in model rejections even when the model is

properly specified (McIntosh, 2006). Secondly, because the chi square statistic is in essence a statistical significance test it is sensitive to sample size which means that the chi square statistic nearly always rejects the model when large samples are used (Bentler and Bonnet, 1980; Jöreskog and Sörbom, 1993). On the other hand, where small samples are used, the chi-square statistic lacks power and because of this may not discriminate between good fitting models and poor fitting models (Kenny and McCoach, 2003). Due to the restrictiveness of the model chi square, researchers have sought alternative indices to assess model fit. One example of a statistic that minimises the impact of sample size on the model chi square is Wheaton et al's (1977) relative/normed chi-square (χ^2/df). Although there is no consensus regarding an acceptable ratio for this statistic, recommendations range from as high as 5.0 (Wheaton et al, 1977) to as low as 2.0 (Tabachnick and Fidell, 2007).

Root Mean Square Error of Approximation (RMSEA)

The RMSEA is the second fit statistic reported in the LISREL program and was first developed by Steiger and Lind (1980; cited in Steiger, 1990). The RMSEA tells us how well the model, with unknown but optimally chosen parameter estimates would fit the populations covariance matrix (Byrne, 1998). In recent years it has become regarded as 'one of the most informative fit indices' (Diamantopoulos and Siguaw, 2000, p. 85) due to its sensitivity to the number of estimated parameters in the model. In other words, the RMSEA favours parsimony in that it will choose the model with the lesser number of parameters. Recommendations for RMSEA cut-off points have been reduced considerably in the last fifteen years. Up until the early nineties, an RMSEA in the range

of 0.05 to 0.10 was considered an indication of fair fit and values above 0.10 indicated poor fit (MacCallum et al, 1996). It was then thought that an RMSEA of between 0.08 to 0.10 provides a mediocre fit and below 0.08 shows a good fit (MacCallum et al, 1996). However, more recently, a cut-off value close to .06 (Hu and Bentler, 1999) or a stringent upper limit of 0.07 (Steiger, 2007) seems to be the general consensus amongst authorities in this area.

One of the greatest advantages of the RMSEA is its ability for a confidence interval to be calculated around its value (MacCallum, et al 1996). This is possible due to the known distribution values of the statistic and subsequently allows for the null hypothesis (poor fit) to be tested more precisely (McQuitty, 2004). It is generally reported in conjunction with the RMSEA and in a well-fitting model the lower limit is close to 0 while the upper limit should be less than 0.08.

Goodness-of-Fit Statistic (GFI) and the Adjusted Goodness-of-Fit Statistic (AGFI)

The goodness-of-fit statistic (GFI) was created by Jöreskog and Sörbom as an alternative to the chi square test and index calculates the proportion of variance that is accounted for by the estimated population covariance (Tabachnick and Fidell, 2007). By looking at the variances and covariances accounted for by the model it shows how closely the model comes to replicating the observed covariance matrix (Diamantopoulos and Siguaw, 2006). This statistic ranges from 0 to 1 with larger samples increasing its value. When there are a large number of degrees of freedom in comparison to sample size, the GFI has a downward bias (Sharma et al, 2005). In addition, it has also been found that the GFI

increases as the number of parameters increases (MacCallum and Hong, 1997) and also has an upward bias with large samples (Bollen, 1990; Miles and Shevlin, 1998). Traditionally an omnibus cut-off point of 0.90 has been recommended for the GFI however, simulation studies have shown that when factor loadings and sample sizes are low a higher cut-off of 0.95 is more appropriate (Miles and Shevlin, 1998). Given the sensitivity of this index, it has become less popular in recent years and it has even been recommended that this index should not be used (Sharma et al, 2005). Related to the GFI is the AGFI which adjusts the GFI based upon degrees of freedom, with more saturated models improving fit (Tabachnick and Fidell, 2007). Thus, more parsimonious models are preferred while penalised for complicated models. In addition to this, AGFI tends to increase with sample size. As with the GFI, values for the AGFI also range between 0 and 1 and it is generally accepted that values of 0.90 or greater indicate well fitting models. Given the often detrimental effect of sample size on these two fit indices they are not relied upon as stand alone indices, however given their historical importance they are often reported in covariance structure analyses.

Root Mean Square Residual (RMR) and Standardised Root Mean Square Residual (SRMR)

The RMR and the SRMR are the square root of the difference between the residuals of the sample covariance matrix and the hypothesised covariance model. The range of the RMR is calculated based upon the scales of each indicator, therefore, if a questionnaire contains items with varying levels (some items may range from 1 – 5 while others range from 1 – 7) the RMR becomes difficult to interpret (Kline, 2005). The standardised

RMR (SRMR) resolves this problem and is therefore much more meaningful to interpret. Values for the SRMR range from zero to 1.0 with well fitting models obtaining values less than .05 (Byrne, 1998; Diamantopoulos and Siguaw, 2000), however values as high as 0.08 are deemed acceptable (Hu and Bentler, 1999). An SRMR of 0 indicates perfect fit but it must be noted that SRMR will be lower when there is a high number of parameters in the model and in models based on large sample sizes.

3.5.2 Incremental Fit Indices

Incremental fit indices, also known as comparative (Miles and Shevlin, 2007) or relative fit indices (McDonald and Ho, 2002), are a group of indices that do not use the chi-square in its raw form but compare the chi-square value to a baseline model. For these models the null hypothesis is that all variables are uncorrelated (McDonald and Ho, 2002).

Normed-Fit Index (NFI)

The first of these indices to appear in LISREL output is the Normed Fit Index (NFI: Bentler and Bonnet, 1980). This statistic assesses the model by comparing the χ^2 value of the model to the χ^2 of the null model. The null/independence model is the worst case scenario as it specifies that all measured variables are uncorrelated. Values for this statistic range between 0 and 1 with Bentler and Bonnet (1980) recommending values greater than 0.90 indicating a good fit. More recent suggestions state that the cut-off criteria should be NFI $\geq$ .95 (Hu and Bentler, 1999). A major drawback to this index is that it is sensitive to sample size, underestimating fit for samples less than 200 in size (Mulaik et al, 1989; Bentler, 1990), and is thus not recommended to be solely relied on

(Kline, 2005). This problem was rectified by the Non-Normed Fit Index (NNFI, also known as the Tucker-Lewis index), an index that prefers simpler models. However in situations were small samples are used, the value of the NNFI can indicate poor fit despite other statistics pointing towards good fit (Bentler, 1990; Kline, 2005; Tabachnick and Fidell, 2007). A final problem with the NNFI is that due to its non-normed nature, values can go above 1.0 and can thus be difficult to interpret (Byrne, 1998). Recommendations as low as 0.80 as a cutoff have been proferred however Bentler and Hu (1999) have suggested NNFI ≥ 0.95 as the threshold.

CFI (Comparative Fit Index)

The Comparative Fit Index (CFI: Bentler, 1990) is a revised form of the NFI which takes into account sample size (Byrne, 1998), performing well even when sample size is small (Tabachnick and Fidell, 2007). This index was first introduced by Bentler (1990) and subsequently included as part of the fit indices in his EQS program (Kline, 2005). Like the NFI this statistic assumes that all latent variables are uncorrelated (null/independence model) and compares the sample covariance matrix with this null model. As with the NFI, values for this statistic range between 0.0 and 1.0 with values closer to 1.0 indicating good fit. A cut-off criterion of ≥ 0.90 was initially advanced however, recent studies have shown that a value greater than 0.90 is needed in order to ensure that misspecified models are not accepted (Hu and Bentler, 1999). From this, a value of ≥ 0.95 is presently recognised as indicative of good fit (Hu and Bentler, 1999). Today this index is included in all SEM programs and is one of the most popularly reported fit index due to being one of the measures least effected by sample size (Fan et al, 1999).

3.5.3 Parsimony Fit Indices

Having a nearly saturated, complex model means that the estimation process is dependent on the sample data. This results in a less rigorous theoretical model that paradoxically produces better fit indices (Crowley and Fan, 1997; Mulaik et al, 1989). To overcome this problem, Mulaik et al (1989) have developed two parsimony of fit indices; the Parsimony Goodness-of-Fit Index (PGFI) and the Parsimonious Normed Fit Index (PNFI). The PGFI is based upon the GFI by adjusting for loss of degrees of freedom. The PNFI also adjusts for degrees of freedom however it is based on the NFI (Mulaik et al, 1989). Both of these indices seriously penalise for model complexity which results in parsimony fit index values that are considerably lower other goodness of fit indices. While no threshold levels have been recommended for these indices, Mulaik et al (1989) do note that it is possible to obtain parsimony fit indices within the .50 region while other goodness of fit indices achieve values over .90 (Mulaik et al, 1989). The authors strongly recommend the use of parsimony fit indices in tandem with other measures of goodness-of-fit however, because no threshold levels for these statistics have been recommended it has made them more difficult to interpret.

3.6 Reporting Fit Indices

With regards to which indices should be reported, it is not necessary or realistic to include every index included in the program's output as it will burden a reader. Given the plethora of fit indices, it becomes a temptation to choose those fit indices that indicate the best fit. This should be avoided at all costs as it is essentially sweeping important

information under the carpet. In a review by McDonald and Ho (2002) it was found that the most commonly reported fit indices are the CFI, GFI, NFI and the NNFI. When deciding what indices to report, going by what is most frequently used is not necessarily good practice as some of these statistics (such as the GFI discussed above) are often relied on purely for historical reasons, rather than for their sophistication. While there are no golden rules for assessment of model fit, reporting a variety of indices is necessary (Crowley and Fan, 1997). Multiple fit indices need to be reported as each index reflects a different aspect of model fit. Although the chi square has many problems associated with it, it is still essential that this statistic, along with its degrees of freedom and associated *p* value, should at all times be reported (Hayduk et al, 2007; Kline, 2005). Threshold levels were recently assessed by Hu and Bentler (1999) who suggested a two-index presentation format. This always includes the SRMR with either the NNFI (TLI), the RNI and the CFI. The various combinations are summarised in Table 3.1 below.

Table 3.1: Hu and Bentler's Two-Index Presentation Strategy (1999)

Fit Index Combination	**Combinational Rules**
NNFI (TLI) and SRMR	NNFI of 0.95 or higher and an SRMR of .09 or lower or NNFI of 0.96 and SRMR of 0.09
RMSEA and SRMR	RMSEA of 0.06 or lower and a SRMR of 0.09 or lower

Kline (2005) speaks strongly about which indices to include and advocates the use of the chi-square test, the RMSEA, the CFI and the SRMR. Boomsma (2005) has similar recommendations but also advises for the squared multiple correlations of each equation

to be reported. These indices have been chosen over other indices as they have been found to be the most insensitive to sample size, model misspecification and parameter estimates. A summary of all fit indices and their recommended threshold levels can be found in Appendix 3B.

3.7 Conclusions

This chapter has provided a comprehensive presentation of the theoretical foundations of structural modelling as well as providing guidelines on the most appropriate way to deal with missing cases. From this point on in the dissertation much of the analyses employs structural equation modelling and will draw on this chapter when making decisions in relation to the technique.

4.0 ASSESSING THE DIMENSIONALITY OF THE SERVICESCAPE

4.1 Introduction

It is evident from Chapter One that the servicescape consists of several conceptual dimensions (Bitner, 1992; Baker et al, 2002; Ezeh and Harris, 2007). Conceptual definitions of the servicescape clearly regard the servicescape as a multi-faceted phenomenon, however in actual practice the servicescape is often modelled as a single construct (Grace and O'Cass, 2004; Hightower et al, 2002; Reimer and Kuehn, 2005; Wakefield and Blodgett, 1994; Wakefield and Blodgett, 1999). Modelling the servicescape in this way results in an incomplete understanding of its consequences thus requiring a multi-dimensional conceptualisation in which each sub-dimension can be examined in relation to other constructs in a model. Multidimensional models are certainly an advancement on these unidimensional studies with both theoretical (Baker, 1987; Bitner, 1992) and empirical (Wakefield and Blodgett, 1996) studies appearing in the literature over the years. However, as discussed in earlier chapters, there appears to be no consistency as to what constitutes the servicescape with huge discrepancies across each individual study. Previous to this, none of these studies have sought to directly estimate the superiority of one model over another, instead authors seem to randomly decide what encompasses the servicescape construct based on the particular research problem at hand. This has led to a need for an evaluation of existing models in the literature to determine whether certain models are superior to others. This chapter addresses this issue by bringing together the literature in this area and determines whether the model proposed in Chapter One is a more accurate representation of the servicescape

construct than other conceptualisations such as Bitner (1992) and Baker's (1987; 2002). This can be achieved through second order confirmatory factor analysis which permits model dimensionality to be assessed. In the sections that follow, confirmatory factor analysis is used to examine the dimensionality of the servicescape. Before beginning the analysis the proposed measurement model is presented along with a rationale for the choice of service. Following this, decisions relating to the mode of administration, format of the questionnaire and interviewing techniques are presented.

4.2 Proposed Measurement Model

In deriving the measurement model, a number of theoretical conceptualisations were brought together. Previous research by Bitner (1992), Baker (2002) and Wakefield and Blodgett (1994; 1996; 1998; 1999) formed the backbone of this model. In terms of both the number and content of their dimensions these authors vary widely. For instance, Bitner's (1992) servicescape representation consists of three dimensions; 'ambient conditions', 'space and function', and 'signs, symbols and artifacts'. Within the 'space and function' dimension there exist a number of sub-categories which include layout, furnishings and equipment. For the 'signs, symbols and artifacts' dimension the sub-categories consist of signage and style of décor. Baker (1987; 1994; 2002) also proposes a three factor structure however its content is quite dissimilar to Bitner's (1992; 2002). Firstly, in terms of similarity, both Baker (1987) and Bitner (1992) agree that ambient cues form their own unique construct. However, the remaining two of Baker's constructs consist of a general 'design' dimension and a 'social factors' dimension. Baker's design construct is an all encompassing dimension that includes stimuli that are 'at the forefront

of our awareness' (Baker, 1987, p. 79). Design elements are then further sub-divided into functional and comfort dimensions. Functional elements are described as including items relating to layout whereas comfort dimensions relate to the style in which the store is presented. Baker's (1987) third dimension is the people component which consists of both the customers in the service environment and the personnel. According to Baker (1987) issues that can impact on behaviour are the number of customers and personnel and also their appearance and behaviour.

Wakefield and Blodgett (1994; 1996; 1999) have also strongly influenced the proposed servicescape schema as they have extensively researched the impact of the servicescape in a variety of service contexts. Two of Wakefield and Blodgett's (1994; 1999) studies combine aspects of the servicescape to form one unidimensional construct. However in their 1996 study, the servicescape was described as comprising of five dimensions namely, Layout, Aesthetics, Seating, Electronic Equipment and Displays, and Cleanliness. Conceptually it seems reasonable to divide the servicescape into such discrete dimensions. However, by including dimensions relating to seating and electronic displays makes their study highly context specific.

In terms of how the servicescape should be depicted, each of the above models were scrutinised and evaluated to determine whether they could be used in the current research project. Firstly, as second order factor modelling is to be used, it is necessary to describe the servicescape in terms of a number of discrete dimensions. This is in keeping with previous empirical work in this area, however, what dimensions should be included

needed to be asked. Rather than taking one author's conceptualisation and trying to replicate it, it was decided that theoretically it is more logical to amalgamate all of the store environment models to form one more comprehensive model. This is particularly useful since to date, each of the previous conceptualisations have been limited in some way.

While it was Bitner (1992) who introduced the term the 'servicescape', her model is not without its faults. Upon finer scrutiny, it becomes evident that some of her dimensions are conceptually unclear. For instance, Bitner (1992) subsumes layout, equipment and furnishings under the 'Space and Function' dimension, servicescape issues which seem to be conceptually quite distinct. Servicescape researchers such as Wakefield and Blodgett (1996) also seem to have realised this and tend to model each of these constructs as separate dimensions. A second ambiguity exists within the 'Signs, Symbols and Artifacts' dimension. This dimension regards signage and communications materials as ways in which businesses communicate with customers about the place. The primary stimuli under this dimension are any signage or communication materials that either describe the name of the company or communicate different rules of behaviour such as no-smoking signs or directions to toilets. Also included in this dimension are more subtle cues such as the quality of the materials and the style of the décor. Conceptually there appears to be little difference between style of décor, which is classified under the 'Signs, Symbols and Artifacts' and furnishings which are placed under the 'Space and Function'. However, there is a difference, albeit quite small; it seems that furnishings under the 'Space and Function' dimension relates to the space between furnishings and where the

furnishings are placed in the environment. Conversely, style of décor under 'Signs Symbols and Artifacts' also encompasses furnishings, but instead of relating to where they are placed in the environment it seems to relate to the aesthetics.

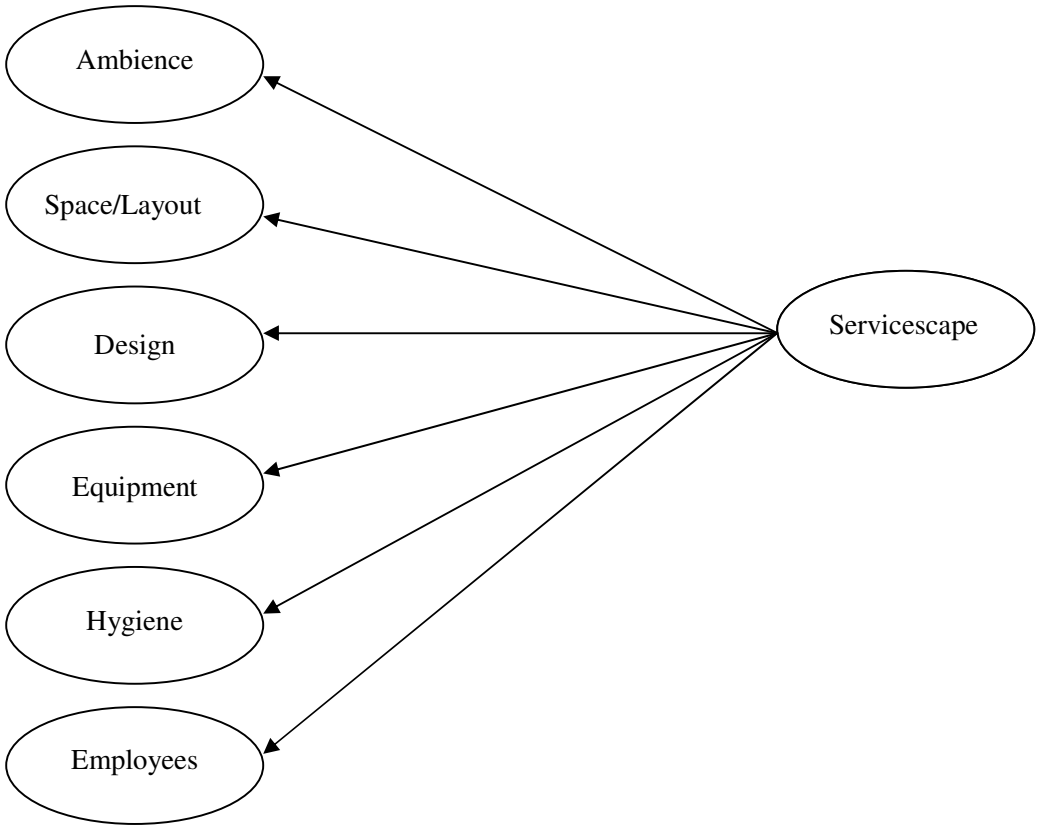


Figure 4.1: Proposed Second Order Factor Model of the Servicescape

From the initial conceptualisation presented in Chapter One it was hypothesised that the Design construct would contain items that reflected both the décor of the interior and signage and communication materials present in the environment. This hypothesis reflected Bitner's (1992) servicescape model as she proposes that these two aspects of the

service environment would fall together. Unfortunately this cannot be tested here as all items relating to signage and communications were removed at the scale development stage. Upon reflection, this is not surprising as others who have included items relating to this dimension have found it to be unreliable and also removed it from their analysis (Hightower et al, 2002). Bitner's (1992) remaining two constructs, 'ambient conditions' and 'space and layout' were also hypothesised to exist here. As per Baker's (1987) models both a design and social element was also proposed. Design elements encapsulate aesthetic cues relating to the interior design and décor. For this study the social element relates only to the role of the employee since all items relating to other customers were also removed at the scale development stage. The final two hypothesised constructs relate to the cleanliness of the environment and the electronic equipment. These two constructs are taken from the work of Wakefield and Blodgett (1994; 1996; 1999) who have included these in their studies. From this literature the following hypothesis was drawn:

H1*: The servicescape is a multi-faceted construct which is aptly described as a second-order factor structure. In particular, constructs which will be subsumed within the servicescape construct are ambient conditions, space/layout, design, employees, hygiene and equipment.*

So that this hypothesis could be tested a consumer survey was conducted. The service chosen for this research was service stations and the following section provides a

rationale for this choice. Following this, a description of the various decisions that were made in relation to the questionnaire itself and its administration will be discussed.

4.3 Chosen Service Context: Utilitarian Service

It was briefly mentioned in Chapter Two that the original focus of this dissertation was to examine the role of the servicescape in professional services, in particular dental services. Unfortunately access could not be gained to any dental service and because of this the emphasis of the dissertation needed to be reappraised. This led to a reassessment of the current research design and also a re-examination of the chosen industry. Upon reflection and further analysis of the servicescape literature, it became evident that there is a dearth of servicescape research on utilitarian services.

The notion of goods and services being classified into being either utilitarian or hedonic in nature was first introduced by Woods (1960) and was later developed by Hirschman and Holbrook (1982). In this treatise of the literature on hedonic consumption, Hirschman and Holbrook (1982) highlighted the differences between hedonic and traditional consumption situations. Firstly, they define hedonic consumption as 'the facets of consumer behaviour that relate to the multisensory, fantasy and emotive aspects of one's experience with products', (Hirschman and Holbrook, 1982, p. 92). Associated with hedonic consumption are heightened levels of emotional arousal and it also effects non-cognitive modalities such as the olfactory, tactile and aural senses. Hedonic products are primarily consumed for sensory gratification, fun and enjoyment (Lim and Ang, 2008) and generate high levels of emotional arousal (Mano and Oliver, 1993). Hirschman

and Holbrook (1982) also compare hedonic consumption with its antithesis, traditional or utilitarian consumption situations which are associated with functionality and rationality where consumer decision making is a logical process.

Classifying phenomena according to whether they are utilitarian or hedonic has become quite topical in recent years, with studies examining hedonic or utilitarian differences in a variety of contexts including; motivation research (Childers et al, 2001; To et al, 2007), advertising appeals (Albers-Miller and Stafford, 1999), attitudinal research (Voss et al, 2003), cross-cultural differences (Lim and Ang, 2008) and value research (Jones et al, 2006; Overby and Lee, 2006). Different types of products and services have also been classified according to whether they are utilitarian or hedonic (Hirschman and Holbrook, 1982; Wakefield and Blodgett, 1999; Dhar and Wertenbroch, 2000; Hightower et al, 2002; Okada, 2005; Chitturi et al, 2008). The difference between the two product or service types stem from the benefits that a consumer receives in return for consuming them (Okada, 2005). Basic necessities such as food and certain types of clothing are utilitarian products that satisfy what Maslow (1943) classed as physiological or safety needs. However, clothing or food products are not limited to being utilitarian products in definition as they can migrate to being classified as hedonic products if they are consumed in a discretionary nature. For instance, some clothing items may be purchased for purely functional purposes such as a rain jacket if the weather is particularly bad. On the other hand, buying haute couture would be deemed hedonic consumption as it is both indulgent and inessential. For this reason, spending on hedonic goods has been described

as being more difficult for consumers to justify than spending money on utilitarian goods (Okada, 2005).

With regards to services, much of the same theory applies, with hedonic services characterised as being more sensual in nature and are consumed for affective pleasure and fun (Hirschman and Holbrook, 1982). Conversely, utilitarian services are those services that are more task oriented and are used to achieve some sort of functional goal (Dhar and Wertenbroch, 2000). Utilitarian services are generally characterised as being low involvement services that are primarily consumed for a particular reason. Stafford and Day (1995) specifically designed a services classification schema that reflects the extent to which a service can be described as being hedonic or utilitarian in nature. This was adapted from Bowen's (1990) three-group classification of services wherein they classify utilitarian services in line with Group 2 in Bowen's (1990) schema. This group is characterised as being low in terms of customer contact with little heterogeneity across different firms' offerings and with an emphasis on the usage of equipment. On the other hand, Bowen's (1990) Group 1 tend to have high degrees of customisation, an extended amount of interaction between customers and employees with high levels of customisation of services across different individuals and also across services (Stafford and Day, 1995).

Wakefield and Blodgett (1999) have also developed a service typology bearing in mind the purpose of the consumption situation in terms of whether it is utilitarian or hedonic and also the length of time the consumer spends in the facility. The authors describe the

reasons that consumers seek out hedonic services as being for pleasure or emotional fulfilment. Given the hedonic and emotive contexts surrounding these services, the authors hypothesise that consumers will be more aware of their surroundings thus making the servicescape a much more potent influencer of consumer behaviour. Given that hedonic services tend to be more extended service transactions in which the emphasis of the service tends to be more sensory, it is likely that the role of the environment will play a greater role than in more utilitarian services. This supposition has led to a wealth of servicescape studies in which hedonic services were chosen as the context. The hedonic services chosen in these studies range from sports stadiums (Wakefield and Blodgett, 1994; Wakefield and Blodgett, 1996), recreation centres (Wakefield and Blodgett, 1999), casinos (Wakefield and Blodgett, 1996; Lucas, 2003; Mayer and Johnson, 2003; Johnson et al, 2004) and restaurants (Ryu and Jang, 2007). Utilitarian services, on the other hand, have received little attention in the literature. Due to this, there is a need to examine whether the servicescape plays an important role in utilitarian services. This is of particular importance since the majority of services that individuals frequent on a regular basis fall into the utilitarian category. Services such as convenience retail outlets and banks are visited by customers on a regular basis whereas the hedonic services chosen by Wakefield and Blodgett (1996; 1999) and others tend to be services that people use only occasionally. Indeed, it must be noted that the atmospherics literature has had a tendency to focus on retail contexts however, this trend has not followed through to more comprehensive evaluations of servicescape environments.

In order to choose a service that can aptly be described as a utilitarian service, both Stafford and Day's (1995) and Wakefield and Blodgett's (1999) typologies were examined. These two typologies can be viewed in Table 4.1 and 4.2 below.

Table 4.1: Stafford and Day's (1995) Classification of Services

	Hedonic Services	**Utilitarian Services**	**Mixture**
Characteristics	Customised, personal services (e.g. Hotel, full service restaurant)	Semi-customized, nonpersonal services (e.g. car repairs, photo processing)	Semi-customised, personal services (e.g. fast food restaurant, budget hotel)
Employee contact	High	Low	Moderate
Employee Involvement	People oriented	Equipment oriented	People/equipment oriented
Perceived difference between firms	High	Moderate	Moderate
Continuous benefits	Moderate	Moderate	Discrete
Service directed at people/things	People	Things	People

Ideally a utilitarian service needed to be chosen that was relatively short in duration and is used by customers to meet some practical purpose. The service would need to have a relatively high level of self-service, with little difference between competitors. It was felt that service stations, (also referred to as petrol stations, filling stations or gas stations) would be a suitable utilitarian service for the current project. Service stations are facilities that sell fuel in the form of petrol or diesel with service stations in Ireland allowing customers to pump their petrol first and to pay afterwards. In addition to selling petrol, many service stations in Ireland also have a convenience store attached in which

food, drinks, magazines and other items are sold. Given the emphasis on convenience and self service this context is befitting for the current research.

Table 4.2: Wakefield and Blodgett's (1999) Service Typology

Time Spent in the Facility	Utilitarian		Hedonic
Short [minutes]	Dry-Cleaner Drive through bank	Fast Food Hair Salon	Game Room Psychologist
Moderate [hours]	Health clinic Law Office	Restaurant Airline	Sporting Event Church Service
Extended [day(s)]	School Hospital	Hotel Conference Centre	Amusement Park Resort

The service stations under scrutiny are part of a large multi-national chain with numerous outlets throughout the country. In general, the interior of these service stations are relatively standardised with similar colour schemes, layouts, and design throughout.

As a service, petrol stations are akin to Stafford and Day's (1995) representation of utilitarian services. Employee contact is generally low for this service as much of the process is delivered by the customer themselves. Equipment also plays a big role in this service as without the petrol pumping equipment the service would be defunct. With regards to Wakefield and Blodgett's (1999) taxonomy, given that customers are likely to

only spend minutes in the service facility, this service also falls into their understanding of utilitarian services.

4.4 Mode of Administration, Sampling Technique and Data Collection

So as to include respondents that had very clear perceptions of the service station, respondents were intercepted as they exited the store. The major advantage of intercepting respondents in this way is that it is much more efficient for the respondent to come to the interviewer than for the interviewer to find respondents who had visited this store. While this approach is quite similar to the traditional mall intercept approach it is in many ways different as the traditional mall intercept approach is generally used to reach customers and decision makers to determine their attitudes and opinions towards a wide range of goods (Frost-Norton, 2005). In mall intercept approaches, specific spaces within malls are often reserved for data collection and it is in these locations the questionnaire is administered. For the current situation, there was no need for a specific location and respondents were questioned as they left the store. In comparison to other methods such as telephone interviewing, personal methods tend to provide answers that are more accurate (Bush and Hair, 1985). However, one disadvantage has been noted with regards to this technique and this is that intercept methods are thought to attract respondents that are somewhat atypical in terms of their demographic profile (Wiseman et al, 1983) and also in relation to their shopping behaviour (Bloch et al, 1989). Further to this, McIntyre and Bender (1986) note that the intercept technique does not consider those who did not make a purchase or use the store and therefore it becomes difficult to determine their motivations for choosing an alternative service provider. Research by

Keillor and Sutton (1993) investigated the profile of shoppers who participated in studies in which an intercept technique was employed. They found that demographically those who participate in mall studies are different to those who avoid participation, however across the two groups no differences were found in their perceptions of the shopping process and their behaviour. This is a key point as it indicates that when one is conducting intercept studies, those who decline to participate view the service in the same way as those who agree to participate. From this, it can be assumed that those who refused to participate in the current study, perceive the service station in the same way as the respondent group.

Intercept methods are inherently intertwined with convenience sampling methods wherein the selection of sampling units is left primarily to the interviewer (McIntyre and Bender, 1986). Unfortunately employing a convenience sampling method could not be avoided as in order to survey the customers of these particular service stations the most effective data collection method was to approach customers on-site. Alternative methods such as telephone, internet, or mail would require that respondents would need to reflect on their last experience with such a service station and this would ultimately introduce other forms of error.

Data was collected by both the researcher and three female students who were paid in return for their help. Having an all-female interviewing panel is particularly advantageous as it has been found that female interviewers can increase response rates in surveys (Hornik and Ellis, 1988). Prior to beginning the data collection, the three female

students were trained on how to administer the survey and the approach that should be taken in asking potential respondents to participate. This was to ensure that the interviewers did not bias or interfere with the validity of the data collection, and to train the interviewers on methods that can be used to reduce nonresponse and refusals. The following section deals with each of these in turn.

4.5 Increasing Response Rates and Interviewing Techniques

The literature is replete with recommendations on how to increase response rates for mail, internet and telephone methods. Common methods for increasing response rates are prior notification (Duhan and Wilson, 1990; Sutton and Zeits, 1992), follow-ups (Roscoe et al, 1975; Mitchell and Brown, 1997; Fox et al, 1998), personalisation of letters to respondents (Eisinger et al, 1974) and incentives (Paolillo and Lorenzi, 1984; James and Bolstein, 1992; Church, 1993). However, all of the above, apart from incentives are inappropriate for personal intercept methods. Although incentivisation could possibly have been used it was not possible due to budgetary restrictions. Given that all of the above techniques were inappropriate for the current situation the researcher needed to rely on other strategies to encourage respondents to take part in the survey. One such strategy are foot-in-the-door techniques (Freedman and Fraser, 1966; Yu and Cooper, 1983; Burger, 1999). Foot-in-the-door techniques are based on the premise that if you ask an individual for a small favour initially and follow it up with a larger favour they are more likely to comply than if you went straight to asking the larger favour. This was employed by the interviewers by first asking potential respondents for a moment of their time which was then followed with an invitation to partake in the survey. A second

method used to increase compliance was to implement a variety of non-verbal communication behaviours. Hornik and Ellis (1988) found that when interviewers made a concerted effort to hold the respondent's gaze, and also lightly touched the potential respondent on the arm, response rates increased. The interviewers were educated on these findings and were asked to introduce them into their approach if at all possible. Obviously, the interviewers were told to rely on their intuition as it is possible that some respondents may react unfavourably to such a strategy.

Once the interviewers were sufficiently prepared, data collection began. This took place between May and June 2007 in three different service stations in the South Dublin region. On permission from the retailer, the interviewers stood in the forecourt of the service station, positioned directly outside the exit of the door. This allowed them to intercept respondents as they exited at which time they drew on their training to recruit respondents and to administer the survey. The interviewers read the questions to the respondents and recorded their answers on the two-page questionnaire (See Appendix 4A). As well as including all questions relating to the servicescape and its dimensionality, other data was also collected at this stage which will be discussed in more detail in later chapters.

4.6 Format of the Questionnaire

The format of the questionnaire was in many ways quite similar to the format used at the scale development stage. For instance, Likert scales were used in which the respondent was required to express their agreement with each statement. However, rather than using

five-point scales, at this empirical stage of the research it was decided that seven-point scales would be more appropriate. Traditional guidelines recommend that the appropriate number of categories should be a seven point scale plus or minus two (Cox, 1980) with scales containing two or three alternatives generally thought to be inadequate. In a more recent study, it was found by Dawes (2008) that there were no differences between 5-point, 7-point or 9-point scales across standard variation, skewness, kurtosis and variance about the mean. This finding suggests that all three formats are comparable if used in confirmatory factor analysis or structural equation modelling (Dawes, 2008). Despite this, the decision was made to use scales with a seven points as it has been found that some statistics such as the correlation coefficient can decrease when the number of categories are reduced (Malhotra, 2006). In addition to this, using scales with a larger number of scale points allows for more variability and for subtle nuances to be discerned (Laroche et al, 2001).

Placing verbal labels on each point on a scale has been recommended as it is thought to clarify the purpose of the scale (Schwarz et al, 1991). However, for seven or more items this is quite impractical as the gradations of agreement become too fine to express in words (Dawes, 2008) and it was thus decided to anchor only the end points of the scale with ‘strongly disagree’ and ‘strongly agree’. The next decision pertained to the value or the description of the scale items. As it was decided to use a seven-point scale, intuitively one would expect that the numeric values of the scale are arbitrary once there are equal distances between scale values. For instance, one would imagine that there should be no difference between using scale points of -3 to +3, 1 to 7 or 0 to 6. However,

research by a number of authors has proven otherwise (Schwarz et al, 1991; O'Muircheartaigh et al, 1995; Sangster et al, 2001) where it has been consistently found that respondents avoid negatively valued items or values on the zero point of a scale. Therefore, continuum scales in which the starting point begins at one are recommended if negative skewness is to be minimised (Sangster et al, 2001). This was the approach taken here with scale points ranging from 1 to 7.

The final formatting decision related to the ordering of the servicescape items on the questionnaire itself. Research has shown that respondents are more likely to perceive items to be similar to one another when placed together (Podsakoff et al, 2003; Tourangeau et al, 2004). While discussing measurement issues in relation to multifaceted constructs, Briggs (1992) also notes that items that are systematically ordered load together more consistently than if they are scattered randomly throughout the measure. Given that the primary purpose of this chapter is to determine whether the servicescape structure proposed in Figure 4.1 holds in real service environments, it was decided not to group the items from the sub-constructs together which may unnaturally create associations between constructs. Instead it was decided to scatter the items randomly throughout the servicescape measure (see Appendix 4A).

4.7 The Survey Instrument

As it is expected that the service environment can be described as a multifaceted construct, the multi-item scale developed in Chapter Two was employed here. The

following table (Table 4.3) displays the items used to represent each construct (a full copy of the questionnaire can be viewed in Appendix 4A).

Table 4.3: Scale Items and Their Related Constructs

Sub-dimension	Item
Ambience	The background music was pleasant
	The lighting was comfortable
	The atmosphere was comfortable
	The store had a pleasant smell
	The background music was appropriate
Space & Layout	The flooring was appropriate
	I found my way around quite easily
	The interior layout was pleasing
Design	I found the interior design visually appealing
	The interior design was attractive
	The colour schemes were pleasant
	The materials used were of high quality
	The architecture was attractive
	I found the physical facilities comfortable
Equipment	The equipment was modern looking
	The electronic equipment was excellent
	The equipment was of high quality
Hygiene	The store was very clean
	The forecourt was clean and tidy
	The service station appeared to be hygienic
Employees	The employees were neat and tidy in appearance
	I found the staff friendly
	The employees were helpful

Given the change in service context, some modifications were made to the scale developed in Chapter Two. The first alteration involved changing the wording from 'the bathroom facilities were clean and tidy' to 'the forecourt was clean and tidy'. The second change involved rewording the 'waiting room was comfortable' item to 'the atmosphere was comfortable'. Two items were also deleted as it was felt that they were both ambiguous and inappropriate for the service station context, these were 'the seating in the

surgery was comfortable' and 'the style of the interior was fashionable'. Finally, an additional cleanliness item was included as only two existed in the final scale which could cause identification problems later on in the analysis.

4.8 Sample

A total of 355 questionnaires were collected, all of which could be used in the final analyses. As structural equation modelling is considered a large sample technique, this sample size more than meets sample-size requirements. This sample was composed of 66.2% male and 33.8% female with a mean age between 26 and 35 years (SD = 1.67). For further details on the age profile of respondents please see Appendix 4B.

4.9 Normality Assumptions, Treatment of Outliers and Missing Values

An inspection of the skewness and the kurtosis of the data revealed that the majority of items did not violate this assumption. However, two items were relatively kurtic, 'I found the staff friendly' and 'I found my way around quite easily' but it was decided to retain these items despite their distribution as research has demonstrated that the Maximum Likelihood Estimator is not as sensitive to deviations from normality as previously thought (Gerbing and Anderson, 1985; Anderson and Gerbing, 1988). Using Prelis, Mardia's Statistic was also run which produced a value of 1.307 which is below the recommended cut-off of 3 (Bollen, 1989) therefore indicating that multivariate normality exists. Outlier analysis was also conducted using Mahalanobis Distance (D^2), Cook's and Leverage Values (Mullen et al, 1995). A number of outliers were detected at this stage, however, no deletions were made. It was decided to hold onto these cases for

the time being and to determine at the CFA stage whether these outliers jeopardized the results in a negative way.

It was found that very few cases contained missing data, the reason for this is probably related to the mode of administration. It is likely that respondents were more inclined to complete the entire survey as the presence of the interviewer may have made them feel more obliged to continue the interview. Despite the relatively small number of missing data values, the dataset was examined to determine whether the pattern of the data was MCAR, MAR or MNAR. As was discussed in the previous chapter, results found that the data was indeed MCAR thus making it ideal for any form of missing data treatment. Given that relatively few missing cases were found, listwise or pairwise deletion could have been used. However, it was decided that the more sophisticated Expectation-Maximisation (EM) Algorithm (Dempster et al, 1977) would be used given its availability in later versions of SPSS.

Once all preliminary analysis was conducted the next stage was to begin using structural equation modelling to determine whether the hypothesised factor structure can be upheld in a real-life utilitarian service environment. This is dealt with in the sections that follow. The section begins by examining a single-order factor structure of the servicescape. Following this, the superiority of a multi-dimensional conceptualisation is demonstrated.

4.10 Confirmatory Factor Analysis

This next stage of the analysis involved testing whether a single-factor representation of the servicescape is superior to a multidimensional model. The one-factor model approach has been taken by a number of authors in the past, some have had success with this (Wakefield and Blodgett, 1994) while others have struggled to find a well fitting solution (Reimer and Kuehn, 2005). In order to decide whether the measurement properties of a one-factor solution or a multiple-factor model is superior can be done so using either exploratory factor analysis or confirmatory factor analysis (Rindskopf and Rose, 1988). As exploratory factor analysis is a data driven approach in which there are no prior specifications regarding the underlying structure of the data it is unsuitable for the current research problem. On the other hand, confirmatory factor analysis (CFA) allows the researcher to specify *a priori* models to decide which of these two models fits the data better and for this reason is regarded as a confirmatory technique. Confirmatory factor analysis is also regularly used by other researchers to assess the dimensionality of various different constructs (Liden and Maslyn, 1998; Laroche et al, 2001; Laroche et al, 2004).

The confirmatory factor analyses were conducted in two stages. As mentioned above, the first stage involved modelling the servicescape as a one-factor model, while the second stage was to determine whether the factor structure proposed earlier in this chapter is a better representation of how consumers perceive the servicescape. Within this second stage there were a number of sub-stages that led to a final servicescape model. Firstly, it was necessary to examine each sub-dimension in isolation to determine whether there

were any items that did not meet reliability criteria. For constructs with three or less items, identification criteria could not be met so for these dimensions they moved straight to the second phase. In this stage, each construct was examined in relation to every other sub-construct in the model to determine whether there were any major misspecifications in the model and whether discriminant validity was found between constructs. This is in line with Anderson and Gerbing's (1988) two-step approach who opine that prior to conducting structural analysis it is important to achieve unidimensionality within each construct first. Unidimensionality occurs when only one construct underlies a set of items and has been regarded as one of the most essential assumptions of measurement theory (Steenkamp and Trijp, 1991). Once unidimensionality was achieved between each sub-construct the full servicescape model was run.

4.10.1 One Factor Model

This first model was analysed using confirmatory factor analysis with LISREL 8.51 (Jöreskog and Sörbom, 1996). The estimator used was the Maximum Likelihood Estimator and a converged, admissible solution was obtained. As with all the models from this point on, all 355 participants were included in the analyses. The following figure depicts the one-factor model on which all 23-items loaded on the servicescape construct.

Fit statistics for this model were poor with a significant chi square (χ^2) value of 1038.34, p = 00000, df = 100. Other fit statistics also did not meet the criteria standards for adequacy of fit: RMSEA = 0.10, CFI = 0.76, SRMR = 0.074, NFI = 0.71. In addition to

this, a number of the standardised loadings fell below the recommended cut-off of .4 further suggesting that perhaps this model has been misspecified. However, all t-values were significant.

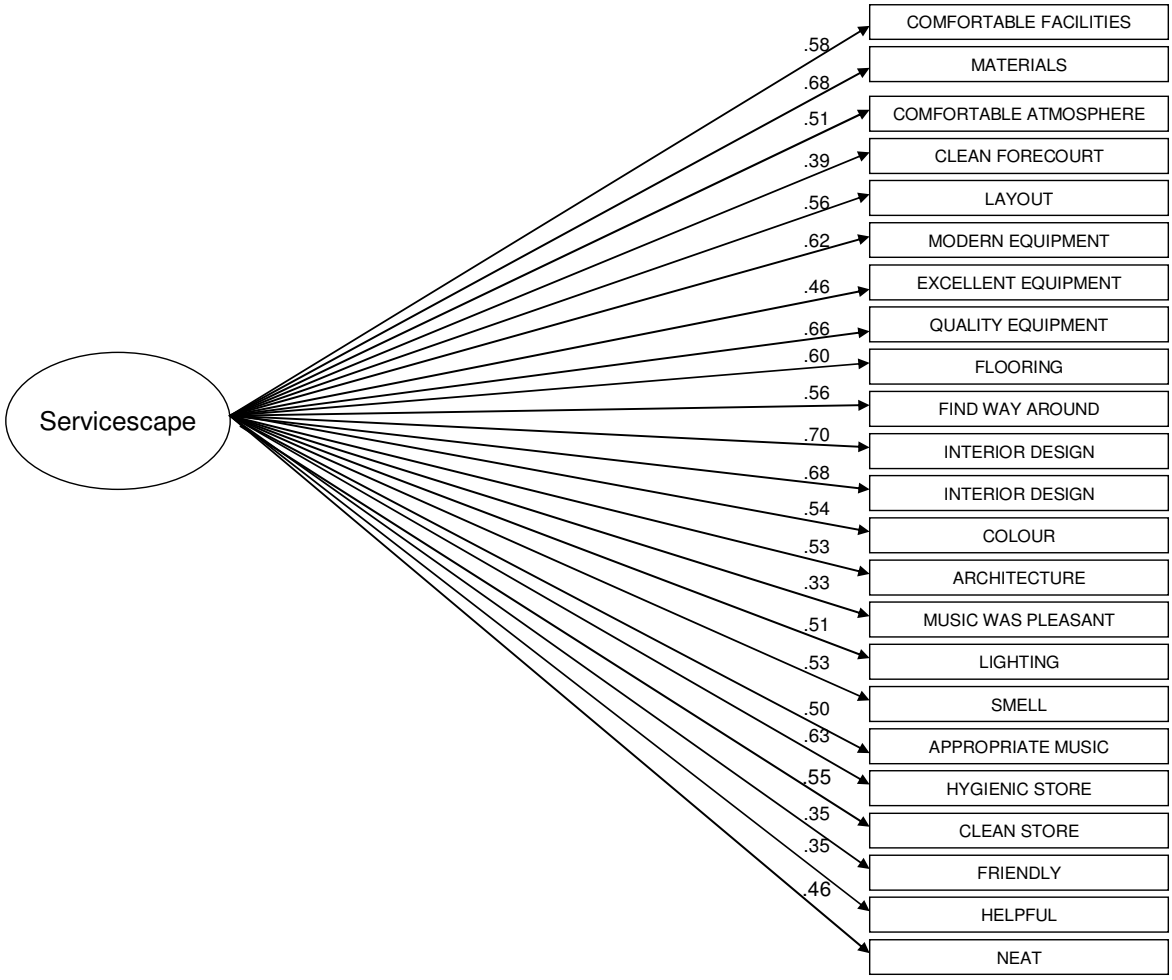


Figure 4.2: Standardised Loadings for a One-Factor Servicescape Model

This finding is very similar to that of Reimer and Kuehn (2005) who found that their thirteen-item representation of the servicescape did not fit the data well. Perhaps the reason behind why the model fits so poorly is because in reality it is possible that consumers perceptually divide the servicescape into a number of different concepts

which in turn are combined in a synergistic manner to form an overall impression of the environment.

4.10.2 Multiple Factor Model

The next stage of the analysis tested this proposition by modelling the servicescape as a second-order factor structure. However, as per Anderson and Gerbing's (1988) recommendations, prior to running the full second order model it was necessary to ensure that all sub-constructs were unidimensional. To achieve unidimensionality both convergent and discriminant validity need to be assessed. Convergent validity is 'is the degree to which multiple attempts to measure the same concept are in agreement', while discriminant validity is the 'degree to which measures of different concepts are distinct' (Bagozzi and Yi, 1991, p. 427). To meet this end each construct was chosen individually and examined in isolation and then in conjunction with other constructs in the model. The reasoning behind this is that issues such as discriminant validity may be obfuscated by examining several constructs at once (Anderson and Gerbing, 1988). Following recommendations by Bagozzi (Bagozzi, 1980; Bagozzi and Yi, 1991; Bagozzi et al, 1991) discriminant validity was tested by determining whether the correlation between two constructs (phi φ) is significantly less than 1.0. At this stage convergent validity (Kline, 2005) was also assessed and this relates to whether the indicators specified all have relatively large standardised loadings on that factor (> . 40). The following section discusses this in detail.

Design

The design construct was run with its six hypothesised indicators loading on it. Table 4.4 summarises the results from this CFA. Overall the fit for this sub-construct is just within recommended limits with an RMSEA of 0.069 which is just below Steiger's (2007) upper threshold of 0.07. The squared multiple correlations (R^2) is analogous to the R^2 obtained in traditional regression analysis therefore, higher values indicate more reliable items (Diamantopoulos and Siguaw, 2000). It is evident from Table 4.4 that some of the Design variables have dubious reliability scores with 'I found the physical facilities comfortable' obtaining an R^2 of 0.22.

Table 4.4: CFA for Design Construct

	Estimates	**Standard-ised Loading**	**t-value**	**R^2**
I found the interior design visually appealing	1.10	0.81	17.05*	0.66
The interior design was attractive	1.10	0.81	17.00*	0.65
The colour schemes were pleasant	0.77	0.58	11.17*	0.34
The materials used were of high quality	0.74	0.62	11.98*	0.38
The architecture was attractive	0.88	0.57	10.87*	0.33
I found the physical facilities comfortable	0.54	0.47	8.58*	0.22
Model fit: Chi-Square=23.98, df=9, P-value=0.00433, RMSEA=0.069, CFI=0.98, NFI=.96 * = significant at the p = 0.01 level				

Following this, the Design construct was then examined in relation to every other construct in the model. A full description of all the statistics from this stage of the analysis can be viewed in Appendix 4C. Overall, design performed quite well in relation to the other constructs. However, the 'I found the physical facilities comfortable', item

consistently cross-loaded onto all other constructs. This instability coupled with its very low squared correlation coefficients made this item an ideal choice for deletion and was subsequently removed from the analysis. 'The materials used were of high quality' was also found to be relatively unreliable due to it crossloading onto the Hygiene dimension and was also deleted from further analyses. Whether this construct met discriminant validity criteria was also calculated using Bagozzi and Yi's (1991) formula and it was found that this construct was independent of other constructs in the model. Full details of discriminant validity calculations can be found in Appendix 4D.

Hygiene

Identical procedures to the Design construct were followed in dealing with this dimension. As originally proposed, Hygiene was expected to have three items loading onto it, namely, 'The store was very clean', 'The forecourt was clean and tidy' and 'The service station appeared to be hygienic'. As this construct was underidentified it was not possible to determine its fit when modelled in isolation. However, convergent validity and R^2 values could be assessed. The following table summarises the results from this confirmatory factor analysis. It is evident from the table overleaf that 'The forecourt was clean and tidy' item is problematic, achieving a standardised loading of merely 0.37 which violated the convergent validity assumption. In addition, the R^2 value is worryingly low at 0.13. Convergent validity problems were also found with this item when Hygiene was modelled with other servicescape constructs. In addition to this it was found that this item crossloaded with the Space construct. These tenuous findings resulted in this item being deleted from further analyses. On the other hand, the

discriminant validity between the Hygiene construct and others was generally satisfactory.

Table 4.5: CFA for Hygiene Construct

	Estimates	**Standardised Loading**	**t-value**	**R^2**
The store was very clean	0.54	0.59	7.41*	0.35
The forecourt was clean and tidy	0.53	0.37	5.67*	0.13
The service station appeared to by hygienic	0.79	0.81	8.31*	0.65
Model Fit: Not identified Chi-Square=0.00, df=0, P-value=1.00000, RMSEA=0.000 * = significant at the p = 0.01 level				

However, this assumption was not met between it and the Space construct with an intercorrelation (φ) of 0.92 between the two. This indicates that perhaps consumers cannot discriminate between these two constructs in their minds. Conventional wisdom would suggest that these two constructs should be merged. However, rather than allowing statistics to drive the analysis, it was decided that these two constructs should remain distinct from one another at this stage of the analysis so that the model represents previous theoretical propositions.

Space

The space dimension originally consisted of three items, 'The flooring was appropriate', 'I found my way around quite easily', and 'The interior layout was pleasing'. The initial CFA of this constructs suggested that perhaps the squared correlation coefficient for 'The interior layout was pleasing' item was too low. Through analysing this construct in conjunction with other servicescape constructs it came to light that this item crossloaded

onto both the Design and Equipment constructs and for this reason was withdrawn from any remaining analyses.

Table 4.6: CFA for Space Construct

	Estimates	**Standardised Loading**	**t-value**	**R^2**
The flooring was appropriate	.91	.75	9.61*	.56
I found my way around quite easily	.53	.56	8.22*	.32
The interior layout was pleasing	.62	.51	7.71*	.26
Model Fit: Chi-Square=0.00, df=0, P-value=1.00000, RMSEA=0.000 * = significant at the p = 0.01 level				

As mentioned above, a problem arose regarding the discriminant validity between this construct and the Hygiene dimension however for theoretical reasons no decision was made to merge these two dimensions at this point in the analysis.

Equipment

The Equipment construct was also underidentified and could not be assessed in terms of fit criteria. All three items on this construct performed well in terms of their reliability and convergent validity.

Table 4.7: CFA for Equipment Construct

	Estimates	**Standardised Loading**	**t-value**	**R^2**
The equipment was modern looking	.69	.59	9.53*	.34
The electronic equipment was excellent	.70	.57	.9.31*	.32
The equipment was of high quality	1.05	.82	11.92*	.67
Model fit: Not identified Chi-Square=0.00, df=0, P-value=1.00000, RMSEA=0.000 * = significant at the p = 0.01 level				

In examining the Equipment dimension in relation to others it also performed very well with no crossloadings on any other constructs. With regards to discriminant validity it was also found that this construct was unique in relation to all others in the model. Therefore, no item changes were made.

Ambience

Originally containing five items in total, the Ambience dimension is the second largest construct in the model. Unfortunately the fit for this model was quite poor achieving an RMSEA of 0.148. Many items obtained only very low levels of reliability in relation to their squared multiple correlation. 'The atmosphere was comfortable' item performed particularly badly with both a low standardised loading and a low R^2.

Table 4.8: CFA for Ambience Construct

	Estimates	**Standardised Loading**	**t-value**	**R^2**
The background music was pleasant	. 65	.82	11.80*	.43
The lighting was comfortable	.42	.48	7.20*	.17
The atmosphere was comfortable	.39	.32	6.65*	.15
The store had a pleasant smell	.42	.49	7.26*	.18
The background music was appropriate	1.01	.81	14.78*	.66
Chi-Square=43.79, df=5, P-value=0.00000, RMSEA=0.148, CFI=0.88, NFI=0.87 * = significant at the p = 0.01 level				

Almost all of the ambience items cross-loaded with other constructs in the model (see Appendix 4C). However, it was 'The atmosphere was comfortable' item that was the most worrying with an extremely high crossloading onto the Employees construct. No theoretical justification could be found to move this item onto the Employees construct and was thus deleted from the analysis at this stage. Removal of this item provides a substantial improvement to fit (Chi-Square = 7.37, df = 2, P-value = 0.02504, RMSEA = 0.087). This however, does not solve the problem relating to other items crossloading onto many other servicescape constructs. Despite the many crossloadings, phi values between the Ambience dimension and other constructs were not excessively high. For substantive reasons it was decided to retain these items given that they are considered such an important part of the servicescape. Further discussion of this is found later in this chapter.

Employees

It was hypothesised that the Employees dimension would contain three items. These being 'The employees were neat and tidy in appearance', 'I found the staff friendly' and 'The employees were helpful'. As with many earlier constructs, model identification was not achieved which limits the ability to assess the strength of the model's specification. Regardless, it was found that standardised loading and R^2 values were all within normal limits. Overall the Employees construct performed quite well. Modification indices between this and other constructs were almost negligible apart from 'The employees were neat and tidy in appearance' item which had a crossloading on the Ambience dimension. Theoretically it would not make sense to place this item onto the ambience construct and

was therefore ignored. Statistical discrimination between this construct and others in the model was also found.

Table 4.9: CFA for Employees Construct

	Estimates	**Standardised Loading**	**t-value**	**R^2**
The employees were neat and tidy in appearance	.43	.52	7.73*	.27
I found the staff friendly	.56	.67	8.93*	.45
The employees were helpful	.55	.58	8.25*	.34
Model fit: Not identified. Chi-Square=0.00, df=0, P-value=1.00000, RMSEA=0.000 * = significant at the $p = 0.01$ level				

Deleted Items: Rationale for their Removal

In total there were five items deleted at this stage of the analysis. The two primary reasons for deleting these items related to excessively high crossloadings with other factors in the model or for obtaining R^2 values that were unacceptably low. Earlier it was briefly mentioned that when deciding to delete an item from the analysis it was not made purely for statistical reasons. The underlying theory which has been developed by previous authors within the work of store environments and servicescape was borne in mind when deciding whether an item should remain in the dataset. The five items will now be discussed in terms of their relatedness to theory and the implications associated with their removal.

The first item to be removed from the analysis was 'I found the physical facilities comfortable'. It would be very difficult to include this item in further analysis given its

high crossloadings with other constructs in the model. Upon reflection, it is possible that this item confused respondents. The word 'comfortable' in this context is perhaps a little ambiguous. However, if this item was used in an alternative context such as professional services perhaps this problem would not arise because in many extended professional transactions the respondent is likely to spend more time in the service thus making how comfortable the respondent is much more important. For instance, if a patient of a dentist found the treatment chair in which they sat was relaxing it would be likely that when a question was posed in relation to comfort levels of the facility when would have a clearer understanding of how to answer it. It was also found that 'The materials used were of high quality' item was relatively unreliable and was also removed. It is felt that taking this item from the analysis is not detrimental theoretically as the items that remain in the dataset adequately capture the essence of the Design construct. As with the previous item that was removed, it is likely that respondents found this question vague and possibly did not quite understand the meaning of the word 'materials'.

Table 4.10: Items Deleted

Item	**Construct**
'I found the physical facilities comfortable'	Design
'The materials used were of high quality'	Design
'The forecourt was clean and tidy'	Hygiene
'The interior layout was pleasing'	Space
'The atmosphere was comfortable'	Ambience

For the hygiene dimension, the item that performed most poorly and which was subsequently removed was the analysis was 'The forecourt was clean and tidy'. Removing this from the analysis leaves the hygiene dimension with only two items.

While having two-indicator constructs is not an ideal situation, it does not jeopardise identification of the second order factor model. From the Space dimension 'The interior layout was pleasing' item could not be retained due to low R^2 and crossloadings. As with other items that were removed from the analysis, it is possible that respondents did not quite know how to respond to this question as the layout of a store may not be something that respondents consider in terms of how pleasant it is. No problems were encountered with regards to the remaining two Space items and it is possible that the phrasing of the two items was somewhat clearer than the item that was deleted from the analysis.

On the whole, many of the Ambience items were found to be problematic. However, as discussed above, only 'The atmosphere was comfortable' item was actually removed. Theoretically this item does not relate to any specific atmospheric stimulus but is a more general item which attempted to measure consumers overall impression of the ambient conditions of the service environment. Unfortunately this item proved highly questionable in terms of its reliability and could not be retained. Unlike many of the ambience items, this question has no direct link to any previous atmospheric research such as the music, lighting or smell items. This linkage between the music, lighting and smell items was a key driving force for retaining these items at this stage. As one of the primary objectives of this dissertation is to provide a representation of the servicescape that reflects the wide body of literature in this area it was felt by taking these ambience items out would detrimentally effect the comprehensiveness of this model.

Second Order Factor Model of the Servicescape

Upon deletion of the above items the next stage was to run the full second order model of the servicescape. It is hypothesised that this model will form a six factor structure with multiple indicators on each construct. In particular it is proposed that the servicescape construct underlies each of these six dimensions. Previous theorists and researchers have shown how each of these sub-dimensions are similar to one another. For consumers it is likely that they do not regard servicescape sub-constructs in isolation but as a number of intertwining concepts that synergistically combine to form their overall impression of the store environment. In essence, this conceptual framework proposes that the six constructs share a common servicescape core that exists at a higher level of abstraction. Figure 4.3 overleaf is a graphical representation of this framework.

Confirmatory factor analysis also allows to test theoretical relationships in which a structure is specified which accounts for the relationships among the factors. This is commonly known as second-order factor analysis and is a more restrictive and parsimonious method of conceptualising relationships than standard confirmatory factor analysis (Rindskopf and Rose, 1988). In addition to this, second order factor models can also be used to test the reliability and validity of measures which allows researchers to determine whether one model specification is superior to another.

Measurement Model Analysis

As with the other analyses in this chapter LISREL 8.51 was used to test the hypothesis that the servicescape is a multi-dimensional construct. Table 4.11 summarises the CFA

results and the covariance matrix from this data can be found in Appendix 4E. Each indicator is specified to load on its latent factor. In overall model fit, the model yields a chi-square of 403.82 (df = 129, p = 0.00). Because the chi-square statistic is known to be overly sensitive to sample size (Bentler and Bonnet, 1980), more emphasis is placed on other statistics such as the RMSEA and incremental fit indices such as the CFI, SRMR, NFI and GFI. For this model the RMSEA, CFI, SRMR and NFI were 0.078 (90% confidence interval: 0.069 - 0.086), 0.86, 0.082, and 0.81 respectively.

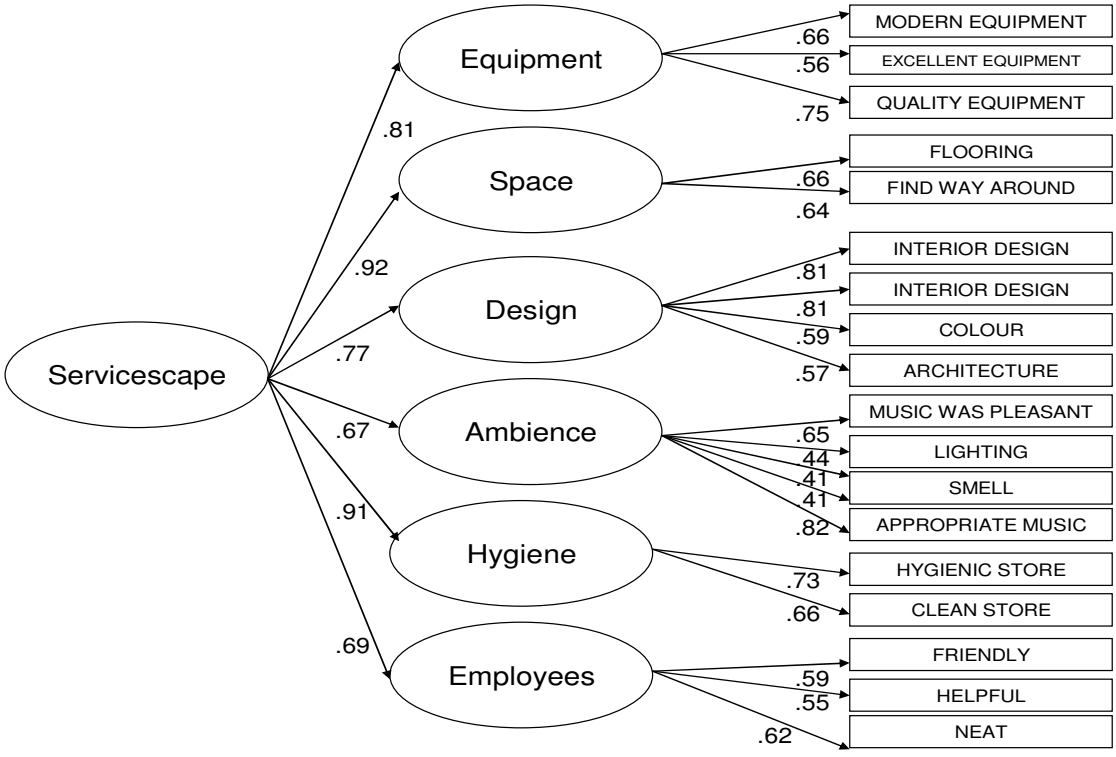


Figure 4.3: Second Order Factor Model of the Servicescape

Unfortunately, recommended threshold levels for each of these statistics were not met (Byrne, 1998; Miles and Shevlin, 1998; Hu and Bentler, 1999) indicating that model fit was poor. The first issue which must be scrutinised is to determine whether the factor loadings of the items are greater than .4 (Kline, 2005). As is evident from Table 4.11, all standardised loadings are greater than .4 indicating that all items converge on some common point. Also, without exception it was found that the estimated loadings are relatively large and all are significant (t-values > 6, $p < .01$). This test is the ratio of the parameter estimate to its standard error and indicates whether parameter estimates are statistically different from zero (Fornell and Larcker, 1981). Solely relying on t-tests to assess convergent validity has been criticised and in response to this Fornell and Larcker (1981) developed two tests that allow us to assess the reliability of a construct. The first of these is their measure of composite reliability which assesses internal consistency and is therefore analogous to that of coefficient alpha. While it is still commonly applied, coefficient alpha is thought to understate reliability whereas Fornell and Larcker's (1981) method is thought to provide a more accurate indication of internal consistency. This is calculated as follows (Fornell and Larcker, 1981, p. 45):

$$\rho_c = \frac{(\sum_{i=1} \lambda_{yi})^2}{(\sum_{i=1} \lambda_{yi})^2 + \sum_{i=1} \mathrm{Var}(\varepsilon_i)}$$

where:

ρ_c = composite reliability

λ = indicator loadings

ε = indicator error variances

Σ = summation over the indicators of the latent variable

According to Bagozzi and Yi (1988) composite reliability statistics need to exceed 0.60 and it is evident from Table 4.11 that for all but the Space construct has this criterion been met. However, with a composite reliability of .59, the Space construct is only marginally less than the recommended cut-off point and is therefore not hugely worrying.

Complementary to Composite Reliability is Fornell and Larcker's (1981) Average Variance Extracted. This formula uses very similar information to the above equation however, this measure shows the amount of variance that is captured by the construct in relation to the amount of variance due to measurement error (Fornell and Larcker, 1981). Average Variance Extracted values less than .50 suggest that measurement error accounts for more variance in the indicators than the underlying latent variable which sheds doubt on the soundness of the indicators or the latent variable itself (Diamantopoulos and Siguaw, 2000). As LISREL also does not calculate values for the Average Variance Extracted (AVE) it is necessary to calculate it using the following equation (Fornell and Larcker, 1981, p. 46):

$$\rho_v = \frac{\sum_{i=1} \lambda^2_{yi}}{\sum_{i=1} \lambda^2_{yi} + \sum_{i=1} \mathrm{Var}(\varepsilon_i)}$$

where λ, ε, and Σ are defined as above. The constructs did not perform as well on Average Variance Extracted (See Table 4.11) with only two constructs having AVE values greater than .50. This suggests that three constructs have marginal reliabilities thus calling into question the validity of these constructs. The next stage of model assessment relates to whether discriminant validity between the model constructs has been achieved.

There are a number of methods for assessing discriminant validity between constructs. One such method is the aforementioned Bagozzi (Bagozzi and Yi, 1991) method which determines whether the correlation between two constructs (phi φ) is significantly less than 1.0. A second more stringent method is based upon the Average Variance Extracted which specifies that if the AVE is greater than the square of the correlation between the latent variables then discriminant validity obtains (Fornell and Larcker, 1981). In testing the discriminant validity in second order models, the AVE is most appropriate as phi values between the constructs are not calculated. Using Fornell and Larcker's (1981) test of discriminant validity it was found that the Space and Hygiene constructs were found to lack discriminant validity with a number of other constructs. A related problem was found at the first stage of the confirmatory factor analysis and has led to problems at this stage of the analysis. Full details of the inter-correlations between constructs can be found in Appendix 4F.

Table 4.11: Results from the Second-Order Factor Analysis of Study Constructs

	Loading	t-value	Coefficient α	Composite Reliability	Average Variance Extracted
Ambience			.63	.68	.36
The background music was pleasant	.65	—[a]			
The lighting was comfortable	.44	6.92*			
The store had a pleasant smell	.41	6.53*			
The background music was appropriate	.82				
Space & Layout			.58	.59	.42
The flooring was appropriate	.66	9.37*			
I found my way around quite easily	.64	—[a]			
Design			.78	.79	.50
I found the interior design visually appealing	.81	10.88*			
The interior design was attractive	.81	10.89*			
The colour schemes were pleasant	.59	—[a]			
The architecture was attractive	.57	8.69*			
Equipment			.69	.70	.44
The equipment was modern looking	.66	10.27*			
The electronic equipment was excellent	.56	8.92*			
The equipment was of high quality	.75	—[a]			
Hygiene			.65	.65	.48
The store was very clean	.65	—[a]			
The service station appeared to be hygienic	.73	10.24*			
Employees			.61	.62	.35
The employees were neat and tidy in appearance	.61	7.03*			
I found the staff friendly	.60	6.88*			
The employees were helpful	.56	—[a]			
Fit Statistics: χ^2 = 403.82, df = 129, p = 0.00. RSMEA = 0.078, CFI = 0.86, SRMR = 0.082 and NFI = 0.81					

[a] = The corresponding coefficient was fixed to set the metric of the latent construct.

* = significant at the p = 0.01 level

The above representation of the servicescape model is a comprehensive attempt at presenting the theory in this area. Nevertheless, it fails to replicate the underlying pattern that exists in the data. The two major issues in relation to the above model relate to poor item reliabilities on the ambience dimension and a lack of discriminant validity between the Hygiene and the Space dimension. The following section discusses these in detail.

In Chapter One previous servicescape studies were criticised for their lack of consistency regarding the items that they chose to include to represent ambient conditions. The basis for this criticism was that theoretical papers on the servicescape are clear that the ambience items included in this study (music, lighting, smell) are key to forming consumers perceptions of ambient conditions. Empirical research in this area has consistently presented measures of ambient conditions that are either lacking some key stimuli (Wakefield and Blodgett, 1999; Baker et al, 2002) or else fail to address this important dimension entirely (Wakefield and Blodgett, 1996; Lee et al, 2008). It is unclear whether the initial measures used in these studies consisted of a more inclusive measure of ambient conditions at the outset which was reduced to a lesser number of items due to reliability analysis. What is known is that in the current study, many of the ambience items were found to have troubling reliabilities. One can only surmise as to why these items achieved such low reliabilities. One possible reason is that perhaps these items are outside respondents' level of consciousness and for this reason they are not aware of their presence. When questioned, it is likely that respondents were unsure how to reply to such questions resulting in extremely varied responses from participants with poor reliabilities ensuing. However, it is possible that only when these stimuli are

extreme will consumers regard their presence. For example, if there was a particularly unfavourable smell in an environment, consumers would be more likely to notice it which would result in more consistent results from respondents (i.e. most respondents would rate the scent as unfavourable). This was not the case here as smells within the environment were not especially pronounced which led to relatively poor reliability levels. This is also likely to have occurred with the lighting item with respondents not making a conscious evaluation of this aspect of the service environment.

While the current study encountered problems in measuring ambient conditions, within the atmospherics literature it has been consistently demonstrated how stimuli such as lighting and olfaction can impact behavioural responses (cf. Turley and Milliman, 2000). It is likely that this has been possible due to the experimental approach that is taken by many of these authors. Experimental studies offer the distinct advantage of permitting the researcher to manipulate independent variables so that consumer behaviour can be monitored. The design also eliminates directly questioning respondents as stimulus modifications are used to assess the respondents' behaviour. For the current study, it was not possible to alter ambient cues within the service environment because of the field approach that was implemented.

Although the current study design required that respondents were directly asked their opinion of the ambient conditions, in hindsight perhaps this is not a suitable method for determining how these stimuli contribute to respondents overall perceptions of the servicescape. This is not saying that ambient conditions are irrelevant to consumers,

rather it is proposing that perhaps the most effective method for examining these variables is through experimental methods. Due to this, it was decided that the two most unreliable of the ambience items be deleted from further analysis. These were: 'The lighting was comfortable' and 'The store had a pleasant smell'. This leaves only two remaining ambience items these being: 'The background music was pleasant' and 'The background music was appropriate'. Interestingly, a number of authors in the past have used only music to represent ambient cues (Baker et al, 2002; El Sayed et al, 2004) and perhaps this is because music is much more obvious to a consumer than the lighting or smells that may be present in the environment.

The second problem in relation to the above conceptualisation is that discriminant validity could not be found between the Hygiene and the Space dimensions. At the first stage of confirmatory factor analysis in which each construct was modelled against every other construct within the servicescape, it became evident that the two constructs did not exhibit discriminant validity. This has also followed through to the second order factor structure causing multiple discriminant validity problems. Lacking discriminant validity between constructs suggests that respondents regard them both the same conceptually. Theoretically, Space and Hygiene dimensions have always remained apart from one another. In fact hygiene has been subsumed by some under Ambient conditions (Ezeh and Harris, 2007; Harris and Ezeh, 2008) which seems rather inappropriate since ambient conditions are generally assumed to relate to the five senses (Bitner, 1992). However, as suggested by the results, amalgamating the Space and Hygiene constructs seems more reasonable as it is likely that the more clean and tidy an environment the greater the

feeling of space. In fact one of the fundamental tenets of the ancient Chinese art Feng Shui is that keeping interiors tidy and reducing clutter are key to maintaining a healthy space (Lip, 1997). Having decided to omit a further two ambience items and to merge the Space and Hygiene dimensions the model was re-run. This model can be seen in Figure 4.4 and the results from this are discussed in detail within the following section.

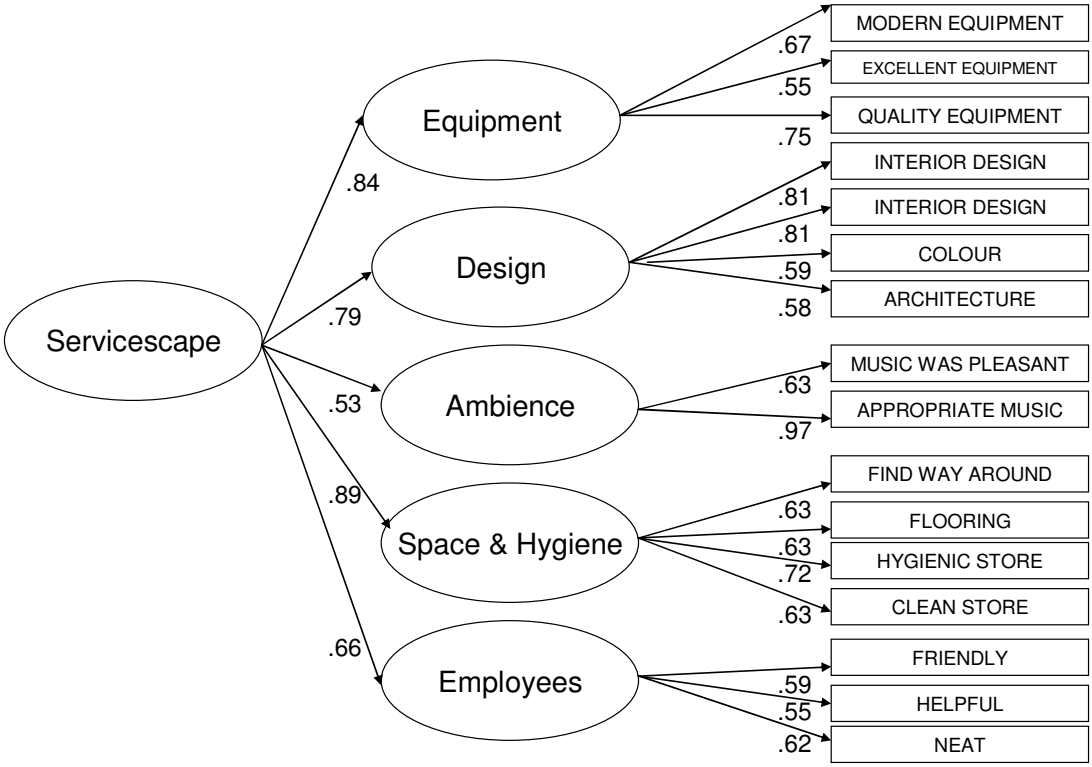


Figure 4.4: Re-specified Second Order Model of the Servicescape

4.10.3 Re-specification of the Servicescape Construct

The five factor re-specified structure for the 355 respondents were obtained. This model retains 17 of the 23 original items and obtains an excellent fit. Given the large sample size it is not surprising that a significant chi square was obtained (χ^2 = 205.15, df = 99, *p*

= .00) however, all other indices point towards a well fitting model (RMSEA = 0.05, CFI = .94, NNFI = .92, NFI = 0.89, IFI = .93, SRMR = 0.054). It is not recommended that all fit indices that are produced by the LISREL program are reported (Hooper et al, 2008) however for those wishing to see a more detailed exposition of the fit indices can refer to Appendix 4G. The parameters for the model (seen in Figure 4.4 and Table 4.12) were all strong and significant with standardised loadings greater than .50. As with the previous model, further tests of convergent validity were made using Fornell and Larcker's (1981) measure for Composite Reliability and their Average Variance Extracted formula. Details of these calculations can be found in Appendix 4H while the results are presented in Table 4.12 below. Bagozzi and colleagues (Bagozzi et al, 1991) recommend that composite reliability statistics must be above .60 if reliability is to be assumed. All of the constructs apart from Space and Hygiene were found to meet this criterion. However, the composite reliability for Space and Hygiene was .57 which is only marginally below the cut-off point. The results for AVE were less positive with only two constructs meeting the recommended threshold. As with the previous model, discriminant validity was assessed using Fornell and Larcker's (1981) stringent test. The calculations for these can be found in Appendix 4I. Discriminant validity was found between all constructs, apart from between the 'Equipment' and 'Space' dimension and also between the 'Space' and 'Design'. Given that Fornell and Larcker's (1981) test for discriminant validity is widely regarded as being very difficult to satisfy, further tests of discriminant validity were employed in order to determine whether these violations are to be considered worrying. The first of these was Bagozzi's, which states that discriminant validity has been found if the correlation between two constructs (phi φ) is significantly less than 1.0. The second

test of discriminant validity was to set the correlation between the two constructs of interest to 1.0 and to determine whether the model fit significantly improves. A significant improvement to the model would indicate that discriminant validity has not been met. In order to utilise these tests it was necessary model the two constructs in tandem using CFA.

The first model between Space and Design had a Phi (φ) of .34 with a standard error of .05. Using the formula described in Appendix 4D it was found that the Phi between these two constructs is significantly less than 1.0. Prior to fixing the correlations between the two latent constructs to 1.0 the model fit was χ^2 = 21.13 with 19 degrees of freedom. Setting the covariances to 1.0 severely effected the fit increasing the Chi square (χ^2) to 61.75 with 20 degrees of freedom. Both of these tests indicate that discriminant validity can be found for the two constructs and were therefore left as discrete dimensions.

Following this, Space and Equipment were also re-assessed using the above formulae. The initial CFA in which no constraints were placed on the intercorrelations between the two latent constructs had an excellent model fit: Chi-Square = 19.13, df = 13, *p* –value = 0.11920. When the two latent constructs were fixed to 1.0 the model fit significantly worsened with a Chi square difference of 47.29 with 1 degree of freedom. Similarly, an unconstrained model also produces a Phi (φ) of .35 with a standard error of 0.06 which also indicates discriminant validity between the two constructs (Bagozzi and Yi, 1991). As discriminant validity was found between all of these constructs they were all left as originally hypothesised.

Table 4.12: Results from the Second-Order Factor Analysis of Study Constructs

	Loading	t-value	Coefficient α	Composite Reliability	Average Variance Extracted
Ambience			.76	.66	.49
The background music was pleasant	.60	—[a]			
The background music was appropriate	.97	14.58*			
Design			.78	.80	.50
I found the interior design visually appealing	.81	10.87*			
The interior design was attractive	.81	10.86*			
The colour schemes were pleasant	.59	—[a]			
The architecture was attractive	.58	8.70*			
Equipment			.69	.70	.44
The equipment was modern looking	.67	10.37*			
The electronic equipment was excellent	.55	8.82*			
The equipment was of high quality	.75	—[a]			
Hygiene and Space			.73	.57	.43
The store was very clean	.63	9.35*			
The service station appeared to be hygienic	.72	10.20*			
The flooring was appropriate	.63	9.35*			
I found my way around quite easily	.63	—[a]			
Employees			.61	.61	.34
The employees were neat and tidy in appearance	.62	6.88*			
I found the staff friendly	.59	6.78*			
The employees were helpful	.55	—[a]			
Fit Statistics: χ^2 = 205.15, df = 99, p = .00, RMSEA = 0.05, CFI = .94, IFI = .94, NFI = .89 SRMR = 0.054.					

[a] = The corresponding coefficient was fixed to set the metric of the latent construct.

* = significant at the p = 0.01 level

Model Discussion

As mentioned above, all parameters in the above model were significant and loaded in the direction expected. In addition, the model fit was found to be acceptable, providing empirical confirmation that the servicescape consists of a multi-dimensional structure. As was expected, all of the servicescape subdimensions significantly contribute to consumers' overall perceptions of the servicescape. In particular it was found that the Space and Hygiene dimension explained more of the variance in the servicescape construct than any of the other constructs in the model. Shortly following behind this is the Equipment dimension which contributed .84 of the variance in the model. The constructs that contributed the least were Employees ($\gamma = .66$) and Ambience ($\gamma = .53$).

As this model achieved excellent fit, it indicates that a multidimensional model of the servicescape exists. This therefore supports Hypothesis 1 which was put forward earlier in this chapter. However, as the Hygiene and Space dimensions have not remained as separate dimensions, this part of the hypothesis remains unproven.

4.10.4 Alternative Conceptualisations of the Servicescape: Testing their Applicability

It became evident in Chapter One that there is very little consistency in the servicescapes literature. Each new paper in this area seems to herald a new alternative perspective on the servicescape which has led to a disjointed body of knowledge. The reason why this has occurred is not very easy to discern, particularly since both Bitner (1992; 2000) and Baker (1987; 2002) are clear as to what constitutes the servicescape. However, these two

authors differ greatly in terms of what they include under the servicescape umbrella which may have led to some of the confusion in the literature. Both authors describe the servicescape in terms of three dimensions, however the content of these sub-constructs are quite different (See Table 4.13 and Table 4.14). In terms of similarity, both authors describe ambient cues as a discrete dimension which encompasses all atmospheric stimuli that impact on any of the five senses. However, a point of difference does exist; Baker (1987) includes cleanliness under Ambient Conditions, a stimulus that is wholly excluded by Bitner (1992). The appropriateness of placing cleanliness under this dimension is unclear as it seems incongruent with what one would imagine would encompass ambient cues. Baker's (1987) second dimension is referred to as 'Design Factors' and includes both aesthetic and functional factors. This dimension is somewhat similar to Bitner's (1992) Space and Function dimension as it relates to the way in which furniture is positioned. The major point of difference is that Bitner (1992) does not regard the style of the furniture under her Space and Function dimension, merely she refers to the way in which the furniture is placed.

Table 4.13: Baker's Servicescape Dimensions

Ambient Conditions	**Design**	**Social Factors**
Temperature	Architecture	(Other Customers)
Air Quality	Colour	Number
Noise	Scale	Appearance
Scent	Materials	Behaviour
Cleanliness	Shape Style	(Service Personnel)
	Layout	Number
	Comfort	Appearance
	Signage	Behaviour

Table 4.14: Bitner's (1992) Servicescape Dimensions

Ambient Conditions	**Space/Function**	**Signs, Symbols & Artifacts**
Temperature	Layout	Signage
Air Quality	Equipment	Personal Activities
Noise	Furnishings	Style of Décor
Odour		

While Baker (1987) places the aesthetics of the décor under her Design dimension Bitner (1992) treats the style of the environment as a separate dimension which she refers to as Signs, Symbols and Artifacts. Along with the style of décor, this dimension refers to the signals that the environment communicates to customers about the place. The third of Baker's (1987) dimensions refers to the social factors and how they can motivate behaviour. This is possibly the largest point of difference between the two authors as Bitner (1992) excludes the impact of social cues entirely.

In order to resolve whether either of these two representations is a more accurate depiction of the servicescape construct structural equation modelling was used to analyse them both. The data that was used to meet this end is the same dataset that was used in the earlier analyses in this chapter. What differs here is that the conceptualisations tested reflect the models proposed by both Bitner (1992) and Baker (1987; 2002). This section begins by testing Baker's (1987) model which is followed by an analysis of Bitner's (1992) competing model.

Baker's (1987) Model

As per Baker's (1987) conceptualisation, a three factor model was run. This three factor model reflects Baker's (1987) original theoretical viewpoint in which the three dimensions are: Ambient Conditions, Design and Social Factors. Table 4.15 shows the indicators that were used to represent each of these dimensions. The fit for this model was particularly poor with a Chi Square (χ^2) of 597.94 with 149 degrees of freedom and a significant *p* value. The RMSEA was also well outside acceptable limits (RMSEA = 0.092) as were all other fit statistics, thus indicating that major misspecifications had been made (CFI = .81, IFI = 0.81, NFI = .76, SRMR = .073). Given that some of the items in the above conceptualisation were not used in the final servicescape representation, the model was re-run eliminating those items that had previously been found to be unreliable. This involved deleting three ambience items (The atmosphere was comfortable, The store had a pleasant smell, The lighting was comfortable), and three design items (The interior layout was pleasing, I found the physical facilities comfortable, The materials used were of high quality). Rather than improving the fit, it was found that when these items were deleted the fit worsened (χ^2 = = 344.24, df = 62, P-value = 0.00000, RMSEA = 0.113, CFI = .81, IFI = 0.81, NFI = .78, SRMR = .081). In addition to this, the ambience construct had a standardised loading greater than 1, both in the initial specification and in the second model. Standardised loadings greater than 1 generally indicate that multicollinearity exists which suggests a strong misspecification within the model.

Table 4.15: Items Used to Test the Applicability of Baker's Model

Ambient Conditions
The background music was pleasant
The lighting was comfortable
The atmosphere was comfortable
The store had a pleasant smell
The background music was appropriate
The store was very clean
The service station appeared to be hygienic
Design
I found the interior design visually appealing
The interior design was attractive
The colour schemes were pleasant
The materials used were of high quality
The architecture was attractive
I found the physical facilities comfortable
The flooring was appropriate
I found my way around quite easily
The interior layout was pleasing
Social Factors
The employees were neat and tidy in appearance
I found the staff friendly
The employees were helpful

Bitner's Model

While Baker's (1987; 2002) model has received some empirical testing, Bitner herself has never directly tested the applicability of her model. In later years, others have tested elements it, however, to date, no direct empirical testing of her model has been conducted. As with Baker's (1987) model in the above section, the data used to test the

fit of Bitner's (1992) model is the same as that used for the conceptualisation in this thesis. So as to reflect Bitner's model as best as possible the following items were used:

Table 4.16: Items Used to Test the Applicability of Bitner's Model

Ambient Conditions
The background music was pleasant
The lighting was comfortable
The atmosphere was comfortable
The store had a pleasant smell
The background music was appropriate
Space and Function
The flooring was appropriate
I found my way around quite easily
The interior layout was pleasing
Signs, Symbols and Artifacts
I found the interior design visually appealing
The interior design was attractive
The colour schemes were pleasant
The materials used were of high quality
The architecture was attractive
I found the physical facilities comfortable

As with Baker's model, the fit for Bitner's conceptualisation was particularly poor (χ^2 = 597.94 df = 61, P-value = 0.00000, RMSEA = 0.104, CFI = .85, IFI = .85, SRMR = .099, NFI = .82). It was also found that the Standardised Loading between the Servicescape and Design was greater than 1 also indicating that there was a major misspecification in the data. As some of the items in this model were found to be relatively unreliable the model was re-run excluding these items (The lighting was comfortable, The atmosphere

was comfortable, The store had a pleasant smell, The interior layout was pleasing, The materials used were of high quality, I found the physical facilities comfortable). Interestingly, this model produced an excellent fit (χ^2 = 21.86, df = 17, P-value = 0.19022, RMSEA = 0.028, CFI = 1.00, IFI = 1.0, NFI = .98, GFI = .98, SRMR = .024). Although the fit for this model is exceptionally good and has the advantage of being parsimonious it is seriously lacking in terms of its comprehensiveness. The items within this model relate only to the style of décor, the ease of navigating through the environment and consumers' perceptions of the music. Important servicescape stimuli such as the cleanliness of the environment and the equipment have been excluded. While statistically Bitner's (1992) model is superior to the servicescape representation outlined earlier in the chapter, it is limited in terms of its usefulness to practitioners as many key stimuli are omitted. In the servicescape model advocated by this research it was found that the Space and Hygiene dimension and the Equipment dimension contributed more of the variance in the servicescape than the other constructs in the model. This indicates that two important dimensions are entirely omitted from the Bitner (1992) model. Future researchers who choose to use Bitner's (1992) model over the model proposed and extended here will choose statistical power to the detriment of gaining a more complete overview of consumers' perceptions of the servicescape. In essence, choosing Bitner's model may obfuscate what truly lies beneath the servicescape construct.

4.10.5 Comparing the Three Models

In terms of fit, Bitner's (1992) model is by far the most superior of the four models (See Table 4.17 for a comparison of the models). However, the sole goal of this dissertation is

not to arrive at a servicescape conceptualisation that fits perfectly, rather the purpose of this thesis is to present the servicescape in a manner that is both useful in terms of ascertaining which aspects of the servicescape are more salient than others and also to demonstrate how each of the servicescapes' sub-dimensions interact with one another. In addition, in later chapters of this dissertation, this servicescape conceptualisation will be analysed in relation to a number of key post-consumption processes thus providing an even more thorough depiction of how the servicescape predicts consumer behaviour. It is also worthy to note that while Bitner's (1992) model is better in terms of fit, the model being proposed here is not an ill fitting model and is within acceptable limits across all key indices.

Table 4.17: A Comparison of the Three Servicescape Models

Fit Statistic	**Single Factor Model**	**Current Model**	**Baker Model**	**Bitner Model**
χ^2, df, and *p* value	1038.34, df = 100, p = 00000	205.15, df = 99, p = 0.00	344.24, df = 62, p = 0.00	21.86, df = 17, p = 0.19
χ^2/df ratio	10	2.07	5.55	1.28
RMSEA	0.10	0.05	0.113	0.028
CFI	0.76	0.94	0.81	1.00
NFI	0.71	0.89	0.78	0.98
IFI	0.77	0.93	0.81	1.00
SRMR	0.074	0.054	0.081	0.024

4.10.6 The Role of the Employee

In Chapter Two the theoretical differences between Bitner (1992) and Baker's (1987) models were briefly discussed. The key divergence between the two theorists stems from their placing of employees. The two theorists are not as clear in their position on the role of employees in the service environment. For instance, in Bitner's (1992) model, the servicescape does not include the employees as a stimulus, instead, the servicescape has

an effect on staff which subsequently manifests itself in various different approach and avoidance behaviours. Baker and a number of others have taken an altogether different view (Baker, 1987; Baker et al, 1994; Baker et al, 2002; Hightower et al, 2002; Ezeh and Harris, 2007; Harris and Ezeh, 2008) and have included employees' behaviour as part of the servicescape dimension. Theoretically, the placing of employees as part of the servicescape construct causes difficulties. When the content of Baker's (1987) employees items are examined, it is revealed that perhaps these items are not measuring tangible aspects of the store environment, but instead measuring some aspect of service quality. For example, two of the employees items, 'I found the staff friendly' and 'The employees were helpful' do not seem to be measuring the servicescape in any way but instead attempting to determine whether the staff provided quality service. This approach to formulating the servicescape seems somewhat incongruent with our understanding of the servicescape as one would expect that only atmospheric and tangible aspects of the physical environment should be included under a servicescape classification. This content validity problem becomes even more pertinent when we consider the service quality literature. Most proponents of service quality place heavy emphasis on the behaviour of employees. In fact, in Parasuraman et al's (1991) SERVQUAL scale all items apart from the tangibles items and item number nineteen (for a detailed account of the SERVQUAL items please see Appendix 5A) are dependent on the ability and the care provided by the staff. Similar service quality measures advanced by Brady and Cronin (2001) also cite the imperative role of the employee in overall service quality evaluations.

From a statistical viewpoint it is interesting to see that in the above analysis, Bitner's (1992) model significantly outperformed Baker's (1987). In examining the content of each of the models it seems that the major point of difference is that Baker (1987) includes an employees dimension whereas Bitner (1992) does not. It is possible that one of the reasons why Baker's (1987) model fit so poorly is due to the placing of employees as part of the servicescape construct. From this, it would seem that perhaps the two problematic employees items ('I found the staff friendly' and 'The employees were helpful') included in the current conceptualisation are best removed from the analysis. The third employees item, 'The staff were neat and tidy in appearance', does seem to relate to tangible aspects of the service environment and would theoretically fit well with the Hygiene and Space dimension. Therefore, a final model was run which incorporated these changes.

Final Re-specification of the Servicescape Model

The final four factor re-specified structure in which the employees dimension was omitted was again analysed using LISREL. Overall model fit for the four factor structure was excellent (χ^2 = 130.17, df = 73, p = .00004, RMSEA = 0.047, CFI = .96, NFI=0.92, NNFI = .95, IFI = .96, SRMR = 0.042). All parameters in the model were strong, significant and in the expected direction (see Figure 4.5 and Table 4.17). Both Composite Reliabilities and Average Variance Extracted were also calculated (See Table 4.17). Composite reliabilities were all above the .60 threshold while all but two of the constructs exhibited sufficiently high enough levels of Average Variance Extracted. Discriminant validity was also found between all constructs. This parsimonious model of

the servicescape demonstrates again that the servicescape can be conceived as a multi-dimensional structure in which the four sub-dimensions all contribute to consumers' overall perceptions of the servicescape. The fit of this model is superior to the last model in which the employees items were included thus suggesting that this final model is the superior from a statistical viewpoint. Theoretically this model is also a much sounder representation, as the omission of the employees dimension results in a servicescape construct that solely reflects the tangible and atmospheric facets of the service environment.

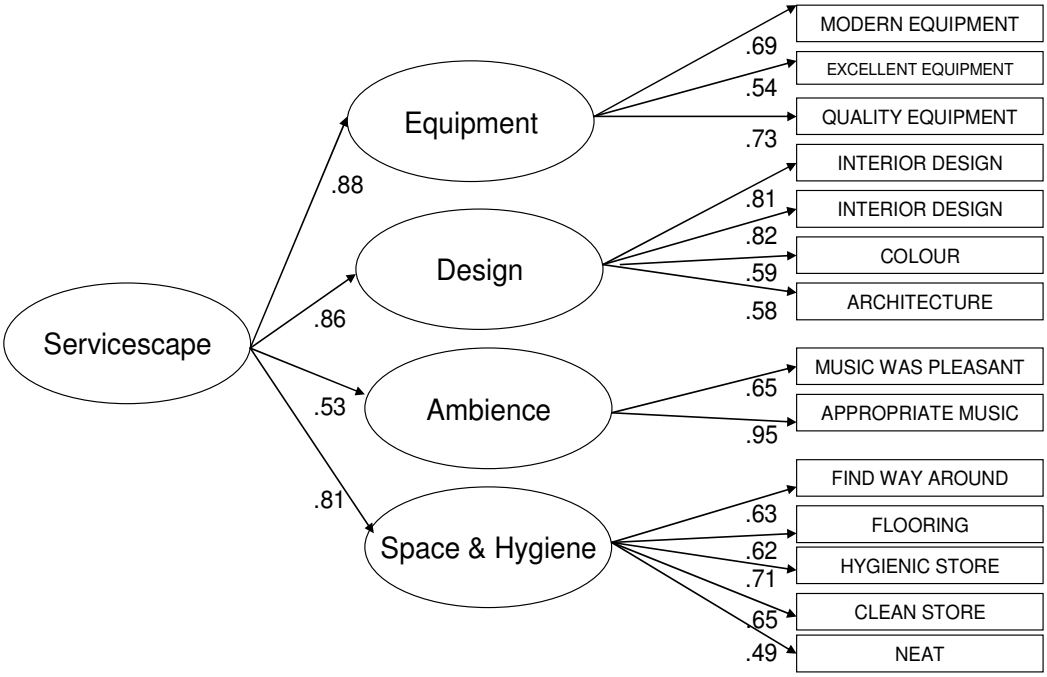


Figure 4.5: Final Servicescape Model

A Chi-Square difference test was also conducted between the above model and an unconstrained measurement model. The measurement model yielded a Chi-Square value

of 122.43 with 71 degrees of freedom. Thus, the overall fit of the second order model, with the additional paths is statistically better at the 1% level (130.17 - 122.43 = 7.74), therefore showing further support for the superiority of the second-order factor model (Kline, 2005).

In comparing this model with alternative conceptualisations such as Bitner (1992) and Baker's (1987) model it is found that Bitner's (1992) model still statistically outperforms this final model. However, as was mentioned earlier, the aim of this dissertation is not to allow statistics to drive the analysis but instead to allow theory and substantive issues to guide the process. Therefore, while Bitner's (1992) model is slightly superior with regards to fit, it is severely limited in terms of its content with only a small number of servicescape issues addressed.

Table 4.18: Comparing the Fit of the Various Models

Fit Statistic	**Single Factor Model**	**Final Model (excl. employees)**	**Baker Model**	**Bitner Model**
χ^2, df, and *p* value	1038.34, df = 100, p = 00000	130.17, df = 73, p = .00004	344.24, df = 62, p = 0.00	21.86, df = 17, p = 0.19
χ^2/df ratio	10	1.78	5.55	1.28
RMSEA	0.10	0.047	0.113	0.028
CFI	0.76	0.96	0.81	1.00
NFI	0.71	0.92	0.78	0.98
IFI	0.77	0.96	0.81	1.00
SRMR	0.074	0.042	0.081	0.024

Table 4.19: Final Servicescape Model (excluding Employees)

	Loading	t-value	Coefficient α	Composite Reliability	Average Variance Extracted
Ambience			.76	.66	.66
The background music was pleasant	.65	—[a]			
The background music was appropriate	.95	7.28*			
Design			.78	.80	.50
I found the interior design visually appealing	.82	10.89*			
The interior design was attractive	.81	10.83*			
The colour schemes were pleasant	.59	—[a]			
The architecture was attractive	.58	8.73*			
Equipment			.69	.70	.44
The equipment was modern looking	.69	10.67*			
The electronic equipment was excellent	.54	8.65*			
The equipment was of high quality	.73	—[a]			
Hygiene and Space			.75	.76	.39
The store was very clean	.65	9.52*			
The service station appeared to be hygienic	.71	10.07*			
The flooring was appropriate	.62	9.61*			
I found my way around quite easily	.63	—[a]			
The employees were neat and tidy in appearance	.49	7.66*			
Fit Statistics: χ^2 = 130.17, df = 73, p = .00004, RMSEA = 0.047, CFI = .96, NNFI = .95, NFI = .92, IFI = .96, SRMR = 0.042					

[a] = The corresponding coefficient was fixed to set the metric of the latent construct.

* = significant at the p = 0.01 level

4.11 Conclusions

Disagreements over the factor structure of the servicescape, along with inconsistent and poor measurement have hindered practical and theoretical advancements in the literature. The purpose of this chapter was threefold; 1) to explore the theoretical dimensionality of the servicescape, 2) to compare the current conceptualisation with alternate servicescape models, and 3) to perform construct validation of the new servicescape measure developed in Chapter Two. To meet this end, structural equation modelling was used to link theory construction and testing.

According to the criteria for goodness-of-fit that was imposed on each of the models, the present study found no empirical support for a unidimensional servicescape model. In contrast the multi-dimensional models were found to provide a significantly better fit. These models comprise a number of discrete servicescape elements and demonstrates how they work together to form consumers' overarching perceptions of this construct. The model proposed by the current research is relatively parsimonious, constituting a four factor structure. Surprisingly, the analyses suggest that statistically Bitner's (1992) model is slightly better than the final model. However, it is argued that despite the slightly worse fit of the new servicescape model, it is theoretically and practically more sound as it informs practitioners and academics alike on a wider range of servicescape dimensions.

Overall support was found for the hierarchical, second-order servicescape model compared with the single-order model. Therefore, at the theoretical level, these findings

were consistent with an underlying hierarchy of servicescape elements which is demonstrated in the acceptable fit of the second order model. Some have considered the Space and Hygiene dimensions to be separate entities (Wakefield and Blodgett, 1996), however the results suggested there that collapsing these two dimensions together provides an improved fit. At first this seemed unusual however upon further scrutiny it became quite logical as it can be justified that consumers' perceptions of the amount of room in the environment are likely to be linked with hygiene as cleaner environments create a greater feeling of space and tend to feel less cluttered and cramped. The most interesting finding was that the more utilitarian dimensions within the servicescape construct were found to contribute most of the variance in the model with the more afferent dimensions playing a much lesser role.

Apart from the theoretical implications of the results, the analysis conducted in this chapter point to several strengths of the servicescape measure developed here versus others currently used in the literature. This measure was capable of discriminating between the four servicescape constructs, indicating that consumers perceive each of the sub-dimensions as separate from one another while still tapping into the overall servicescape construct. Future researchers may choose to use scale composites or parcels as the fit for this model has been demonstrated to be very good. Further research may also seek to employ this measure to ascertain whether the role of equipment and hygiene will dominate the factor structure as found with this particular service context.

Classifying schema is important to the development of an area, because by organising phenomena into classes or groups they become more amenable to systematic study and theory development (Baker, 1987). This has certainly been the situation here as by deconstructing the servicescape into separate dimensions has allowed us to determine which of these servicescape elements contributes most of the variance in the construct. The next stage of this dissertation aims to extend the current research by firstly analysing how this construct relates to respondents perceptions of service quality. Following this, the construct will be examined in conjunction with a number of key post-consumption variables which will both assess the construct's nomological validity and also to describe how it predicts consumer behaviour.

5.0 SERVICE QUALITY

5.1 Introduction

In the previous chapter the servicescape construct was explored in detail. Using structural equation modelling, the multi-dimensional nature of this construct was assessed. This chapter moves on from the servicescape and focuses on the service quality construct. Within this chapter the literature on this important topic is reviewed. The chapter will begin by defining service quality. The two major streams of research are examined, these being the American school of thought and the Scandinavian approach. The differences in these models are assessed and more current thinking on this subject is introduced. It is from this more recent research that a two-dimensional view of service quality is presented. This two dimensional model is then tested using structural equation modelling and it is subsequently modelled as being preceded by the servicescape. From this analysis a number of interesting findings are gleaned and their implications for this dissertation are discussed.

5.2 Defining Quality

Service quality is widely regarded as being elusive and difficult to define (Parasuraman et al, 1988). Definitions of this construct vary widely ranging from value based definitions, quality vis a vis excellence, quality as a manufacturing requirement, and quality as an outcome of meeting or exceeding customer expectations (Levitt, 1972; Parasuraman et al, 1985; Parasuraman et al, 1994; Reeves and Bednar, 1994; Reeves and Bednar, 1995; Wetzels, 1998). The outcome definitions of service quality views the construct as an

overall appraisal of a product or service that is dependent on consumers prior expectations (Grönroos, 1984; Bitner and Hubbert, 1994). This consumer-centric definition is the one that seems to prevail most commonly in the literature and has now become the academic standard (Reeves and Bednar, 1995). It is also this definition that forms the basis of much of the service quality literature and has now become the most widely agreed definition of service quality.

5.3 Two Schools of Thought: The Scandinavian versus the North American Tradition

Within the services quality literature, two complementary streams of research have evolved. These two traditions can be broadly categorised as those being of either the Scandinavian or American tradition. Both of these traditions agree that consumers arrive at an evaluation of service quality that is based upon disconfirmation theory (Ekinci et al, 1998). This being that prior to consuming a service, consumers hold preconceived ideas of how the service will perform. Once the consumer has experienced the service, they compare the performance of the service to their *a priori* expectations in a subtractive manner to determine the level of perceived service quality (Lehtinen and Lehtinen, 1991). The above definition is employed by both of these perspectives, however they differ in terms of how they conceptualise finer elements of the construct.

Within the Nordic stream of research, Grönroos (1984) has been one of the most prolific contributors to our understanding of the service quality concept (Brogowicz et al, 1990). Grönroos (1984) proposes that service quality can be described as a two factor structure

comprising of both functional and technical elements. The functional element relates to the way in which the service is delivered, while the technical element refers to what the consumer receives from the service (Brogowicz et al, 1990). This functional aspect of service delivery has also been referred to as peripheral to the process while the technical element conceptually constitutes the core or outcome components of the service delivery process (Tripp and Drea, 2002). Given the higher level abstraction associated with functional quality dimensions, it is thought that their evaluation is much more subjective (Hill, 1986; Caceres and Paparoidamis, 2005). Walker (1995) also discusses this and gives the example of how comfortable your night's sleep in a hotel as the core (technical) service, with supplemental components such as the pleasantness of the staff encompassing the peripheral (functional) aspects of service quality. Moderating the two quality dimensions is company image which helps consumers form expectations as to the standard of service which they can expect from the company (Grönroos, 1984).

Writings on this model of service quality have been largely theoretical (Ekinci et al, 1998), however, in recent years a number of authors have sought to link technical and functional quality dimensions to a variety of constructs such as trust, commitment, satisfaction and loyalty (Lassar et al, 2000b; Lassar et al, 2000a; Caceres and Paparoidamis, 2005) and general support has been found for the two dimensional conceptualisation of service quality.

The second, more widely cited service quality approach is that of Parasuraman and his colleagues (Parasuraman et al, 1985; Parasuraman et al, 1988; Parasuraman et al, 1991b;

Parasuraman et al, 1994; Zeithaml et al, 1996). Parasuraman et al (1985) concur with Grönroos (1984) that service quality can be conceived in terms of the difference between expectations and performance, however they build on this line of work through the development of their 'Gaps Model' (See Figure 5.1). From a number of focus group and executive interviews, five 'gaps' that can cause quality problems for organisations were identified. The first gap is the consumer-management perceptions gap. This gap is the difference between what management regard as essential to achieving high standards of service quality and what consumers believe is important to them. The second gap, known as the management perception-service quality specification gap refers to operational difficulties that were encountered by management in trying to implement service quality standards. The third gap, known as the gap between service quality specifications and service delivery gap, is quite similar to the second gap in that it too relates to operational difficulties. This gap relies heavily on the influence of employees and how individuals' variability can lead to fluctuations in service delivery standards. The fourth gap is described as the service delivery-external communications gap. In short, this relates to how and what the firm communicates to prospective and current customers on the level of service provided. Failure to meet promises made in advertising communications creates this gap in what the customer expects to receive and what they receive in reality. The final gap (Gap 5) has been termed the expected service-perceived service gap and is probably the most frequent gap that service firms fail to bridge. Drawing on disconfirmation theory (Oliver, 1980), this gap will occur in any situation where a consumer is dissatisfied with the service they receive. Conversely, if a consumer

receives over and above what they had expected this gap will also occur except that in this situation it is a positive rather than a negative gap.

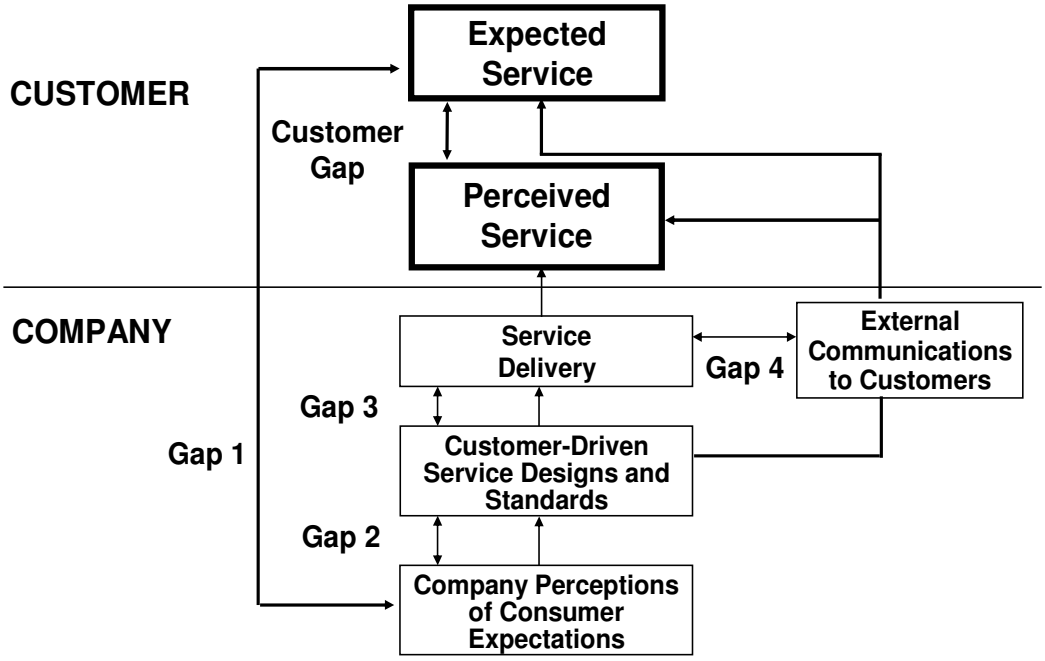


Figure 5.1: Service Quality Gaps Model (Parasuraman et al, 1985, p.44)

This final gap in the model (Gap 5) has been stressed as being of utmost importance and forms the basis of Parasuraman et al's (1985) Service Quality Model.

5.4 The SERVQUAL Model

The initial Service Quality Model (Parasuraman et al, 1985) proffered that there were ten dimensional antecedents of service quality which were described as: reliability,

responsiveness, competence, access, courtesy, communication, credibility, security, understanding and tangibles. In a later study (Parasuraman et al, 1988) it was found that some of the initial ten dimensions were highly correlated with one another which resulted in a refined five dimensional model which consisted of assurance, empathy, reliability, responsiveness and tangibles (See Table 5.1). These five dimensions were uncovered as part of their detailed scale development procedure which resulted in the SERVQUAL scale (Parasuraman et al, 1988) which has since been widely used by both academics and practitioners.

Table 5.1: Service Quality Dimensions

Assurance	Employees' knowledge and courtesy and their ability to inspire trust and confidence
Empathy	The caring, individualised attention a firm provides its customers
Reliability	A firm's ability to perform the promised service dependably and accurately
Responsiveness	Employees' willingness to help customers and provide prompt service
Tangibles	A firm's physical facilities, equipment, personnel and communication materials

Source: *Adapted from Parasuraman et al (1988*)

This scale contains 22 pairs of Likert-type items (see Appendix 5A), 22 of the items measure expectations with the remaining 22 intending to measure the perceived level of service provided by the organisation (Babakus and Boller, 1992). For example, expectations are measured in the following format, 'they should have up to date facilities' with and perceptions phrased in the following way, 'XYZ has up to date facilities. From this, service quality is measured by subtracting the difference between the expectation score and the perceived service quality score which is then used to determine the

respondents' service quality perceptions of each dimension. Following this, using all respondents scores, an average score for each dimension is calculated. Positive scores indicate that respondents' expectations were exceeded whereas negative scores indicate poor service quality with expectations not being met. An average score of zero means that quality is satisfactory, however given the number of questions in the scale it is doubtful that this would ever occur.

The SERVQUAL scale has been commended for its intuitive appeal and on its ability to diagnose service quality problems. However, despite the widespread acknowledgement this instrument has received, it has been subject to substantial criticism and debate. Criticisms of this measure relate primarily to the construct's dimensionality and problems associated with the use of an expectations measure.

5.4.1 Criticisms of the SERVQUAL Instrument

In relation to the constructs' dimensionality, there is disagreement with regards to both the number and the content of the service quality dimensions. While Parasuraman et al (1988, p. 24) believe that this scale has 'sound and stable psychometric properties' attempts to replicate the dimensional solution found by Parasuraman and his colleagues has proven quite difficult. Those who have made verbatim attempts at reproducing the five-dimensional factor structure found by Parasuraman et al have generally failed to do so (Finn and Lamb, 1991; Cronin and Taylor, 1992). The number of dimensions found by authors seems to vary widely from one (Brown et al, 1993; Mels et al, 1997), two (Babakus and Boller, 1992; Smith, 1999) three (Durvasula et al, 1999) five (Carman,

1990), to as many as nine (Carman, 1990) dimensional structures found. The dimensions have also been criticised for being overly generic (Babakus and Boller, 1992; Dabholkar et al, 2000) and not being specific enough for various different services.

The calculation of gap scores also been challenged by a number of authors (Gounaris, 2005). Issues that are raised generally relate to reliability and discriminant validity problems (Cronin and Taylor, 1992; Brown et al, 1993; Cronin and Taylor, 1994). Brown et al (1993) argue that low reliabilities are almost impossible to avoid when using expectations standards as positive correlations between components of the score attenuate the reliability of the resulting difference score. In addition to this, low measure reliability attenuates correlations between constructs thus suggesting that discriminant validity exists however this is merely an artefact of achieving low reliability (Buttle, 1996). A second discriminant validity issue is that difference scores are generally highly correlated with one of the component measures used to obtain the difference thus calling into question the construct's validity. Nomological validity is also a problem with SERVQUAL, with SERVPERF (a performance only model) outperforming SERVQUAL as a predictor of behavioural intentions.

Operational issues can also hinder the quality of the data collected using difference scores (Buttle, 1996). As per Parasuraman et al's (Parasuraman et al, 1988; Parasuraman et al, 1991a; Parasuraman et al, 1994) recommendations, both the expectations and perceptions measures are administered post-consumption. The likelihood is that respondents cannot accurately report their prior expectations as they will be tainted by their experience.

Clow and Vorhies (1993) demonstrated this in a longitudinal study where it was found that if consumers are asked after a negative experience what they expected before the experience more than 50 per cent of the consumers answers will be different. A similar, but not as strong effect was found in the positive group leading Clow and Vorhies (1993) to conclude that expectations should be measured prior to consumption. The implementation of this may prove to be very difficult as it is unlikely that respondents will be willing to participate in two survey administrations (Robinson, 1999). Indeed, even in situations where respondents are willing to complete two surveys the likelihood is that their evaluation of the service experience is not wholly accurate as demand artefacts will influence their responses.

These criticisms of the SERVQUAL measure have led to a reduction in its use in recent years. It also seems that some of the original authors of the model have conceded that there may be some overlap between the items measuring responsiveness, assurance and empathy (Parasuraman et al, 1994). In an article which was co-authored with a number of others, Zeithaml, also makes the admission that the expectations view of measuring service quality is perhaps not appropriate which is demonstrated in the following quote: 'Our results are incompatible with both the one-dimensional view of expectations and the gap formation for service quality. Instead we find that service quality is directly influenced only by perceptions of performance' (Boulding et al 1993, p. 24). Even a recent electronic service quality measure developed by Parasuraman et al (2005) has excluded an expectations measure. This thus suggested that perhaps the original developers of this scale now agree that the expectations measure is flawed. Further to

this, it also implies that they do not believe that the SERVQUAL measure is a generic scale which can be adapted to a broad range of contexts.

5.4.2 Alternative Perspectives on Service Quality

Given the many criticisms levelled at the SERVQUAL measure, numerous theorists have sought alternative methods of measuring this important construct. The first of these measures is Cronin and Taylor's (1992) SERVPERF scale. This 21-item scale uses all but one of the performance measures proposed by Parasuraman et al's (1988) SERVQUAL scale. The major advantage of this scale is that it does not contain an expectations measure and therefore does not suffer from the various problems associated with it. In fact in studies where SERVQUAL has been compared with SERVPERF it has been found that SERVPERF outperforms the former measure (Cronin and Taylor, 1992; Durvasula et al, 1999). There has also been growing support for service quality models that are based upon Grönroos's (1984) model (Rust and Oliver, 1994; Brady and Robertson, 2001). The first of these is Rust and Oliver's (1994) theoretical article which proposes that service quality consists of three elements of service quality, namely, the service product, the service delivery and the service environment. The service product refers to what is to be delivered and the standards to which the service is to be delivered. This is akin to Gronroos's (1984) definition of technical quality which involves *what* is provided. The second of Rust and Oliver's (1994) dimensions refers to the service delivery and relates to how the employees engage in service performances. This dimension is heavily dependent on the role of the employees and the manners and behaviour of the personnel involved in the delivery of the service product (Lassar et al,

2000b). The final of Rust and Oliver's (1994) dimensions includes two main themes, the internal environment and the external environment. The internal environment deals with the structures that are put in place by the organisation that strive towards providing quality service. Included in this category are new customer generation strategies, customer retention initiatives and various different internal marketing activities. The second sub-dimension within the service environment is similar to Bitner's (1992) servicescape model and includes environmental dimensions such as the ambience, space and symbolic elements.

While Rust and Oliver's (1994) article was wholly theoretical, an attempt to empirically test and expand it was made by Brady and Cronin (2001). This study modelled service quality as a hierarchal structure in which service quality consists of three sub-dimensions: interaction quality, service environment quality and outcome quality. Empirical support was found for the model and it has provided support for both Gronroos's (1984) and Rust and Oliver's (1994) theoretical models. Other departures from the SERVQUAL model have tended to focus on Gronroos's conceptualisation (Lassar et al, 2000b; Lassar et al, 2000a)and it has been found that his technical/functional conceptualisation performs better than the SERVQUAL measure.

A more recent and conceptually very similar conceptualisation is proffered by Berry (Berry et al, 2002; Wall and Berry, 2007). In the first of two papers, Berry et al (2002) do not explicitly refer to service quality, but instead propose that practitioners need to manage the 'customer experience'. In order to manage the total customer experience,

Berry et al (2002) advocate that all aspects of the service delivery process needs to managed. In particular, companies need to pay attention to various different 'clues' which when combined contribute to the customer's overall evaluation of the service offering. These clues are categorised as being either functional, humanic or mechanic. Confusingly, Berry et al's (2002) functional clues are very similar to Gronroos' technical dimension (Gronröos, 1984) or the outcome dimension proposed by Brady and Cronin (2001) and tend to relate to the way the core service was delivered. Functional type issues would be whether the airplane brought you to your destination on time or whether a hotel had a comfortable bed in which to sleep. Humanic cues (Berry et al, 2002) are the cues emitted by people and determine the interpersonal relationships in the buying experience (Carbone and Haeckel, 1994). The mechanics cues are easier to manage and involve the manipulation, design and control of all aspects of the physical environment. In many ways these two dimensions are similar to both Brady and Cronin (2001) and Oliver and Rust's (1994) three-dimensional view of service quality. Using an experimental design, Wall and Berry (2007) tested how the humanic and mechanic aspects of a restaurant contribute to overall service quality perceptions and it was found that the human dimension played a more significant role in how consumers evaluated the service offering.

5.5 The Servicescape/Service Quality Overlap

When reviewing the literature on service quality it becomes strikingly obvious that many of the service quality theorists subsume the service environment under the service quality umbrella in some form or another. Parasuraman et al (1985; 1988; 1991a; 1994;

1994)include a tangibles dimension within their framework which incorporates the appearance and design of the physical facilities, the appropriateness of the equipment used and also how the employees are presented. Similarly, Brady and Cronin (2001), Rust and Oliver (1994) and more recently Wall and Berry (2007) all describe how service quality can be inferred from the quality of the service environment. For the current study this poses a particular problem, as it becomes unclear whether the servicescape should also be postulated as a service quality dimension. Upon scrutinising Bitner's (1992) original work within which the servicescape construct was introduced, it becomes even more unclear as to where this construct should be placed. Bitner (1992) makes no explicit references to any service quality model, instead she notes that 'customers…tend to use extrinsic cues (such as the physical surroundings) to infer quality' (p. 63). From this it can be seen that Bitner (1992) believes that the servicescape contributes to service quality but how this should be modelled is unclear. If we look to the major studies on the servicescape(Baker et al, 1994; Wakefield and Blodgett, 1994; Wakefield and Blodgett, 1996; Wakefield and Baker, 1998; Wakefield and Blodgett, 1999; Harris and Ezeh, 2008) it can be seen that little research has been conducted which integrates or tests the relationships between the servicescape and service quality. As part of a larger model Hightower et al (2002) found that the servicescape is a predictor of perceived service quality, however, the theoretical implications of this were not discussed. One study which somewhat addresses this issue is Reimer and Kuehn's (2005). This study presented an alternative view of the service quality concept which places the tangibles dimension of the SERVQUAL scale (Parasuraman et al, 1988; Parasuraman et al, 1991a; Baker et al, 1994) as an antecedent of the remaining four intangible dimensions

(reliability, responsiveness, assurance and empathy). However, the authors encountered difficulties in finding a multiple factor structure for service quality and subsequently merged all intangible service quality dimensions to form a single intangible construct. Using path analysis, the relationship between tangibles and intangibles and how they impact on overall service quality was examined. As per their hypotheses, it was found that the servicescape both directly and indirectly impacts overall service quality perceptions through the intangibility dimension. This study by Reimer and Kuehn (2005) provides evidence that the servicescape is an antecedent to the remaining employee related aspects of service quality and it is from this that the following hypothesis is borne:

H1: *Service quality is predicted by the four servicescape dimensions, namely, Design, Hygiene and Space, Equipment and Ambience.*

5.6 A Two Dimensional Conceptualisation of Service Quality

A second issue which this chapter seeks to address is to present a two dimensional structure of service quality. This proposed two dimensional structure borrows from the work of Grönroos (1984), Brady and Cronin (2001) and Wall and Berry (2007). Grönroos's (1984) original model describes service quality as consisting of two dimensions, the technical and functional service quality dimensions. The functional dimension relates to *how* the service is delivered whereas the technical aspect refers to what the customer actually receives. More recently, Brady and Cronin (2001) built upon this, through their description of a three factor model which extends Gronroos's model through the inclusion of a physical environment dimension. Finally, a theoretically

similar view is presented by Wall and Berry (2007) who promote a three dimensional model which includes a human, mechanic and functional dimension. Of these authors, all but Grönroos (1984) describe service quality as containing three dimensions, one related to the physical environment, the second relating to human and social interactions and the third relating to the outcome of the service situation. As mentioned above, it is believed that the first of these constructs, the servicescape or physical environment will precede service quality evaluations. The removal of the servicescape from the service quality construct leaves two sub-dimensions remaining, namely, the service outcome and the interactional or social element. Therefore, it is proposed that service quality will consist of two discrete dimensions which is reflected in the following hypothesis:

H2: *Service quality construct will consist of a two factor structure, one relating to the employees and the second relating to overall service quality evaluations.*

5.7 Measurement of Service Quality

In order to test these hypotheses it was necessary to choose an appropriate service quality measure. The literature is replete with various different service quality batteries. The original of these is the aforementioned SERVQUAL scale which has many problems inherent in its design. A number of authors have used this scale in the past with limited success (Finn and Lamb, 1991; Babakus and Boller, 1992; Cronin and Taylor, 1992; Gagliano and Hathcote, 1994). The most widespread and serious criticism aimed at this scale is the use of the expectations measure. In an attempt to overcome this fault Cronin and Taylor (1992) developed the SERVPERF measure which has been used by a number

of authors (Cronin and Taylor, 1992; Gagliano and Hathcote, 1994; Knowles et al, 1999; Gounaris, 2005). The SERVPERF scale is identical to the SERVQUAL measure in that it uses all the performance items from the scale but excludes the expectations items. As the content is more or less the same, SERVPERF also suffers from the disadvantage of being overly generic and lacking a consistent dimensional structure. These problems have been overcome by researchers through the inclusion of additional variables(Dabholkar et al, 2000; Sureshchandar et al, 2002; Wong and Sohal, 2003; Choi et al, 2004; González et al, 2007) or through the removal of certain items. Semantic differentials based on a shortened version of SERVPERF have also been used in the literature (Gotlieb et al, 1994; Voss et al, 1998; Brady and Cronin, 2001; Brady and Robertson, 2001). Quite a lengthy scale has also been developed by Brady and Cronin (2001) which attempts to reconcile the problems associated with both SERVQUAL and SERVPERF however this does not seem to have moved authors away from these scales and no replications of this scale seem to have appeared in the literature (Carrillat et al, 2007). In fact, subsequent papers by Brady (Brady and Robertson, 2001; Hightower et al, 2002; Brady et al, 2005) have not used this scale and have used alternative service quality measures which in some way indicates that perhaps this scale suffers from some problems. Others have used performance based scales that are based upon SERVQUAL and capture the key dimensions of the construct (Taylor and Baker, 1994; Grace and O'Cass, 2004). Even more recently, many authors have also taken to designing their own service quality measures that are constructed for the specific study at hand (Anderson and Fornell, 1994; Ekinci, 2001; Aydin and Özer, 2005; Caceres and Paparoidamis, 2005; Ibanez et al, 2006; Pei, 2008).

5.7.1 Measuring Service Quality in Service Stations

For the current research problem a number of issues needed to be considered in deciding which service quality scale to use. As many of the problems relating to the SERVQUAL scale have not yet been resolved (Hartline and Ferrell, 1996) it makes choosing an appropriate measure very difficult. Overall the consensus seems to be that SERVQUAL and the expectations-minus-performance measure is inappropriate from both a psychometric and operational viewpoint and was thus discounted as a possible alternative. SERVPERF, is a performance based version of SERVQUAL in that it excludes the expectations measure was a possible contender, however its length (22 items) makes it a somewhat cumbersome choice. This is particularly pertinent given the large number of other constructs that were measured in this administration (the servicescape, behavioural intentions, customer satisfaction, emotions and some demographic information). Given that length is an issue here it also eliminates Brady and Cronin's (2001) 35 item service quality measure.

Semantic differentials have also been used however, it was decided against using a semantic differential format for this question as it has been found previously that this format of scale can be difficult to administer and involves more complex analysis on the part of the respondent which reduces their reliability (Stephens, 1991). Developing a service quality scale to suit this particular service was also not appropriate for a number of reasons. Firstly, the psychometric properties of such a scale would be unknown prior to its usage and could result in a scale with limited reliability and validity. Secondly, it is

hoped that in the future similar studies will be conducted in varying contexts which means that a relatively general scale is needed which would translate well to other service situations. A number of relatively short, performance only service quality scales were possible alternatives given their relatively general wording. Both Hartline and Ferrell (1996) and Brady et al (2005) used almost identical performance only items that are shortened versions of the SERVPERF scale. The items from these scales were examined, however it was decided that these scales are not suitable for the current context for a number of reasons (these are presented in Table 5.2 below). The first issue that is immediately obvious is that the items all relate solely to the behaviour of the employees. This is a problem as it proposed that the service quality construct will consist of two dimensions, one relating to the employees and the second an outcome measure of the service received.

Table 5.2: Scale Items used by Hartline and Ferrell (1996) and Brady et al (2005)

Their employees offer the personal attention I need from them
The behaviour of their employees instills confidence in me
Their employees are courteous
I receive enough individual attention from their employees
I can depend on receiving prompt service from their employees
I feel safe conducting business with their employees
Their employees are able to answer my questions
Their employees are never too busy to respond to my requests
Their employees have my best interests at heart
Their employees understand my specific needs

A second problem with these items is that they are unlikely to transfer well to a service station context. For instance, questions such as, 'Their employees are able to answer my questions', 'Their employees have my best interests at heart', and 'Their employees

understand my specific needs' seem more appropriate for more extended service situations or for professional service situations than for service stations.

For the above reasons, these SERVPERF variation scales were not chosen. The scale finally chosen to measure service quality was a scale adapted from a study conducted by Grace and O'Cass (2004). In their original study, convergent and discriminant validity was found for the construct and was also found to have stable psychometric properties in a follow-up study by the same authors (Grace, 2005). Slight adjustments were made to their core service measure which was chosen to represent outcome service quality. The first change that was made was to drop the word 'core' from the 'Good core service' and 'Quality core service' items. It was felt that the word 'core' is somewhat ambiguous and may confuse respondents. The second change involved lengthening the items so that they formed full statements (in Grace and O'Cass's (2004) paper the items were abbreviated when reported). Finally, the item 'The service was of a very high quality' was added to the item pool to ensure validity. All items used to measure the outcome dimension of service quality can be viewed in Table 5.3 below:

Table 5.3: Service Quality Outcome Items

The store provided good service
The service was of a very high quality
The service here suited my needs
The service here is reliable
The service station provided quality service

The remaining service quality items relate solely to the behaviour of the employees and were expected to form a unique dimension of their own. These items were again taken

from Grace and O'Cass's (2004) study however further modifications were necessary. Of the seven original items used by Grace and O'Cass (2004) all but one, 'Feel safe in transactions', were thought to be suitable. The service chosen by Grace and O'Cass (2004) was banking and is an obvious choice of item given the context where large amounts of money change hands and where personal information is disclosed. However in the current situation this is not likely to be an issue. As with the previous items it was also necessary to create full statements for each of the items as they were again reported by Grace and O'Cass (2004) in an abbreviated format. There were also two 'helpfulness' type items in Grace and O'Cass's (2004) scale which were, 'Willing to help', and 'Gives personal attention'. Rather than using both of these it was decided to just use one helpfulness item which was phrased, 'The staff were helpful'. The 'gives personal attention' item is not likely to be that relevant to the current context. The 'never too busy for me' item was also changed to 'the service was efficient'. Other slight rewordings were made which are evident upon examination of Table 5.4 below.

Table 5.4: Original Grace and O'Cass (2004) Employees items and Final Items Used

Original Grace and O'Cass Items	**Final Items**
Provides prompt service	The service in store was delivered promptly
Willing to help	The staff were helpful
Never too busy for me	The service was efficient
I can trust employees	The staff were trustworthy
Feel safe in transactions	**----------**
Employees are polite	The staff were polite
Gives personal attention*	The staff were helpful*

* 'Gives personal attention' was felt to be slightly superfluous so was decided to use 'The staff were helpful' to represent this.

The format of the items was the same as that used for the servicescape with all items using a Likert style design with 7-point scales anchored with 'Strongly Disagree (1) to Strongly Agree (7).

5.7.2 The Sample and Data Collection

The sample was drawn from the same pool of respondents that were used in the analysis of the previous chapter. The respondents were intercepted as they exited the store of the service station from which a total of 355 useable questionnaires were collected.

5.7.3 Assessment of the Measures using Confirmatory Factor Analysis

The first of the above hypotheses proposes that the servicescape will be an antecedent of service quality. By removing the servicescape from the service quality construct, two service quality dimensions remain, namely the Interactional/Employees dimension and an Outcome dimension. Prior to assessing whether the servicescape is a direct antecedent of the service quality construct it is necessary to assess the service quality measurement model. As discussed above, it is hypothesised that the service quality construct will consist of two distinct dimensions. Figure 5.2 below is a graphical representation of the hypothesised service quality structure:

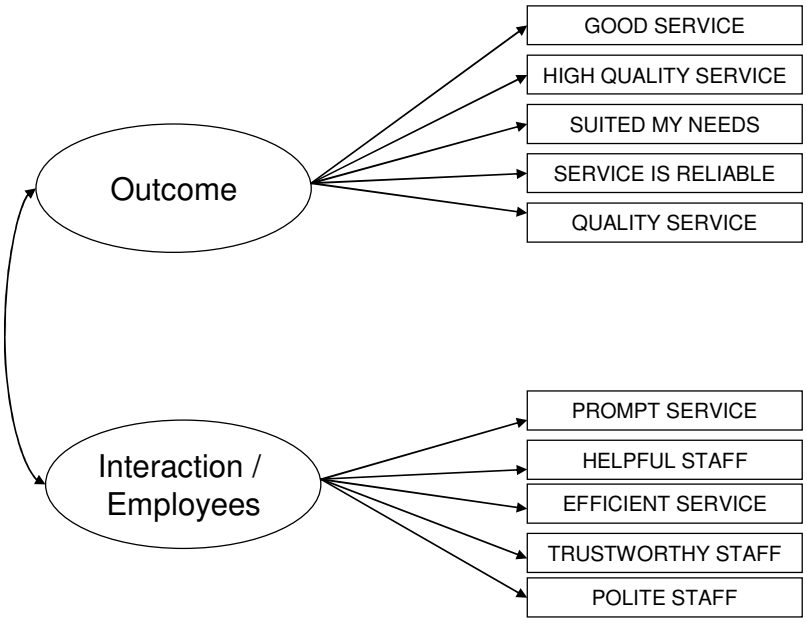


Figure 5.2: Proposed Measurement Model for the Service Quality Construct

In line with Anderson and Gerbing's (1988) recommendations, a two-step approach was taken. The Interaction/Employees dimension was analysed first. A five item confirmatory factor analysis of the construct revealed quite a poor fit (Chi-Square = 35.48, df = 5, P-value = 0.00000, RMSEA = 0.131, NFI = 0.96, CFI = 0.96). Multiple squared correlations of all items were not problematic, however, large error variances were found between the 'Efficient service' item and 'The staff were trustworthy' items. Removal of the 'Efficient service' item resulted in a marked improvement in model fit (Chi-Square = 3.51, df = 2, P-value = 0.17331, RMSEA = 0.046, NFI = 0.99, CFI = 1.00) leaving four items remaining.

A confirmatory factor analysis was also conducted for the Outcome dimension. The model was run with the five items loading onto the Outcome dimension and as with the

Interaction/Employees dimension a very poor fit was found (Chi-Square = 92.30, df = 5, P-value = 0.00000, RMSEA = 0.222, NFI = 0.87, CFI = 0.88). Inspection of the error variance between the items again suggests that perhaps the 'Good Service' item should be deleted. Upon deletion of this item the fit improved (Chi-Square=9.44, df=2, P-value=0.00890, RMSEA=0.103, NFI = 0.98, CFI = 0.98), however this can still not be regarded as acceptable fit. It was next decided to model this four-item Outcome construct with the four-item Interaction/Employee construct to determine whether any additional anomalies in the data surfaced. Modelling the two constructs in conjunction with one another resulted in an improved fit (Chi-Square = 67.42, df = 19, P-value = 0.00000, RMSEA = 0.085, NFI = 0.94, CFI = 0.96). Modification indices indicate that the 'The service was of a very high quality' item was unstable and was thus removed from the analysis. This is not unexpected as this item was added to the Grace and O'Cass's original measure to ensure there were enough items in the scale. The removal of this item resulted in the final service quality model which is shown in Figure 5.3 below.

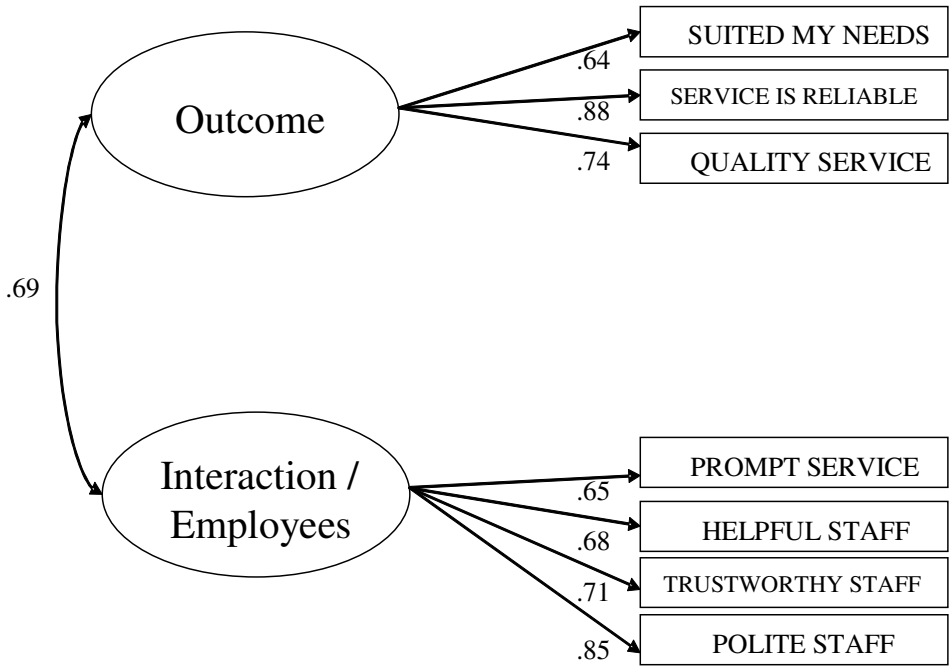


Figure 5.3: Measurement Model for Service Quality

The fit for this model is extremely good across all key fit indices (Chi-Square = 24.80, df = 13, P-value = 0.02453, RMSEA = 0.051, CFI = .99, NFI = .98, SRMR = 0.029) and supports the hypothesis that service quality is a two factor structure. Discriminant validity was also assessed between the two measures and it was found to be satisfactory. A more detailed account of the findings for this model can be viewed in Table 5.5 below:

Table 5.5: Results for the Two Dimensional Service Quality Model

	Loading	t-value	Coefficient α	Composite Reliability	Average Variance Extracted
Outcome Service Quality			.79	.80	.58
The service here suited my needs	.64	11.95 [b]			
The service here is reliable	.88	—[a]			
The service station provided quality service	.74	13.80 [b]			
Interaction/Employees			.79	.82	.53
The staff were polite	.85	12.31 [b]			
The staff were helpful	.68	10.63 [b]			
The staff were trustworthy	.71	10.99 [b]			
The service in store was delivered promptly	.65	—[a]			
Fit Statistics: χ^2 = 24.80, df = 13, P-value = 0.02453, RMSEA = 0.051, CFI = .99, NFI = .98, SRMR = 0.029					

[a] = The corresponding coefficient was fixed to set the metric of the latent construct.
[b] = Significant at the $p < .05$ level.

5.8 Assessing the Relationship between the Servicescape and Service Quality

The above sections have shown that the service quality construct consists of two distinct dimensions. As the dimensionality of this two factor structure has been proven, the next step is to demonstrate how the servicescape is an antecedent to service quality rather than a component or sub-dimension of service quality. This stage of the analysis will also assess the nomological validity of the servicescape. Nomological validity examines the adequacy of the concept by determining the relationships between the construct itself and measures of other constructs (Cronbach and Meehl, 1955). This can be achieved through structural equation modelling whereby the servicescape will be modelled as an antecedent of service quality to determine whether there is a relationship between the two. In the sections that follow, Hypothesis 1 is tested and the relationship between the constructs is discussed.

5.8.1 Modelling the Servicescape as an Antecedent to Service Quality

As with previous analyses, structural equation modelling was used to test the relationships between the servicescape and service quality. Figure 5.4 below depicts the proposed relationships.

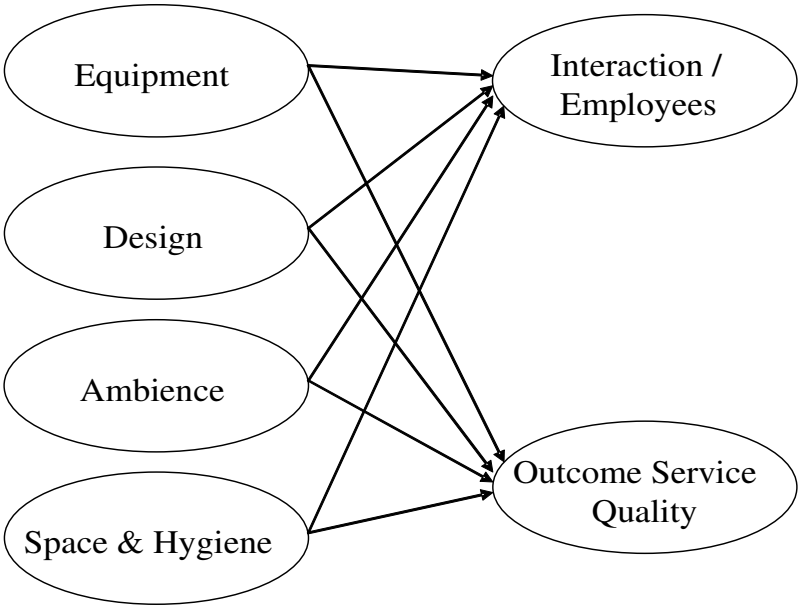


Figure 5.4: The Relationship between the Servicescape and Service Quality

5.8.2 Structural Equation Modelling

The syntax for relationships depicted above was written and the model was run. At the global level the fit for the model was good with model fit statistics pointing to a well fitting model (Chi-Square = 336.72, df = 175, P-value = 0.00000, RMSEA = 0.051, NFI = 0.89, CFI = .94, SRMR = 0.045). However, the standardised paths between many of the constructs were quite weak with some paths not reaching statistical significance. An anomaly was also found in the data with the path between the Hygiene construct and the

Interaction/Employees construct being greater than 1.00. Inspection of the data revealed that the modification indices were suggesting the addition of a path from the Interaction/Employee construct to the Outcome Service Quality construct. This final model produced almost identical fit however, the standardised path greater than one resolved itself. The fit statistics for this model can be viewed in Table 5.6 below.

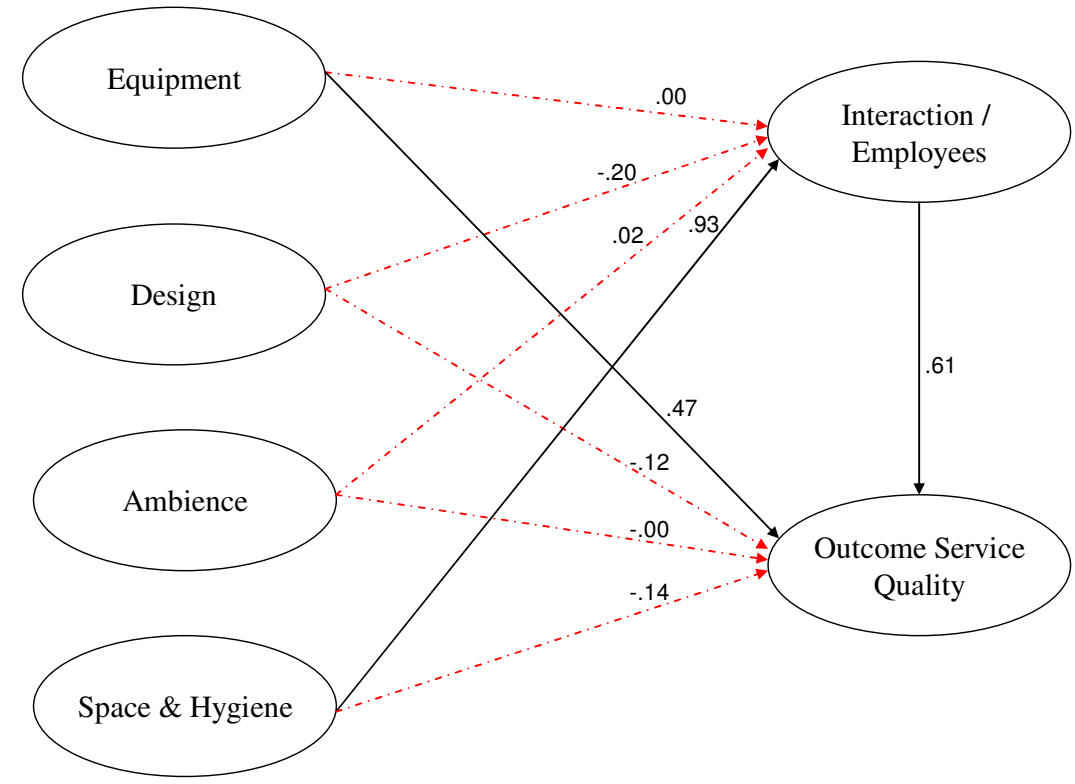
Table 5.6: Estimates of Path Coefficients and Results of Hypothesis Testing

	Loading	t-value	Coefficient α	Composite Reliability	Average Variance Extracted
Outcome Service Quality			.79	.80	.58
The service here suited my needs	.65	11.95[b]			
The service here is reliable	.86	—[a]			
The service station provided quality service	.75	13.80[b]			
Interaction/Employees			.79	.82	.53
The staff were polite	.83	12.31[b]			
The staff were helpful	.68	10.63[b]			
The staff were trustworthy	.73	10.99[b]			
The service in store was delivered promptly	.65	—[a]			
Ambience			.76	.79	.36
The background music was pleasant	.63	—[a]			
The background music was appropriate	.98	7.27[b]			
Design			.78	.79	.35
I found the interior design visually appealing	.81	15.15[b]			
The interior design was attractive	.81	—[a]			
The colour schemes were pleasant	.59	10.88[b]			
The architecture was attractive	.58	10.20[b]			
Equipment			.69	.69	.36
The equipment was modern looking	.67	8.64[b]			
The electronic equipment was excellent	.55	—[a]			
The equipment was of high quality	.75	9.14[b]			
Hygiene and Space			.75	.75	.38
The store was very clean	.61	—[a]			
The service station appeared to be hygienic	.69	10.20[b]			
The flooring was appropriate	.60	9.25[b]			
I found my way around quite easily	.66	9.85[b]			
The employees were neat and tidy in appearance	.55	8.56[b]			
Fit Statistics: χ^2 = 319.88, df = 174, P-value = 0.00000, RMSEA = 0.049, CFI = .95, IFI = .95, NFI = .89 SRMR = 0.044.					

[a] = The corresponding coefficient was fixed to set the metric of the latent construct.

[b] = Significant at the $p < .05$ level.

The path estimates for the measurement part of the model were all satisfactory. However, many of the structural paths between the constructs were insignificant. Figure 5.5 below depicts the relationships between these constructs. The dashed lines signify the hypothesised paths that no support was found for.



Solid line: Significant Path

Dashed red line: Insignificant Path

Figure 5.5: Linking the Servicescape to Service Quality

5.8.3 Analysis and Results

Maximum likelihood estimates for the various parameters in the model are presented in Table 5.7. Hypothesis H1 proposed that each of the servicescape dimensions would have a significant and positive effect on both of the service quality dimensions. Within this section the results from the structural equation modelling are presented, taking each servicescape dimension and how it relates to service quality in isolation.

Table 5.7: Standardized Parameter Estimates

Path	**Standardised Loading**	**T-Value**	**Hypothesis**
Equipment → Interaction/Employees	0.00	0.02	Not Supported
Equipment → Outcome Service Quality	0.47	3.32*	Supported
Design → Interaction/Employees	-0.20	-1.86	Not Supported
Design → Outcome Service Quality	-0.12	-1.02	Not Supported
Ambience → Interaction/Employees	0.02	0.33	Not Supported
Ambience → Outcome Service Quality	0.00	-1.86	Not Supported
Hygiene & Space → Interaction/Employees	0.93	6.62*	Supported
Hygiene & Space → Outcome Service Quality	-0.14	-0.74	Not Supported
Interaction/Employees → Outcome Service Quality	0.61	4.30*	Supported

* = Significant at the $p < .05$ level.

The relationship between Equipment and Interaction/Employees and Outcome Service Quality

The path between the Equipment construct and the Interaction construct was not supported with a low standardised coefficient and also a non-significant t-value. On the other hand, the path between Equipment and the Outcome Service Quality was found to be positive and significant.

The relationship between Design and Interaction/Employees and Outcome Service Quality

No support could be found for the hypothesised paths between the Design construct and the two Service Quality constructs with both low loading standardised paths and non-significant t-values.

The relationship between Ambience and Interaction/Employees and Outcome Service Quality

As with the Design construct, none of the paths between Ambience and the two endogenous variables in the model were found to have either high or significant loadings thus finding no support for the hypothesised relationships.

The relationship between Space and Hygiene and Interaction/Employees and Outcome Service Quality

Of the two hypothesised paths, only the relationship between Space and Hygiene and the Interaction/Employees dimension was found to be significant thus finding support for the hypothesised link.

5.8.4 Model Discussion

Overall it is evident that only limited support can be found for the above model. Although many of the paths are not significant, those that are significant demonstrate that aspects of the servicescape have a positive effect on the dependent variables. Interestingly, it was found that respondents' perceptions of the equipment had no effect

on their perceptions of the employees yet had a strong positive effect on Outcome Service Quality. This is a logical finding given that much of the equipment used by the customers is self service equipment which is in place as a substitute for employee service. A surprising finding is that design elements had no impact on either consumer's perceptions of the employee service or the outcome service quality. The reason for this is unknown but it may be related to the type of service situation under scrutiny here. It was also found that the ambience/music dimension had no impact on either outcome variables. This may be due to a number of factors; firstly, it is likely that the music being played in the service station was not noticed by the respondents as it was neither too loud nor inappropriate for the context. This means that the music being played in store has fallen outside the respondents' threshold of consciousness and is thus imperceptible to consumers. A second reason for this finding may purely be that music is quite literally irrelevant in its contribution to the overall service experience. Again this is likely to be linked to the service context as in more extended service situations such as restaurants the role of music is likely to be much greater.

The final servicescape construct to be examined is the role of Hygiene and Space. For this construct it was found that no significant relationship was found between it and Outcome Service Quality, however it was found to have a significant effect on perceptions of the Interactions/Employees (γ = .93). The reason behind this finding may be because customers believe that the upkeep, cleanliness and presentation of the store and its environs are the responsibility of the staff and it therefore has a direct effect on their perceptions of them. While the direct path between the Hygiene and Space

dimension and the Outcome Service Quality dimensions was found to be insignificant, the indirect effects of this dimension on the Outcome Service Quality dimension was positive and significant through the Interaction/Employees construct. It also must be noted that the path between Hygiene & Space and the Interaction/Employees construct is quite high (γ = .93). The reason for this is likely to be attributed to the presence of the 'The employees were neat and tidy in appearance' item on the Hygiene & Space construct. As this item relates to the neatness of the employees and is not a behavioural item it is sensible to retain this item on the Hygiene and Space construct.

5.9 A Formative versus Reflective Model

In recent years philosophical issues relating to the causal relationships between constructs and their measures has received growing interest (Edwards and Bagozzi, 2000; Edwards, 2001; Diamantopoulos and Siguaw, 2006). There are two primary ways in which the relationship between constructs and their measures can be described, these being commonly known as reflective or formative measurement models. Reflective measures represent manifestations of a construct and are in line with classical test theory and factor analysis (Nunnally, 1978; Bollen, 1989; Edwards and Bagozzi, 2000).

In order to determine whether the servicescape is best viewed as either a reflective versus a formative model, structural equation modelling was used. Firstly, the servicescape dimensions described in the previous chapter and earlier in this chapter were parcelled so that a single item represented each dimension. Both a formative and a reflective model was run and it was found that the formative model had a slightly superior fit to that of the

reflective model (Formative Model: Chi-Square=65.13, df=36, P-value=0.00208, RMSEA=0.048; Reflective Model: Chi-Square=148.26, df=42, P-value=0.00000, RMSEA=0.085). Despite this, there are two issues that need to be considered with regards to these models. Firstly, while the formative indicator model provides better fit, it has high modification indices which suggest that direct paths between the servicescape dimensions and both service quality dimensions. Secondly, modelling the servicescape as either a reflective or formative model poses diagnostic problems as it hides the amount of variance that each of the dimensions contribute to the two service quality dimensions. Taking this into consideration, it was decided that all further analyses would proceed to model the servicescape as multiple dimensions so that the individual contributions of the servicescape dimensions can be seen. LISREL output for both the formative and reflective models can be viewed in Appendix 5B.

5.10 Conclusions

This chapter introduced the Service Quality construct and reviewed the shortcomings of the literature in this area. The SERVQUAL measure was presented and from this more current conceptualisations of service quality were introduced. These more recent conceptualisations of SERVQUAL generally describe the construct as a three factor structure. The current research model proposes that a two dimensional structure in which the servicescape serves as an antecedent is more appropriate.

Using confirmatory factor analysis, this two dimensional structure of service quality was tested and support was found for its existence. Following this, a full structural model in

which the servicescape dimensions found in the previous chapter were modelled preceding service quality. The findings from the modelling were mixed with only some of the hypothesised paths supported. Rather than dismiss these constructs as irrelevant all constructs from this stage of the analysis will be retained for further inquiry. This is particularly important since in the chapters that follow new dependent variables such as emotion, customer satisfaction and behavioural intentions will be introduced.

6.0 EMOTIONAL AND COGNITIVE POST-CONSUMPTION PROCESSES

6.1 Introduction

This chapter brings together research on the servicescape and the literature from customer satisfaction research. The moderating variable linking these two constructs are consumption emotions and are thus the central theme of this final literature chapter. In order to fully understand the theoretical underpinnings of the role of emotions in consumer behaviour, the chapter will first provide an overview of key emotion theories which have been developed within the field of cognitive psychology. Following this, store atmospherics and servicescape studies which have incorporated the impact of emotions will be reviewed. The chapter then turns to the customer satisfaction literature and it is in these sections that it becomes apparent that emotions are also seen as a key component in the satisfaction response. Customer satisfaction studies integrating emotion are then reviewed. Within this section on customer satisfaction two related constructs are also discussed. The first of these is service quality. While service quality was discussed in depth in the previous chapter, in this chapter the relationship between customer satisfaction and service quality is addressed. The final section looks at behavioural intentions. Through emotional responses, environments are widely regarded as having the ability to elicit behavioural responses. Similarly, it is generally believed that high levels of customer satisfaction have a direct impact on consumers' intentions to repatronage a service or retail unit. This section deals with these two areas and from this hypotheses are drawn.

6.2 Defining Emotion

Many have maintained that emotions constitute the primary motivational system of human beings (Izard, 1977). This widely held view has led psychologists to debate and research this pertinent area in great detail. Although the study of emotions is recognised as being integral to understanding human processes, it has been described by psychologists as a tiresome (Evans, 1988), mysterious (Russell, 2003), confusing (Hebb, 1949) and unsatisfactory (Plutchik, 1980) subject. These complaints can be attributed to the subject's vague and subjective nature and these idiosyncrasies have in turn led to little agreement on the subjects' definition (Zeitlin and Westwood, 1986; Buck and Georgson, 1997; Izard, 2007). When defining emotion, it is often referred to as being 'subjective', 'behavioural' and 'possessing physiological aspects' (Plutchik, 1980, p. 79). However it has been pointed out that other phenomena such as hunger, thirst and pain possess these same characteristics and is thus not an adequate definition of emotion (Plutchik, 1980). Another indication of the lack of unanimity regarding emotion, is that many major psychologists who have written extensively on the subject side-step providing an actual definition of emotion, while some theories are merely attempts at defining an emotion. In a synthesis of the literature by Plutchik (1980), dozens of definitions of emotion are provided which date as far back as 1884. So as to give an indication of the contradictions and variations within the literature Plutchik (1980) provides a lengthy list of emotion definitions.

According to Plutchik (1980), many of these definitions refer to the bodily and physiological changes associated with emotions, some solely refer to the physiological

aspects, whereas most of the remaining theorists refer to the physiological aspect of emotions as a component rather than their sole characteristic (Plutchik, 1962; Izard, 1972; Lazarus, 1975). One characteristic which also appears in the definitions is that emotions are an adaptive response to a stimulus which helps us cope with situations (Carr, 1929; Lazarus, 1975). The relationship between emotion and behaviour (or action as a result of experiencing a particular emotion) is also referred to by a number of authors (Young, 1943, Izard, 1972; Lazarus). Surprisingly, a very small number refer to whether or not emotion is a psychological phenomenon (Young, 1943; Lazarus, 1975), although emotion is referred to as a neurophysiological response by some (Carr, 1929; Izard, 1972). Two definitions that are most conspicuously unlike the others are that of Watson (1924) and Skinner (1938). Watson for instance, takes a very primitive view of emotions stating that they are hereditary, whereas Skinner (1938) fundamentally disagrees with many of the definitions by stating that emotion is not a response to a stimulus but a motivational drive that is innate within individuals.

6.3 Psychological Theories of Emotion

The very different approach to defining emotion provided by Watson and Skinner is a reflection of the underlying behaviourist theory from which it is derived. In more recent years, the dominant theories of emotion have moved away from this view and now regard cognitive psychology as a more acceptable framework. This is the primary stance taken by researchers such as Lazarus (1982; 1984), Frijda et al (1989) and Schachter and Singer (1962). These authors believe that emotions are centred on the concept of cognitive appraisal whereby the way one interprets ones situation at any given moment is crucial to

the emotional response (Lazarus, 1982; Frijda et al, 1989). Most consumer behaviour studies which have examined the role of emotions on both cognitive and behavioural responses have drawn from this field as they have the 'advantage of specifying the antecedents of emotions and thus have predictive ability' (Nyer, 1997, p.297). In the sections that follow, the various different emotion theories that have influenced consumer theory will be described. Although many different emotion theories exist (Plutchik, 1980) the frameworks reviewed here have been included as they have formed the foundation of notable work by prominent consumer behaviourists such as Oliver, Westbrook and Bitner (Westbrook, 1987; Westbrook and Oliver, 1991; Bitner, 1992; Mano and Oliver, 1993; Oliver, 1993; Oliver et al, 1997; Rust and Oliver, 2000). The cognitive appraisal theories of emotion will be discussed foremost as it is these that have formed much of the backbone of consumer research in this area. Following this, evolutionary theories of emotion (Plutchik, 1980; Izard, 1977) will be discussed. While the theoretical underpinnings of these evolutionary theories are not entirely congruent with cognitive theories of emotion, their contribution to theory are pertinent given the use of their measures in various consumer behaviour studies. Following this discussion, structural theories of emotion will be discussed (Mehrabian and Russell, 1974; Russell, 1980; Watson and Tellegen, 1985). These theories are grounded in cognitive appraisal models however the central focus of these authors work is not the causal paths between emotion, cognition and behaviour but on the categorisation of emotion. Unlike Izard and Plutchik these authors do not believe there exist fundamental emotions, instead they seek to categorise emotions along dimensions which are plotted in a three or two dimensional space. In this sense, some emotions are thought to be conceptually similar to one another

while emotions that are placed far away from other emotions are thought to be distinctly different.

6.3.1 Cognitive Appraisal Theories: Lazarus and Frijda

Lazarus (1982) generated a renewed interest in the study of emotion when he proposed that emotions are the outcomes of the cognitive appraisal of an event. He believes that 'cognitive activity is a necessary precondition of emotion because to experience an emotion, people must comprehend' (1984, p. 124). This comprehension is what Lazarus views as appraisal and is core to his theory (1982; 1984). In sum, the appraisal process causes an emotional response based on a cognitive review of the situation. The stimulus or events that cause an appraisal are not what determines the emotional response, instead it is subjective in that it is an individuals' internal reaction to that stimulus that causes the appraisal. The work of Frijda (1989) is also grounded in the cognitive theories of appraisal. His view is that all emotions involve appraisal of two types; primary and secondary appraisal. Primary appraisal refers to the emotional meaning of an event whereas secondary appraisal is concerned with the evaluation of the resultant emotion. According to Frijda, (1989) we are continually monitoring our environments where a constant feedback mechanism is at play (Strongman, 2003). It is important to note, that he believes that not all stimuli elicit emotion, only stimuli that have relevant concerns to us will cause an emotional reaction.

In consumer behaviour it is this cognitive appraisal approach that forms how researchers understand motivational and evaluative roots (Watson and Spence, 2007). Indeed, some

advocate that these appraisal theories offer a more complete explanation of behavioural responses (Bagozzi et al, 1999; Johnson and Stewart, 2005). Apart from Nyer (1997), very few have attempted to integrate this theory into empirical research. Despite this, the basic premise of the theory is prevalent in consumer behaviour theory. For example, servicescape researchers advocate that consumers' appraisal of environments leads to emotional responses. Similarly, the customer satisfaction literature proposes that consumers' evaluations of products and services leads to emotional responses which subsequently impact on overall satisfaction.

6.3.2 Evolutionary Emotion Theories: Izard and Plutchik

While cognition is an important part of evolutionary theories the difference between this and the cognitive appraisal approach is that evolutionary theorists believe that non-cognitive activators can cause emotion, thus viewing emotion as entirely independent of cognition (Strongman, 2003). A second difference is that they view emotions as biologically based and universally experienced which manifest to aid us cope with adaptation (Richins, 1997). A final distinction is that evolutionary theorists propose that there exists a small set of basic emotions that motivate us biologically and mentally and are fundamental across cultures and species.

Izard's (1977) Differential Emotions Theory (DET)

Throughout the history of psychology it was thought that the physiological drives of hunger, thirst, sleep, sex and pain were the primary motivators (Izard, 1991). Izard (1977) however argues that emotions constitute the primary motivational system for

human beings and that these emotions are an evolutionary phenomenon that are hardwired into our biological roots (Izard et al, 2000). He thus defines emotions in biological terms, where emotions are viewed as a complex electrochemical activity which occurs within the brain and the nervous system. At the muscular level, emotion manifests in the form of facial activity and bodily (for example gestural, postural) responses. The feedback from these visceral activities is then transformed into a discrete fundamental emotion such as fear, joy or anger. Izard (1977; 1991) believes that, at the consciousness level, emotions act as motivators which hold significance for the individual. He also believes that these emotional processes can interact and exert influence on perceptual cognitive and motor processes. This notion of emotion prior to cognition, or emotion occurring independently of cognition, is an important point of difference between Izard and the stimulus-organism-response views of cognitive psychologists such as Lazarus (1984).

Izard (1977) regards emotions at the same level of generality, that is, ten *basic* emotions exist that are discrete from one another. Each of these emotions are defined in terms of unique subjective feeling, pattern of facial expression and behavioural consequence (Reeve, 1997). The ten basic emotions are; interest-excitement, happiness-joy, surprise-astonishment, sadness-grief, anger-rage, disgust-revulsion, fear-terror, contempt-scorn, shame-shyness and guilt-remorse. Of these ten emotions only two are positively valenced (interest and happiness), seven are negative (sadness, anger, disgust, fear, contempt, shame and guilt), and surprise-astonishment is classified as neutral. Izard (1977, 1991) does allow for the fact that more than one emotion can occur at the one

time, however, they remain conceptually discrete. Izard (1977) argues that such widespread states of love, hate, anxiety or depression are not fundamental as they do not produce coinciding, distinct facial expressions. Moreover, Izard (1977) believes that hate is an emotion pattern that derives from the combination of anger, disgust and contempt. Other such emotion combinations would be love, which is thought to be a combination of interest that is repetitively followed by joy. One major criticism of Izard's (1977) approach is that it has not provided an adequate rationale for choosing his fundamental emotions, other than that these emotions manifest in specific facial expressions (Plutchik, 1980). Despite this shortcoming, Izard (1977; 1991) has made some important contributions in relation to matching facial expressions with verbal labels and physiological indices of muscular patterns of the face. From a marketing research perspective, his most important addition to theory has been the development of a thirty item Differential Emotions Scale (DES II). This is a simple self-report adjective checklist and has been used in a variety of consumption research settings (Liljander and Strandvik, 1997; Oliver et al, 1997).

Plutchik's (1980) Psychoevolutionary Perspective on Emotions

Coinciding with Izard's (1977) DET, Plutchik (1980) also views emotions as adaptive processes which have evolved as a form of coping mechanism. Also in a similar vein to Izard (1977), Plutchik (1980) asserted that there are a small number of basic primary emotions. These eight basic emotions were; joy, acceptance, fear, surprise, sadness, disgust, anger and anticipation. Izard (1977) conceptualises primary emotions as interacting in combinations, or patterns, where the basic emotions retain their qualitative

identity (i.e. love is a result of a combination of interest repeatedly followed by joy, where the primary emotions remain independent). Plutchik (1980), on the other hand, advanced the notion that new emotions can be created by blending affects that occur simultaneously. For example, when anxiety is mixed with fear it will produce guilt, or joy and acceptance will produce love. Although Plutchik (1980) focuses heavily on the biological and evolutionary roots of emotions, he does allow for an interplay between cognition and emotion in his framework. To synopsise, Plutchik (1980. p. 833) refers to emotions as an 'inferred sequence of events set in motion by some stimulus'. To illustrate; a cognition (danger) precedes subjective feeling (fear) which leads to physiological arousal (rapid heart beat), followed by a motor impulse (to run) and finally manifesting in behaviour (running). This sequence of events transfers logically to consumer behaviour. For example, an evaluation of a poor service experience (inferred cognition) will lead to anger (emotional response) and an increased heart-rate (physiological response) which in turn makes the individual want to leave the service, (motor impulse) so they then leave (overt behaviour).

Based on his psychoevolutionary theory of emotion, Plutchik (1980) provides the Emotions Profile Index (EPI), which is a measure of emotions and their combinations. The EPI is a forced-choice test consisting of twelve adjectives which are paired with all possible combinations yielding 66 pairs. However, four of these pairs were dropped as they were found to be overlapping, thus leaving 62 pairs in the final measure. Despite the empirical origins of this scale it has not been used as extensively as other emotion scales (Izard, 1977; Watson et al, 1988), perhaps this is because Plutchik's measure is at least 40

items longer than other popular scales. However it has been used by Holbrook and Westwood (1989) who modified it for use in consumer research by creating a shorter version where respondents rate the level of felt intensity for each adjective (3 adjectives per emotion).

6.3.3 Structural Theories of Emotion

Mehrabian and Russell's Pleasure-Arousal-Dominance (PAD) Theory of Emotion

Unlike Izard (1977) and Plutchik (1980), Mehrabian and Russell (1974) do not define emotions in terms of those that are basic or primary, instead they suggest that emotions can be described in terms of three orthogonal dimensions: Pleasure, Arousal and Dominance. The first dimension, Pleasure is defined as a feeling state and is conceptualised as theoretically distinct from preference, liking, positive reinforcement or approach-avoidance. Arousal is an excitement dimension, which is also defined as a feeling state, however this varies from sleep to frantic excitement. The third emotional dimension, dominance is again a feeling state, but refers to the extent to which an individual feels restricted versus free in a variety of ways. For example, an individual has greater freedom watching television in the comfort of their own home (high feeling of dominance), compared to watching a film in the cinema. When factor analysis was conducted, this dimension was found to be the most tenuous as it accounted for the smallest amount of variance in their studies (12% in Study 1 and 14% in Study 2). In addition, later studies which have attempted to replicate the three-dimensional structure of emotion (Russell, 1980; Russell and Pratt, 1980; Donovan and Rossiter, 1982), have only found support for a two-dimensional structure (pleasure and arousal).

Russell's Pleasure-Arousal Theory

This led Russell to develop a two-dimensional circumplex model (1980) which is also known as the Pleasure-Arousal Theory (PAT) of emotions. His rationale for excluding the dominance dimension was that it was found to be unstable across samples and situations (Plutchik, 1980). Russell's PAT proposes that emotional experience is adequately described by two continuous, bipolar, and orthogonal dimensions, *pleasure-displeasure* (P) and *arousal-sleep* or *activation-deactivation* (A).

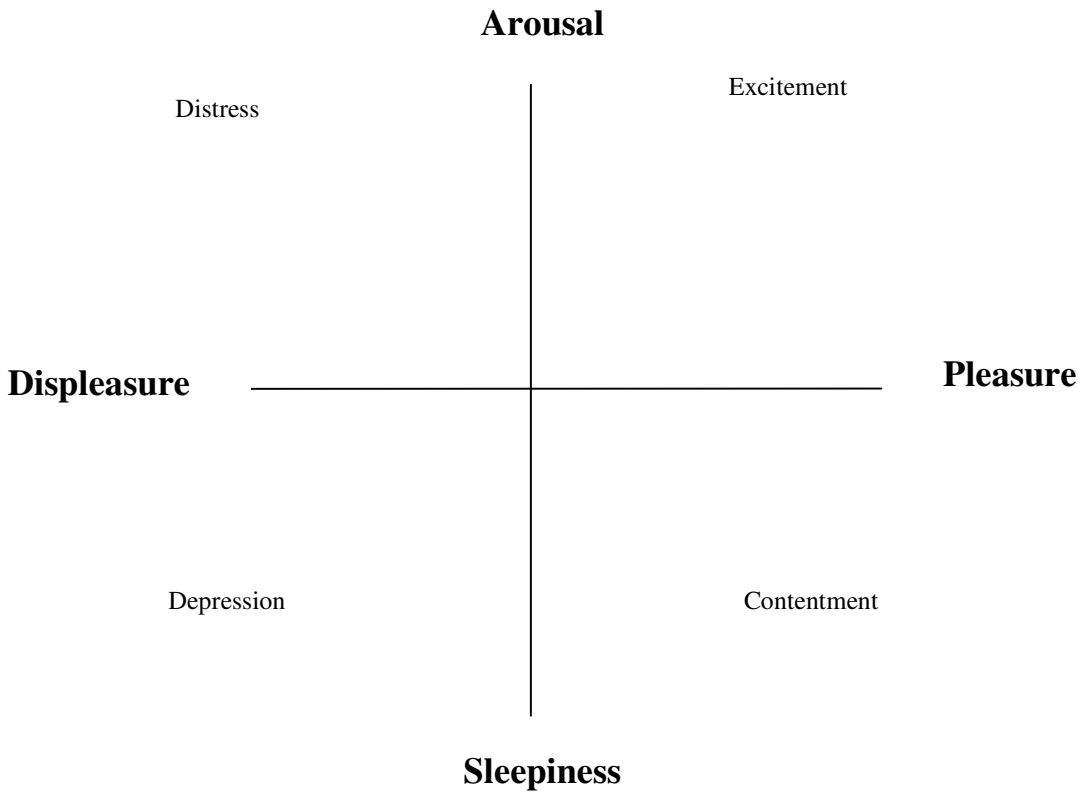


Figure 6.1: Russell's Pleasure-Arousal Circumplex (1980, p. 1164)

While primary emotion theorists assume that emotions are unrelated and discrete (Izard, 1977; Plutchik, 1980), this circumplex structure presumes at the most fundamental level that some emotions are similar to one another while others are distinctly different (Larsen and Diener, 1992). Circumplex structures also assert the order in which emotions fall around the circumference of the circumplex space.

Watson and Tellegen's Positive/Negative Affect Circumplex

Watson and Tellegen (1985) provided an interpretation of the structure of emotions that is based upon Russell's (1980) Pleasure-Arousal Theory. This conceptualisation also provides a two-dimensional structure and is the most widely advocated emotional framework (Larsen and Diener, 1992). This model rotates the axes by 45 degrees and labels the two primary dimensions positive and negative affectivity (Figure 6.2).

Although these dimensions initially appear to be opposites, the authors propose that they are in fact independent of one another and uncorrelated (Watson and Tellegen, 1985). The first factor, Positive Affect relates to the extent to which a person states a zest for life and ranges from High to Low. On the other hand, Negative Affect relates to how upset or unpleasantly aroused the individual is and ranges from High Negative Affect to Low Negative Affect. The high ends of both dimensions represent a state of emotional arousal, while the low ends reflect an absence of activation with such emotions as *calm, relaxed, quiet* and *placid* falling into these areas.

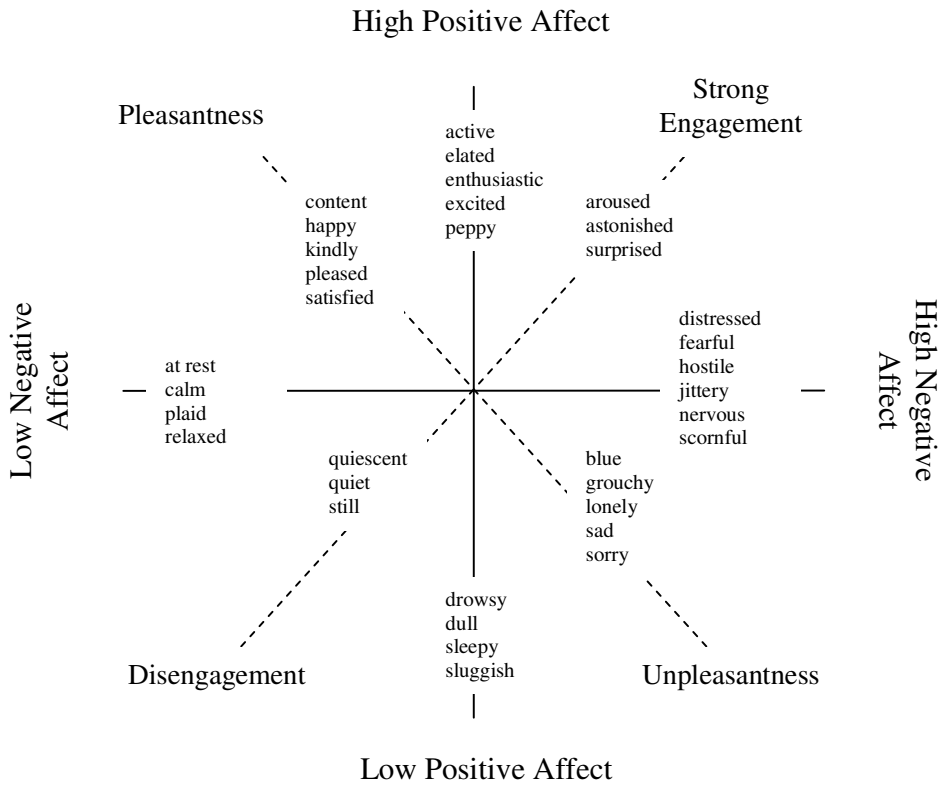


Figure 6.2: Watson and Tellegen's Two-Factor Structure of Affect (1985, p. 221)

Watson and Tellegen (1985) have further segmented the dimensional space with an additional two dimensions which are named engagement and pleasantness. In the North-East of the circumplex lies Strong Engagement, or otherwise known as Positive Affect with High Activation (Carroll et al, 1999). This refers to moderately positively valenced items that are moderately high in arousal such as *astonished* or *surprised*. At the other end of this dimension lies Disengagement, or Negative Affect with Low Activation (Carroll et al, 1999). Emotions falling into this segment are *quiet* or *still*. The Pleasantness dimension ranges from the North-West octant to the South-East octant. The North-West octant, or pleasantness, represents terms that are a mixture of high Positive

Affect and Low Negative Affect. The final octant lies between Low Positive Affect and High Negative Affect and is referred to as Unpleasantness, emotions associated with this segment are *lonely*, *unhappy* and *sad*.

Although this model is the most widely cited and implemented model of affect, there are those who have criticised it on a number of bases. The first problem relates to the naming of the dimensions. Larsen and Diener (1992) find the interpretation of the dimensions Positive Affect (PA) and Negative Affect (NA) to be misleading. They argue that this interpretation does not correspond with what other researchers have designated as Positive Affect and Negative Affect. For example, in Watson and Tellegen's (1985) circumplex, the Positive Affect dimension is anchored on one end with *elated*, *excited* and *peppy*, while at the other pole adjectives with an absence of affective involvement lie; examples of which include *dull*, *drowsy* and *sluggish*. They contend that referring to this dimension as Positive Affect is a misnomer as it includes both pleasant and unpleasant elements (Larsen and Diener, 1992). They attempt to resolve this issue by providing their own interpretation of the affective labels which are in keeping with Russell's categorisations (see Figure 6.1). In the place of positive affect, they name this dimension Activation (ranging from high to low). They also renamed the negative affect dimension Pleasantness (ranging from unpleasantness to pleasantness). This construction of the affect dimensions is more reflective of the actual content of the affective states, while remaining straightforward and fully dimensional (Larsen and Diener, 1992). A second contention these authors have raised relates to Watson et al's (1988) measure of positive and negative affect (PANAS). This scale is a twenty-item mood scale, however

it merely contains adjectives from the high activation end points of the dimensions. Larsen and Diener (1992) reason that in order for the scale to be a true bipolar measure of affect it would include adjectives from the other end of the dimension. Despite their criticism of this scale, it remains one of the most popularly used measures of emotion to date.

Although structural models of emotion have been criticised in the past, their advantages surpass their limitations as they provide a clear structure for the effects of emotion on behaviour. In addition, they have been found to account for the majority of variance in affect measures (Russell and Pratt, 1980; Watson and Tellegen, 1985) which adds validity to their use.

6.4 Emotion and Environments

Emotional or affective reactions are thought to be central to our understanding of how environments shape behaviour. This is a central tenet of Mehrabian and Russell's (1974) theory of environmental psychology which describes in detail how various environments have direct effects on emotional responses. As mentioned earlier in this chapter, Mehrabian and Russell (1974) specify that emotional reactions to environments can be conceived in terms of three orthogonal dimensions: Pleasure, Arousal and Dominance. It is from this theory that Bitner's (1992) taxonomy and much of the atmospherics literature is based.

Apart from a limited number of studies (Wakefield and Blodgett, 1994) most studies which have attempted to apply Mehrabian and Russell's (1974) theory to consumption situations have tended to overlook the mediating role of emotions and have examined only cognitive outcomes such as customer satisfaction (Wakefield and Blodgett, 1994; Wakefield and Blodgett, 1996; Lucas, 2003; Johnson et al, 2004) or service quality evaluations (Reimer and Kuehn, 2005). However, within the atmospherics literature an abundance of research which has investigated the relationship between key environmental stimuli (such as music, scents, and lighting) and emotional responses (pleasure, arousal and dominance). This next section will discuss studies which have focused on emotional responses to the overall environment. The review of these studies will begin with those that applied Mehrabian and Russell's (1974) PAD method of measuring emotions. These PAD studies will be discussed chronologically so as to track the progress of these studies over the last twenty-five years. Following these studies, those who have used derivatives of the PAD will be discussed. These will again be discussed in chronological order.

6.4.1 Servicescape and Environmental Psychology Studies Employing the PAD Measure of Emotions

Donovan and Rossiter (1982) were the first to suggest that the physical environment can influence emotional states. This study applied the full Mehrabian and Russell (M-R) (1974) model, using their information rate, PAD and approach-avoidance measure. Mehrabian and Russell's (1974) information rate was used as a measure of the atmospherics or servicescape of the store. This measure is a general, non-specific

measure of the environment. Although the information rate meets the criteria for discriminant and convergent reliability, it does not allow us to determine which in-store stimuli elicit particular types of emotional responses. The information provided by Donovan and Rossiter (1982) on the relationship between the environment and the emotional responses is quite limited. For some reason, only the arousal dimension was regressed on the information rate and in disagreement with Mehrabian and Russell's theory it was found that the information rate measure was not a good predictor of arousal. In an extension of their earlier work Donovan et al (1994) again attempted to determine whether the PAD emotional responses mediate evaluations of store atmosphere and behavioural responses in a field environment. Unlike their earlier study (Donovan and Rossiter, 1982), information rate was not used to measure the atmospherics of the store here, however the environments were broken down into those that were pleasant or unpleasant. It appears that this decision was made using the pleasure scale, where those that reported feelings of unpleasantness were thus classed as finding the environment unpleasant and vice versa for those that reported feelings of pleasure.

A study which devised a direct measure of the environment and individuals emotional responses to it is that of Sherman et al (1997). The environment was measured using Dickson and Albaum's (1977) semantic-differential scale. Items were also included which were based upon Kotler's (1973-1974) article on atmospherics. These environmental stimuli were categorised into four groups; social influencers, image, design and ambience. Although these scales are an advancement on Mehrabian and Russell's (1974) information rate, they do not measure many of the environmental

variables set out by previous authors such as Bitner (1992; 2000) and Baker (1987). Despite this, the study made a methodological contribution through their use of structural equation modelling and a conceptual one through their integration of all three stimulus-organism-response dimensions. On the whole, the findings from this study were mixed with many of the paths between the environment and the emotion dimensions found to be insignificant or else in directions not expected. Insignificant paths were those between social, image and ambience to arousal and image to pleasure. It was also found that design had a highly significant, negative effect on arousal.

6.4.2 Studies Employing Derivatives of the PAD Measure of Emotions

Other researchers have attempted to explore the influence of servicescapes on consumer responses using different emotion measures, of note is the work of Wakefield, Blodgett and Baker (Wakefield and Blodgett, 1994; Wakefield and Baker, 1998). The first of these studies examined the impact of the servicescape on excitement and subsequent quality perceptions (Wakefield and Blodgett, 1994). According to Wakefield and Blodgett (1994, p. 69), excitement is, 'an emotional response characterized by feelings ranging from exhilarating, sensational, stimulating, and interesting to dull, dreary, boring, unstimulating, and monotonous'. The study in quite unclear about the actual emotion measures used however, it seems that this excitement measure is likely to have been taken from a larger scale by Russell and Pratt (1980). Most of the relationships between excitement and other constructs were found to be significant and in the direction expected. In particular, it was found that individuals who perceived the servicescape to be of high quality also experienced greater levels of excitement thus demonstrating how

environments can cause emotional responses. As this study measured emotions on quite a general level, it cannot be discerned from their results which particular emotions impact upon satisfaction. In addition, their evaluative measure of the servicescape did not explicitly identify the key elements that determine servicescape quality but was a one-item measure asking respondents to rate the overall quality of the servicescape. In a follow up study, Wakefield and Baker (1998) applied a similar model to a shopping mall context. They again examined the role of excitement, but extended their earlier work by investigating the relative influence of various aspects of the servicescape on levels of excitement. The specific environmental characteristics considered were design, music, layout, light/temperature and décor. They hypothesised that given the hedonic nature of shopping experiences, positive perceptions of the mall environment would have a positive effect on excitement and subsequent mall repatronage intentions. The environmental characteristics were regressed on the dependent variable of excitement. As with many of the studies utilising the PAD measure, the results were mixed. Ambience factors (lighting and temperature) had a negative effect on excitement, while interior décor had no effect. Although music and layout had positive effects on excitement, the combined R^2 value merely accounted for 26.9% of the variance in excitement which suggests that perhaps excitement is not an adequate measure of emotion.

Although both environmental psychologists (Mehrabian and Russell, 1974) and service marketing theorists (Bitner, 1992; Bitner, 2000) both strongly advocate that environments have the ability to elicit emotional responses. The above, holistic servicescape studies

have found it difficult to agree on the nature of the relationship. On the other hand, the atmospherics literature is replete with studies that have shown how individual environmental stimuli influence emotions (Turley and Milliman, 2000). This poses a problem as it unclear why this has occurred. This research seeks to extend current research by examining the relationship between the servicescape and consumption emotions. The contribution that is to be made here is twofold; firstly, no known study to date has sought to examine the relationships between a coherent multidimensional conceptualisation of the servicescape and emotions. Secondly, the chosen context also makes a contribution. To date, much of the holistic servicescape studies have focussed on hedonic service situations where emotions are more likely to be an important element of the consumption experience. Utilitarian service situations are relatively under-researched, with all studies in this area falling into the atmospherics group of studies. Atmospherics studies have shown that in utilitarian retail services that emotions are an important part of the evaluative process and therefore the following hypotheses were developed:

H1a: Consumers' perceptions of Servicescape Design will have a positive effect on positive emotions.

H1b: Consumers' perceptions of Servicescape Design will have a negative effect on negative emotions.

H1c: Consumers' perceptions of Servicescape Ambience will have a positive effect on positive emotions.

H1d: Consumers' perceptions of Servicescape Ambience will have a negative effect on negative emotions.

H1e: Consumers' perceptions of Servicescape Equipment will have a positive effect on positive emotions.

H1f: Consumers' perceptions of Servicescape Equipment will have a negative effect on negative emotions.

H1g: Consumers' perceptions of Servicescape Hygiene & Space will have a positive effect on positive emotions.

H1h: Consumers' perceptions of Servicescape Hygiene & Space will have a negative effect on negative emotions.

6.5 Customer Satisfaction

Understanding the factors that contribute to customer satisfaction is imperative for firms as having high levels of customer satisfaction is thought to represent an important source of income (Oliver, 1997). Empirical research has demonstrated a positive association between customer satisfaction and shareholder value (Anderson et al, 2004) and also

market share and profitability (Anderson et al, 1994). As a result of this, customer satisfaction is seen as fundamental to the marketing concept (Oliver, 1993; Spreng et al, 1996). The inclusion of customer satisfaction in this dissertation is significant for three reasons. Firstly, given the general acceptance in the literature that satisfactory purchase experience leads to continued interest in a product or service it is thus a construct that is of interest to the management of the service station under study. Secondly, by including customer satisfaction it allows for its relationship to other constructs studied here to be examined. For instance, one of the purposes of this dissertation is to determine whether there are aspects of the servicescape that are more salient to customers and this can only be done by including an overall evaluative measure in the model. Finally, how emotions contribute to the satisfaction response are also to be assessed. Several researchers have demonstrated how emotions constitute a large part of the satisfaction response, however services researchers have tended to focus on more extended service transactions (Liljander and Strandvik, 1997; Wirtz and Bateson, 1999) or hedonic service situations (Bigné et al, 2008). In this dissertation, an attempt is made to extend the current literature by examining the role of emotions in a utilitarian service situation. The structure of this section of the chapter is as follows: Firstly, customer satisfaction will be defined. Following this, the dominant theory of customer satisfaction will be introduced. Subsequent to this, the role of emotions in the satisfaction response will be discussed along with those studies which have focussed on service situations. From this hypotheses are drawn. The conceptual differences between customer satisfaction and service quality are also highlighted and from this a further set of hypotheses are developed.

6.5.6 Defining Customer Satisfaction

Despite the widely held view that customer satisfaction (CS) is of strategic importance, few agree on the constructs' definition (Oliver, 1997; Giese and Cote, 2000; Vanhamme, 2000). There are those that categorise satisfaction as an outcome and those who categorise it as a process (Yi, 1990) with outcome definitions seeing customer satisfaction as a product of the consumption experience, 'an emotional response to the experiences provided by, associated with particular products or services purchased, retail outlets, or even molar patterns of behaviour such as shopping and buyer behaviour, as well as the overall marketplace', (Westbrook and Reilly, 1983, p. 256). On the other hand, process definitions span the entire consumption experience; 'an evaluation rendered that the consumption experience was at least as good as it was supposed to be' (Hunt, 1977, p. 459). Some have defined satisfaction as a purely cognitive state (Churchill and Suprenant, 1982), while others have considered customer satisfaction to be an emotion (Woodruff et al, 1983). Nowadays, it is generally accepted that satisfaction contains both cognitive and affective elements (Liljander and Strandvik, 1997; Oliver, 1997; Vanhamme, 2000) and is seen as the result of a comparison between a subjective experience and an initial reference base.

6.5.7 Customer Satisfaction Models

In this next section the various customer satisfaction models will be presented. The principal customer satisfaction theory is the disconfirmation of expectations model and will be reviewed first. This is followed by a review of extensions to this model such as equity theory, attributional theory, value-percept and experience based norms.

The Dominant Paradigm

Several theories and models have been put forward regarding customer satisfaction, the most commonly referred to and most widely accepted is the disconfirmation of expectations model (DE model) (Day, 1977; Woodruff et al, 1983; Spreng et al, 1996; Oliver, 1997; Vanhamme, 2000; Phillips and Baumgartner, 2002). This views customer satisfaction as an attitude (Yi, 1990; Fornier and Mick, 1999). According to this purely cognitive paradigm, consumers form expectations prior to purchase and these act as a standard on which performance is judged. Post-usage, the consumer compares their expectations with their perceptions which results in either confirmation or disconfirmation. Confirmation occurs when the product or service perceptions exactly meet expectations, whereas disconfirmation will be the result of a discrepancy between expectations and perceptions. Two types of disconfirmation can be identified: positive disconfirmation occurs when product performance exceeds prior expectations. Negative disconfirmation occurs when expectations exceed performance. Confirmation and positive disconfirmation are likely to result in satisfaction, while negative disconfirmation leads to dissatisfaction. There is a voluminous amount of empirical evidence that supports this theory (Churchill and Suprenant, 1982; Bearden and Teel, 1983; Oliver and Bearden, 1983; Oliver and DeSarbo, 1988; Tse and Wilton, 1988; Oliver and Swan, 1989; Yi, 1990; Bolton and Drew, 1991; Westbrook and Oliver, 1991; Phillips and Baumgartner, 2002). The concept of the DE model is presented in Figure 6.3. below.

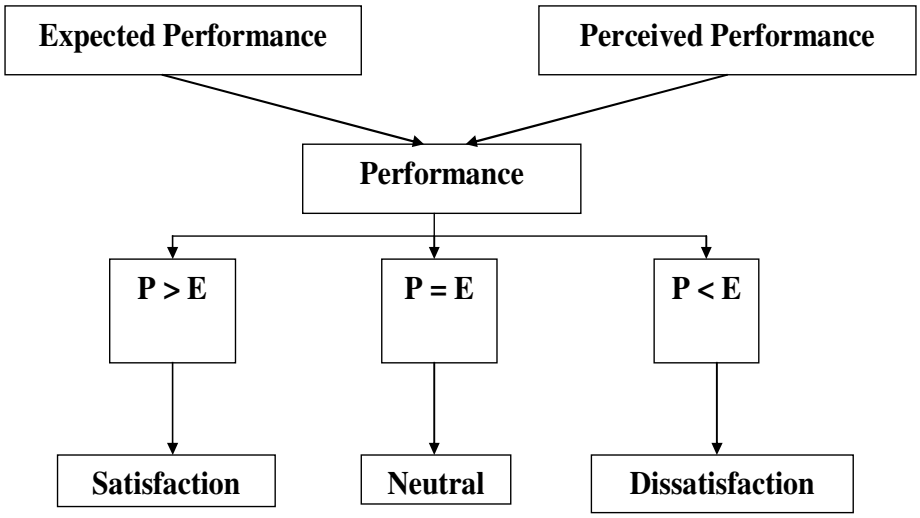


Figure 6.3: The Disconfirmation Model of Consumer Satisfaction

Further Criticisms and Extensions to the Disconfirmation of Expectations Model

Numerous alternative perspectives on the Disconfirmation of Expectations (DE) Model have been proffered over the years. These have included norms (Woodruff et al, 1983; Cadotte et al, 1987), attribution (Folkes, 1984; Folkes and Kostos, 1986; Folkes, 1988) and equity approaches (Oliver and Swan, 1989; Oliver, 1997; Bolton and Lemon, 1999) In this section these alternatives to the DE Model will be briefly discussed.

The experience-based norms model was put forward by Woodruff et al (1983) in response to scepticism in relation to the disconfirmation of expectations model. Woodruff et al (1983) debated whether expectations were an adequate measure on which to evaluate performance, they stated that expectations cannot be disregarded altogether as a standard for measuring customer satisfaction, but that it is probably not the sole

standard utilised by consumers. The model offers an expanded view of how prior experience influences confirmation/disconfirmation. It was proposed that expectations could be framed on the basis of historical norms or experiences, these could be derived from personal experience, positive/negative word-of-mouth behaviour or from marketing efforts of companies (Woodruff et al, 1983). Empirical work by the authors finds support for both the basic DE Model and also acknowledge that the paradigm is useful for comparing norms and standards (Cadotte et al, 1987).

Attribution theories also surfaced as a means of explaining the satisfaction response (Folkes, 1984; Curren and Folkes, 1987; Folkes, 1988). Prior to their application to consumer behaviour contexts, attribution theories were first developed as a means of explaining individuals reactions to other individuals which sought to distinguish between whether people attributed the causes of certain outcomes to be internal or external (Kelley, 1967; Kelley and Michela, 1980; Weiner, 1985; Weiner, 1986). According to the theory there are three dimensions of attribution: the internal/external dimension, the stability dimension and the controllability dimension. In consumer behaviour terms, attribution theory can easily be used to describe purchase outcomes and the causal relationships that people infer from outcomes (Oliver, 1997). For instance, the first dimension refers to whether a consumer attributes an outcome as being either customer related (themselves) or marketer related. Folkes (1984; 1986) has conducted extensive research in this area and has generally found support for the notion that when product failures occur consumers attribute the cause of the failure to external sources (i.e. the marketing entity). The stability dimension relates to whether the causes can be

considered stable and predictable, such as an individual's aptitude or whether they were highly variable such as a teacher's ability to teach a subject matter (Oliver, 1997). The final dimension is the controllability dimension which was added to account for variables that can be modified by the individual. An example of a controllable variable would be effort as it is something that someone can control.

A final addition to the DE model is inclusion of equity as a variable which influences overall satisfaction with a product or service (Oliver and DeSarbo, 1988; Oliver and Swan, 1989; Oliver, 1997; Bolton and Lemon, 1999). These models posit that if the consumer perceives that there is a negative balance, inequity (dissatisfaction) will occur. Inequity is thought to exist in two instances: Firstly, equity/inequity occurs where the consumer compares what they have spent (time, money) with what they receive from the seller (product/service). It has also been found that consumers consider how other consumers are treated in comparison to themselves; this has been coined interpersonal inequity.

While the above models have all made interesting and notable additions to current discourse on customer satisfaction, their impact has been overshadowed by the extensive amount of research into the role of emotion in the customer satisfaction response. Given the extensiveness of the research in this area, the next section of this chapter will deal with the literature in this area in isolation.

6.5.8 Emotion as an Antecedent to Customer Satisfaction

A major refinement to the disconfirmation of expectations model has been the inclusion of affective reactions (Westbrook and Oliver, 1991; Mano and Oliver, 1993; Oliver, 1993; Oliver, 1997; Oliver et al, 1997). Affective variables, and more specifically emotions have been shown to considerably add to the explanation power of the satisfaction model (Westbrook and Oliver, 1991; Oliver, 1993). As previously discussed, early studies on consumer satisfaction were based on the cognitive processes of confirmation of expectations, inequity judgments, and causal attributions (Oliver, 1980; Folkes, 1984; Oliver and DeSarbo, 1988; Oliver and Swan, 1989; Mano and Oliver, 1993; Mudie et al, 2003). More recently, studies have highlighted the importance of the processes that underlie consumption emotions and satisfaction (Westbrook, 1987; Westbrook and Oliver, 1991; Mano and Oliver, 1993). According to these authors, consumption emotions refer to the set of emotional responses elicited specifically during the consumption experience (Westbrook and Oliver, 1991; Wong, 2004). Emotions have also been viewed as resulting primarily from the cognitive appraisal processes a person employs when making sense of, evaluating, determining how to react and then responding to circumstances and events (Lazarus, 1982; Lazarus, 1984).

Before discussing the empirical studies which have examined affect and emotions in the satisfaction process, it is important to note that the terms affect, emotions, mood and attitudes have been frequently used interchangeably in the marketing literature (Bagozzi et al, 1999). Bagozzi et al (1999) distinguish between these constructs: Affect is considered an all-inclusive word which describes mental processes such as mood,

emotions and attitudes. It refers to the feeling side of consciousness as opposed to information processing which is more cognitive in nature (Mudie et al, 2003). Oliver (1997) gives a more detailed definition of affect where it is defined as the feeling side of consciousness. Affect (feeling) includes, 'pleasure and displeasure, happiness and sadness, liking and disliking, and the psychological and visceral sensations brought on by the neutral-hormonal bodily systems' (Oliver, 1997, p. 294).

Emotion, on the other hand, is described as a mental state of readiness which is the outcome of cognitive appraisals of events or thoughts (Lazarus, 1982; Lazarus, 1984; Oliver, 1997; Bagozzi et al, 1999). Emotions are often coupled with physiological processes (sweating, quickened heart rate) and can be expressed physically (facial expression, gesticulation) (Bagozzi et al, 1999; Atkinson, 2000). A distinction between the concept of a consumer's mood and their emotions also needs to be made. Mood is seen as a longer lasting state of being, whereas, emotion is more transient in nature (Westbrook and Oliver, 1991; Bagozzi et al, 1999; Mudie et al, 2003). In addition, mood is viewed as a less intense affective state than emotional responses which are higher in intensity (Westbrook and Oliver, 1991; Neo and Murrell, 1993; Bagozzi et al, 1999; Maute and Dubé, 1999; Mudie et al, 2003). A further point of differentiation between the two states is that emotions are seen as intentional, while moods are seen as unintentional (Bagozzi, 1999). For example, emotions are often a direct consequence of an object or referent (Burns and Neisner, 2006), while a mood is a global state of mind that is not accompanied with action tendencies (e.g. delight might be coupled with repurchase intentions). These distinctions must be kept in mind, particularly with regards to the

work of Oliver and Westbrook (1987, 1991; 1993) as they believe that as a measure of non-cognitive response to a consumption experience, affect is about consumer emotions (Mudie et al, 2003).

6.5.9 Empirical Studies on Emotions in Consumer Satisfaction

Within the customer satisfaction literature, the first study to explore the influence of positive and negative affect on satisfaction is Westbrook's (1987) study on two product categories; cars and cable television. It was found that positive and negative affect contribute independently and significantly to satisfaction judgements beyond the effects of the DE model. For both products, positive affect was positively and negative affect negatively related to satisfaction. This was the first study of its kind and led the authors to question whether the disconfirmation of expectations model should include an affective state. A follow up study by Westbrook and Oliver (1991) attempted to extend Westbrook's (1987) earlier findings through the examination of the affective structure of satisfaction. While they found that three emotional responses were important antecedents of satisfaction of newly purchased cars; pleasant surprise, interest and hostility, other findings in their study were very mixed. For instance, their analysis revealed that there existed a group that were positively satisfied yet had their experience in purchasing the car was largely devoid of affective reaction. According to this finding, it seems that only those who are at extreme ends of the satisfaction scale appear to elicit emotional responses. Another expansion on the determinants of satisfaction was put forward by Oliver (1993) who included positive affect (interest and joy) and negative affect (anger, disgust, contempt, shame, guilt, fear, sadness) as well as disconfirmation of beliefs.

Using two samples, car users and college course students, Oliver (1993) merged affect, disconfirmation and attribute-based satisfaction judgements. These were found to explain 85 per cent of the variance for the car user sample and 81 per cent for the course sample. He also found that consumers can experience both positive and negative affect at the same time, where each product attribute is evaluated by the consumer and their evaluation of the attribute is a potential source of positive or negative affect. In general, this stream of research has generally found that if a product exceeds the customers' expectations, positive emotions will result and vice versa (Oliver, 1993; Oliver and Westbrook, 1993). This valence-congruent direction has long been accepted within the customer satisfaction literature and the direct link between positive emotions and negative satisfaction is also robust (Westbrook, 1987; Dubé and Menon, 2000; Phillips and Baumgartner, 2002). More recent product research also supports this, however it has been found that the role of affect actually decreases over time with the role of cognition becoming more prominent (Homburg et al, 2006).

6.5.10 Empirical Studies on Emotion in Service Satisfaction

A number of studies have also examined the influence of emotions on customer satisfaction within a service context. One of the first studies to examine the role of emotional responses to service encounters was that of Price et al (1995). These authors believed that because the consumer is actively engaged in the service encounter emotions should be more significant than for other purchases. Unfortunately it was found that on average, consumers have little or no emotional response to service encounters. Contradictory to Price et al's (1995) findings, quite different results were found by Alford

and Sherrell (1996) where it was revealed that general and service provider affect had a direct effect on performance evaluations which subsequently influences satisfaction. In an attempt to resolve the issue, Liljander and Strandvik (1997) investigated whether customers experience emotions while consuming a service and whether they are related to perceived satisfaction in a labour force bureau. Assessment of the dimensionality of emotion revealed that negative emotions were more highly correlated with each other than with the positive emotions and vice versa. The results from a Principal Components Analysis were slightly less clear as a three factor solution was found which consisted of positive affect, guilt/humiliation and anger/depression. The analysis also seems to have been constrained to produce a two factor solution where all negative emotions loaded positively and all positive emotions loaded negatively on factor 1 (positive affect). Regression analysis was performed and it was found that emotions explained more of satisfaction than any other single construct (performance, disconfirmation). Interestingly, negative emotions had a stronger effect than positive emotions on satisfaction. From a managerial perspective, Liljander and Strandvik's (1997) results are critical as it was found that positive emotions did not influence satisfaction with the service but if negative emotions are experienced by the customer they will have a negative effect on their satisfaction with the service. In recent years, studies on satisfaction emotions in services have become even more confusing with some finding that cognitive evaluations are more salient than emotions (Mudie et al, 2003; Burns and Neisner, 2006) while others find that emotions are significant predictors of satisfaction (Krampf et al, 2003; van Dolen et al, 2004; Wong, 2004)

Hypotheses Development

While the services marketing literature is quite mixed with regards to the influence of emotion on satisfaction, the overarching view appears to be that satisfaction can be viewed as having affective and cognitive elements. From this the following hypotheses were developed:

H2a: Positive emotional responses will have a positive influence on satisfaction.

H2b: Negative emotional reactions will have a negative influence on satisfaction.

6.5.11 The Relationship between Service Quality and Customer Satisfaction

In the previous chapter, the service quality construct was assessed and its relationship to servicescape evaluations determined. However, as this dissertation contains both satisfaction and service quality judgements it is important that the delineation between them is addressed. Within the services marketing literature the differentiation between service quality and customer satisfaction has received a considerable amount of attention. The reason why these two constructs are often confused with one another is due to the fact that both constructs are based on the 'canonical model' (Wetzels, 1998). This model posits that consumers make evaluative judgements using *a priori* standards as a comparison. From this judgement a subsequent effect usually follows. Although the two constructs are very similar in terms of how they are cognitively processed, Oliver (1997) proposes that service quality and customer satisfaction differ on theoretical grounds. Firstly service quality is thought to reflect a number of distinct dimensions while

customer satisfaction can result from any aspect of a service encounter. Secondly, the expectations standards used for service quality is that of an 'ideal' or standard of excellence (Parasuraman et al, 1991) while customer satisfaction is based on an array of different standards such as predictive norms and experience based norms (Zeithaml et al, 1993). The third point of difference made by Oliver (1997) is that in order for an individual to attribute a level of service quality to a firm there is no requirement for the individual to have had firsthand experience with the service while for customer satisfaction, it is a basic necessity that the individual should have experienced the service. The final distinction made between the two is that customer satisfaction is thought to consist of both cognitive and affective attributes while service quality is primarily cognitively orientated.

The services marketing literature has also devoted considerable time to determining the sequential order of service quality and satisfaction. While early conceptualisations believed that customer satisfaction preceded service quality perceptions (Bitner, 1990; Bolton and Drew, 1991) due to both conceptual and empirical work, customer satisfaction is now widely regarded as the superordinate construct (Woodside et al, 1989; Cronin and Taylor, 1992; Anderson et al, 1994; Gotlieb et al, 1994; Rust and Oliver, 1994; Dabholkar et al, 2000; Brady and Robertson, 2001). The theoretical reasoning behind this causal sequence is that service quality, the more cognitive construct (White, 2006) should precede the more affective customer satisfaction as this is in line with Lazarus's (1982; 1984) appraisal-emotional response framework (Brady and Robertson, 2001).

In keeping with the literature, it is proposed that customers' evaluations of service quality will precede and have a direct effect on customer satisfaction. In the previous chapter it was found that service quality consists of two distinct constructs; employee/interaction service quality and outcome service quality thus requiring the following two hypotheses:

H3a: Outcome service quality evaluations will have a direct effect on customer satisfaction.

H3b: Employee service quality evaluations will have a direct effect on customer satisfaction.

6.6 Behavioural Intentions

Attitude theories, in particular Ajzen and Fishbein's (1980) theory of reasoned action and the theory of planned behaviour can be regarded as one of the original streams of research dealing with intentions (Söderlund and Öhman, 2005). Behavioural intentions are one of the major components of these theories and for Ajzen and Fishbein (1980) intentions are regarded as an individuals' commitment to behaviour. In the theory of reasoned action, Ajzen and Fishbein (1980) postulate that behaviour can be predicted from intentions that correspond directly to that behaviour. Within the services marketing literature, behavioural intentions (BI) are operationalised as indications of whether an individual will return to a particular service. Although behavioural intentions and Ajzen and Fishbein's (1980; 2001) models are widely recognised within the literature,

behavioural intentions and the loyalty concept have become entangled and are often used interchangeably. Although they are closely linked, theoretically and operationally they are distinct. Positive behavioural intentions imply that the consumer *intends* to remain loyal to the product or service. Whereas loyalty on the other hand refers to *actual* behaviour (Jacoby, 1971; Jacoby and Kyner, 1973). In the literature, loyalty has not been defined as simply as this, as there are those who conceptualise loyalty in terms of behaviour (Hallowell, 1996). and those who describe it as an attitudinal construct (Bloemer and de Ruyter, 1999; de Ruyter and Bloemer, 1999). The behavioural approach has been criticised by some for taking an overly narrow view as the amount of re-purchasing that is done by a particular individual may be as a result of a variety of situational factors such as non-availability, variety seeking and lack of provider preference (Day, 1969; Bloemer and de Ruyter, 1999). The attitudinal view of loyalty includes customers' intention to repurchase and their willingness of recommend the service to prospective customers and is very similar to our understanding of behavioural intentions. So as to avoid confusion, this dissertation will from this point on refer to attitudinal loyalty as behavioural intentions.

6.6.1 Types of Behavioural Intention

Behavioural intentions can be defined as being either social or economic in form. Social intentions are behaviours that impact the responses of customers to the firm such as word of mouth communication (Arndt, 1967; Buttle, 1998; Szymanski and Henard, 2001; Ladhari, 2007). Word of mouth communication has been acknowledged for many years as a major influence in what people do and know (Arndt, 1967). Indeed, as far back as

the early 1970s studies found that word of mouth communication carries more influence than marketer controlled media such as print or advertising (Day, 1971). Word of mouth communication can be directed at numerous aspects of the product or service including the product or service itself, the brand or the whole organisation. The forms which it may take can also vary from traditional face to face communication to electronic communication on various internet discussion forums. The valence of word of mouth communications is also of importance with negative word of mouth communications thought to have detrimental effects on brand image and future purchase intentions for prospective customers. According to Bendall-Lyon and Powers (2004) the second form of behavioural intention can be classified as economic intentions and have direct effects on financial performance (Zeithaml et al, 1996). Although economic intentions indicate whether a future relationship between the provider and seller will exist, it is often constrained by factors such as cost or physical access to the service. For this reason, word of mouth behaviours, which are free of such constraints, are of the utmost importance.

6.6.2 Antecedents of Behavioural Intentions

The most widely agreed upon antecedent of behavioural intentions is customer satisfaction with an extensive number of studies showing that behavioural intentions are an outcome of the satisfaction process (Oliver and Swan, 1989; Anderson and Sullivan, 1993; Anderson et al, 1994; Taylor and Baker, 1994; Danaher and Haddrell, 1996; Spreng et al, 1996; Mittal and Lassar, 1998; Anderson and Mittal, 2000; Jones et al, 2000; Caruana, 2002; Hellier et al, 2003). These studies have established that overall

customer satisfaction with a service is strongly associated with behavioural intentions to return to the same provider. As customer satisfaction is seen as a form of attitude, positive levels of this attitude are likely to influence behaviour. As discussed above, service quality evaluations are thought to directly effect perceptions of customer satisfaction. This research has been extended by examining how behavioural intentions influence this relationship and it has been found that service quality affects behavioural intentions indirectly through customer satisfaction (Anderson and Sullivan, 1993; Gotlieb et al, 1994; Patterson and Spreng, 1997). In more recent years service quality has also been regarded as a direct antecedent of behavioural intentions with a number of authors demonstrating this relationship (Parasuraman et al, 1988; Parasuraman et al, 1991; Boulding et al, 1993; Taylor and Baker, 1994; Zeithaml et al, 1996; Cronin et al, 2000). As part of this research the relationship between service quality and behavioural intentions will also be examined. Given the flexibility of structural equation modelling, whether the relationship between service quality and behavioural intentions is direct or indirect can easily be assessed. As per the recent work of Cronin et al (2000) it is believed that customer satisfaction and service quality will directly lead to favourable behavioural intentions. From this the following hypotheses are drawn:

H4: Customer satisfaction will have a positive effect on behavioural intentions.

H5a: Outcome service quality will have a positive effect on behavioural intentions.

H5b: Employee service quality will have a positive effect on behavioural intentions

6.7 Research Model

This chapter and in those chapters leading up to it have been building towards the development of a structural model in which the servicescape is seen as the primary antecedent of emotions and service quality evaluations while indirectly influencing customer satisfaction and behavioural intentions. The model can be divided into the affective and cognitive sub-models. The affective or emotion model posits that consumption emotions mediate the relationship between servicescape evaluations and customer satisfaction. This aspect of the model makes a number of notable advances. Firstly, the conceptualisation of the servicescape as a number of discrete dimensions permits the researcher to determine which of the dimensions, if not all, have the ability to elicit emotional responses. While much of the literature on environments and the servicescape view behavioural intentions as a direct outcome of emotional responses (Bitner, 1992; Sherman et al, 1997; Harris and Ezeh, 2008) it is proposed that customer satisfaction will mediate this relationship.

The second sub-model proposes that that the mediating role of emotional responses is insignificant and suggests that consumers engage in a more cognitively oriented processing system. In the previous chapter some support was found for the relationship between the servicescape and service quality, again this relationship will be tested so that a comparison between it and the affective model can be made. These paths are referred to as Hypotheses 6a to 6h (in Figure 6.4). As it was found in the previous chapter that

Employee/Interaction Service Quality is a direct antecedent of Outcome Service Quality this path is also included in the hypothesised model as H7.

H6a: Consumers perceptions of Servicescape Equipment will have a positive effect on perceptions of Employees

H6b: Consumers perceptions of Servicescape Equipment will have a positive effect on perceptions of Outcome Service Quality

H6c: Consumers perceptions of Servicescape Design will have a positive effect on perceptions of Employees

H6d: Consumers perceptions of Servicescape Design will have a positive effect on perceptions of Outcome Service Quality

H6e: Consumers perceptions of Servicescape Ambience will have a positive effect on perceptions of Employees

H6f: Consumers perceptions of Servicescape Ambience will have a positive effect on perceptions of Outcome Service Quality

H6g: Consumers perceptions of Servicescape Space & Hygiene will have a positive effect on perceptions of Employees

H6h: Consumers perceptions of Servicescape Space & Hygiene will have a positive effect on perceptions of Outcome Service Quality

H7: Employee/Interaction Service Quality will have a positive and direct effect on Outcome Service Quality

As per Figure 6.4, it is postulated that evaluations of the servicescape precede and have a direct effect on consumers' perceptions of service quality which in turn influences customer satisfaction evaluations. The hypothesised relationships between all of these constructs are illustrated in Figure 6.4 below. The affective paths in the model are shown in green with the cognitive paths in blue.

As well as testing whether an affective understanding of the servicescape is more important or if a cognitive model is more salient this model also makes a contribution in that no study to date has investigated the multiple direct links between the servicescape, emotion, service quality, customer satisfaction and behavioural intentions. To demonstrate this Table 6.1 has been compiled which lists the extant servicescape studies to date. It is also evident from Table 6.1 that many of the servicescape studies have tended to be conducted within more hedonic service situations such as casinos, restaurants and sports stadiums. Although some of the studies in Table 6.1 have been conducted in retail situations, some of them have chosen experimental designs and for this reason they lack external validity (Baker et al, 1994; Baker et al, 2002) others have

chosen student samples (Donovan and Rossiter, 1982) and some have chosen incomplete representations of the servicescape (Sherman et al, 1997; El Sayed et al, 2004). Indeed none of the studies in Table 6.1 have comprehensively measured the servicescape nor have these studies integrated each of the variables under scrutiny here with every study omitting at least one of the five constructs analysed here (servicescape, emotion, service quality, satisfaction and behavioural intentions).

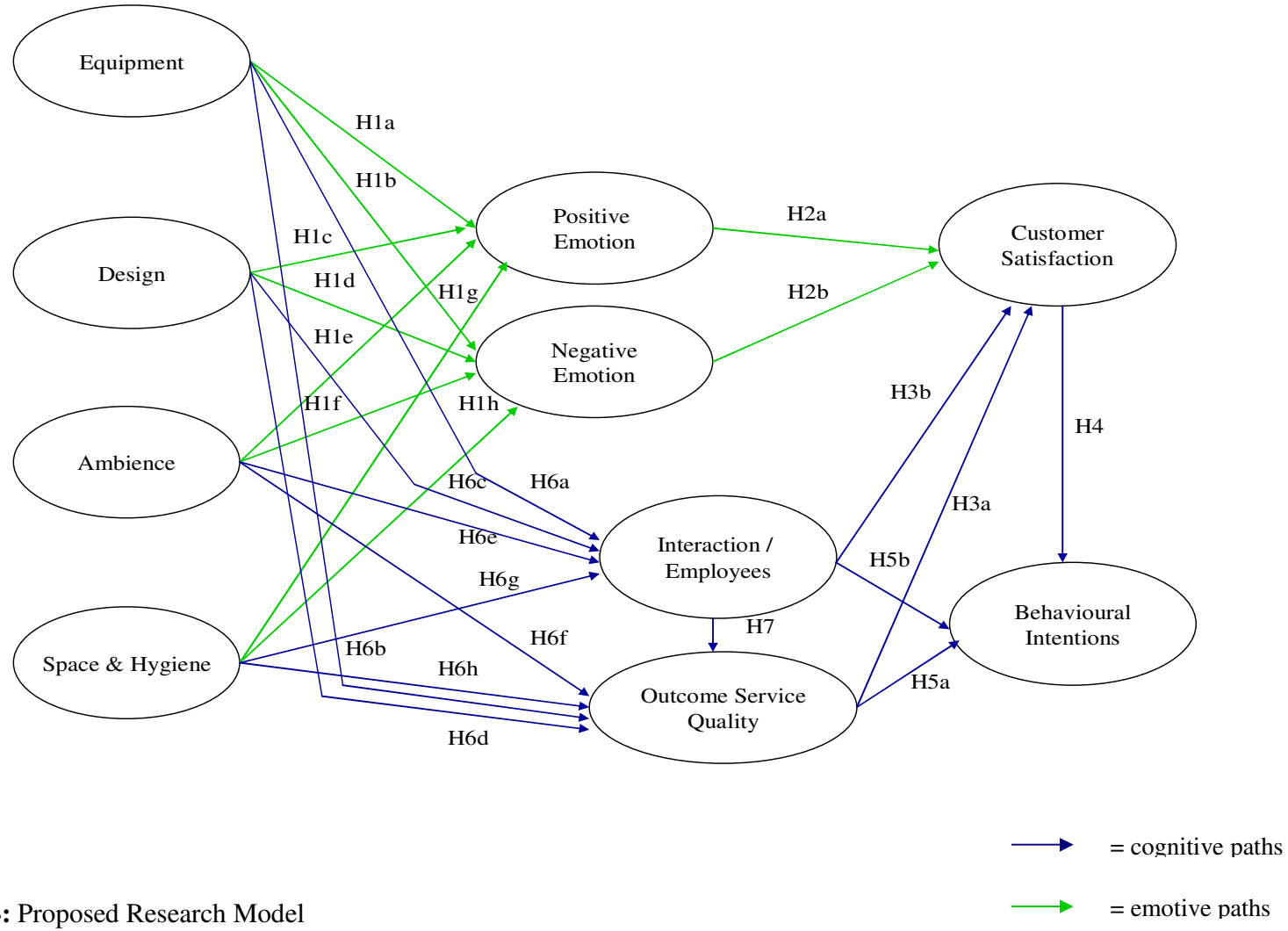


Figure 6.4: Proposed Research Model

Table 6.1: Servicescape Studies to Date

Author	Constructs Studied	Additional Constructs	Context
Baker et al (1994)	Environment, SQ	Merchandise quality, store image	Simulated Store
Baker et al (2002)	SS, Interpersonal SQ, BI	Merchandise quality, price perceptions, time/effort, psychic cost, merchandise value	Simulated Store
Donovan and Rossiter (1982)	SS, PAD, BI		Retail
El Sayed et al (2004)	SS, PAD, BI		Shopping mall
Grace and O'Cass (2004)	SE, CS, Arousal	Brand Attitude	Bank
Hansen (2002)	Servicescape elegance, PA Food quality		Simulated Restaurant
Harris and Ezeh (2008)	SS, BI	Variety seeking, perceived sacrifice, environmental factors	Restaurants
Hightower et al (2002)	SS, SQ, CS, BI	Waiting Time, Involvement	Sports stadium
Lee et al (2008)	SS*, Emotion, CS, BI		Festival
Lucas (2003)	SS, CS, BI		Casino
Pullman and Gross (2004)	Experience design, Emotion, BI		VIP hospitality tent
Reimer and Kuehn (2005)	SS, SQ		Bank Restaurant
Ryu and Jang (2007)	SS, PA, BI		Restaurant
Sherman et al (1997)	SS, PA, Purchase Behaviour		Retail
Tai and Fung (1997)	Information rate, PA, BI		Retail
Tang et al (2001)	SS, Excitement, BI	Resource expenditure	Game Centres
Wakefield and Baker (1998)	SS, Excitement, BI	Variety, Involvement	Shopping mall
Wakefield and Blodgett (1994)	SS, Excitement, BI	Satisfaction with the Servicescape, Crowding, Involvement	Sports stadium
Wakefield and Blodgett (1996)	SS, BI	Satisfaction with the Servicescape	Sports stadiums Casino
Wakefield and Blodgett (1999)	SS, Excitement, BI		Sports stadium Recreation centre cinema
Wall and Berry (2007)	Environment, SQ		Simulated restaurant

*Whether the items used to represent the servicescape in this study can be regarded as servicescape items is questionable. See Chapter Two for a discussion on this.
SS = Servicescape, PA = Pleasure and Arousal, PAD = Pleasure, Arousal and Dominance, SQ = Service Quality, BI = Behavioural Intentions

6.8 Conclusions

Servicescape theories and customer satisfaction theories have both asserted that emotional responses play a key role in their evaluation. Therefore emotion is the nexus which unites these two theories. Service quality serves as a second concatenator by proposing that servicescape evaluations are wholly cognitive in nature whereby servicescape evaluations influence perceptions of service quality and subsequently customer satisfaction and behavioural intentions. This integration of theoretical approaches has led to the research model described in Figure 6.4 and will be tested using structural equation modelling in the next chapter.

7.0 TESTING THE RELATIONSHIPS BETWEEN THE SERVICESCAPE AND KEY POST-CONSUMPTION PROCESSES

7.1 Introduction

In this chapter the proposed research model from the previous chapter is to be tested. This model brings together a variety of different literatures and in doing so makes a series of contributions to the literature. Some of the constructs in the model have not been included in any of the preceding analyses in the dissertation, therefore suitable measurement instruments for these constructs needed to be chosen. The chapter will begin by discussing the scales available in the literature and will provide a rationale for the final scales chosen. In accordance with Anderson and Gerbing's (1988) recommendations a two-step approach to the analysis will be taken, the first stage will assess the measurement model and the second stage then testing the hypothesised structural paths. From this series of analyses a number of interesting findings are gleaned and a discussion on the implications of these findings will conclude the chapter.

7.2 Measurement Instruments

The research model to be tested proposes relationships between the servicescape, service quality, emotion, customer satisfaction and behavioural intentions. As both the servicescape and service quality were analysed in previous chapters it would be repetitious to discuss the items used in their measurement here. Therefore, the items used in this final model to measure the servicescape are those fourteen items used in previous chapters. Similarly, service quality is seen as a two factor model where seven items

represent these latent constructs. In this section, the various different scales available to the researcher to measure emotion, customer satisfaction and behavioural intentions will be addressed. A rationale will also be provided for the final scale choice.

7.2.1 Emotion

Consumer researchers who seek to measure the relationship of emotions to other constructs have a variety of approaches from which to choose. They first must decide whether or not to use autonomic or self-report measures (Poels and Dewitte, 2006). Both of these methods have been applied in consumer research although they are quite different in their approach. Autonomic approaches measure visceral responses such as facial expression, frowning or physiological measures (such as sweating), which are often measured using psychogalvanometers to track changes in the electrical resistance of the skin (Poels and Dewitte, 2006). On the other hand, self-report measures use adjective checklists where the respondent must reflect on the emotions they experienced when exposed to a particular situation or stimulus (Bagozzi et al, 1999). Due to the relatively simple and concise way in which these scales measure emotions they have become the most frequently used procedure for measuring emotions (Buck and Georgson, 1997; Oliver, 1997). Autonomic methods require contrived lab settings and thus do not allow the objectives of this dissertation to be met as the real-life effect of the store environment and the service received cannot be measured. Given that a self-report measure is to be used here the next decision pertains to which of the many emotions scales to use. The remainder of this section will document the various scales that are available to a

researcher and will conclude with a discussion on the scale chosen here and the rationale for its selection.

As a number of self-report methods for measuring emotions exist, the first choice is to develop an ad hoc emotions scale for the specific problem at hand. This alternative suffers from the very same disadvantages that all ad hoc measures are hindered by. These disadvantages were discussed in detail in the servicescape scale section and refer primarily to the lack of rigour, validity and reliability of such makeshift measures. The second option is to adopt a scale that has been used by previous consumer researchers which was developed solely for marketing purposes (e.g. Richins, 1997). The final alternative is to go beyond the marketing literature to the field of social psychology where a variety of emotion scales have been published (Mehrabian and Russell, 1974; Izard, 1977; Watson et al, 1988).

Consumption Emotion Scales

Consumer emotion research has its beginnings in advertising effectiveness studies where authors explored whether the elicitation of particular emotions had an effect on outcome variables. In an attempt to link advertising content with emotional responses and attitude towards the brand, Holbrook and Batra (1987) developed their own scale based on an in-depth review of the literature resulting in a total of 94 items. A similar advertising study by Edell and Burke (1987) also compiled their own emotion list and these emotions were found to contribute uniquely towards advertisement attitude and beliefs about the brands' attributes. Their 52-item measure can be summarised into three underlying dimensions:

upbeat feelings, negative feelings and warm feelings. The problems with applying these measures to a service context are twofold. Firstly, the emotions that are experienced while watching an advertisement are likely to be lower in intensity as they are more secondary in nature than emotions that are experienced as a direct consequence of a consumption experience. Secondly, these scales were developed in an attempt to measure emotions that can be elicited from advertisements and thus may be context specific which would question their use in a services situation (Richins, 1997).

It was not until Richins's (1997) consumption emotion scale was introduced that any attempt was made to overcome the context specificity of the above scales. An objective of Richins (1997) was to develop a scale that would include a broad spectrum of emotions that are frequently experienced in a wide range of situations. After extensive empirical research using multi-dimensional scaling, the Consumption Emotion Set (CES) was derived. The full version of the scale includes 64 items, however Richins (1997) does recommend that it is not necessary to use the scale in its entirety if certain emotions are not likely to be experienced. This rather arbitrary approach to the scales' application seems to negate the point of so carefully constructing the scale. Further to this, when the scale content is analysed, it becomes evident that a considerable number of the emotion sub-groups may not be applicable here (e.g. 'envy', 'loneliness', 'romantic love', and 'love').

While Richin's (1997) scale was developed for use in any consumption situation, Yoo et al (1998) sought to develop a scale that can be applied specifically to service contexts.

Using ethnographic interviews, Yoo et al (1998) identified a number of service specific emotions. These emotions were then translated into a scale which consisted of an equivalent number of both positive (pleased, attractive, excited, contented, pride) and negative (ignored, anxious, nullified, displeased and angry) emotions. Using structural equation modelling, the relationship of these emotions to store attitude and a number of antecedents (atmosphere, after sale service, salesperson service, value of merchandise etc.) was explored. The model was found to fit very well, however the standardised estimates for paths between all of the antecedents of emotions were quite low ranging from .07 (atmosphere) to .30 (value). This suggests that this scale is not good at measuring the kind of emotions that can be elicited from an environment or from services. It also implies that emotional responses were of little importance in this study, a finding that goes entirely against the dominant thinking in this area.

Other consumer researchers have taken an even less rigorous approach in their choice of emotion scale. One such study is that of Alford and Sherrell (1996) in which they investigated the role of emotion on consumers' satisfaction with services. Emotions were measured on a general level and also with respect to the emotions that are elicited in response to their contact with a salesperson. These emotions were measured using semantic-differentials indicating their like/dislike for the category (likeable, pleasant and nice). Liljander and Strandvik (1997) also quite randomly chose their affective attributes, using the literature and earlier studies as bases for their inclusion. Despite this, their seven-item scale (happy, hopeful, positively surprised, angry, depressed, guilty and humiliated) was found to have good reliabilities in both their original study (Liljander

and Strandvik, 1997) and in later replications (Yu and Dean, 2001). A further modification to this scale was made by White and Yu (2005) who added 'regret' and 'disappointment', thus creating an overall scale of three positive emotions (happy, hopeful and positively surprised) and six negative emotions (angry, depressed, guilty, humiliated, disappointed and regretful). The use of more negative emotions than positive emotions biases the results in its very design. Others have done quite the antithesis of this by including positive emotions only, thus ignoring a whole dimension of emotional responses (Bloemer and de Ruyter, 1999; Homburg et al, 2006). The reason why these authors chose positive emotions only is unclear and seems to have been used for convenience purposes. As prominent theories of emotion such as those presented by Russell (1980) and Watson and Tellegen (1988) present models which contain an equal number of emotions that capture both positive and negative affects, it seems sensible to follow in this line of thinking whereby a balance between the two is achieved.

Psychology Emotion Scales

Emotion scales from the field of psychology fall into two groups: discrete emotion sets and dimensional structures of emotion. Izard's (1977) Differential Emotions Set (DES II) is a scale that was developed to measure what Izard (1977) considered to be fundamental emotions. These emotions are thought to be fundamental across species with the ten primary emotions consisting of: interest, joy, surprise, sadness, anger, disgust, contempt, fear, shame and guilt. This scale has been applied in both service and product situations and has been found to be reliable (Westbrook, 1987; Westbrook and Oliver, 1991; Krampf et al, 2003; van Dolen et al, 2004). However, the major fault of

this scale is that it contains a greater number of negative emotions (seven) than positive emotions (three). A recent article by Wood and Moreau (2006) claimed that they overcame this problem by adding four positive items to the scale thus balancing the number of positive and negative emotions. However, upon examination of the items used by Wood and Moreau (2006), only four of Izard's (1977) original ten items were actually included (sad, joyful, angry and scared), with the reason for this discrepancy not included in their methodology. A theoretically similar approach to Izard (1977) is taken by Plutchik (1980) whose conceptual model is based on the idea that emotions must be considered from an evolutionary point of view. In this model Plutchik (1980) describes eight basic patterns of emotion (acceptance, joy, anticipation, anger, disgust, sadness, surprise and fear) where mixtures of these emotions can result in many hundreds of mixed emotions. The scale accompanying this theory is outlined in his text and consists of 43 emotion pairs. Consumer behaviour applications of this scale are few, however Zeitlin and Westwood (1986) developed and tested a 43-item communication checklist that was based upon Plutchik's (1980) work. However, in a comparison of Plutchik's (1980) emotional spectrum to Mehrabian and Russell's (1974) Pleasure-Arousal-Dominance (PAD) model, it was found that across a number of forms of analyses that the PAD model captured more information about the emotional character of consumption experiences (Machleit and Eroglu, 2000). This finding, and the scales' length, seems to have hindered its use in marketing studies and it is not popular as a result.

The first dimensional emotion scale to appear in the literature was that of Mehrabian and Russell (1974) who developed a measure of emotions to assess individuals' approach or

avoidance tendencies for certain environments. They assume that pleasure, arousal and dominance, which are measured using semantic-differential scales, are the three basic human emotional reactions. Many applications of this scale have been made in consumer research and it is used extensively in store atmospherics studies (Donovan and Rossiter, 1982; Donovan et al, 1994; Sherman et al, 1997; Tai and Fung, 1997; 1999; Chebat and Michon, 2003; Gilboa and Rafaeli, 2003; El Sayed et al, 2004; Bigné et al, 2005). Despite this scales' widespread use, it has been criticised for serving only to measure individuals' reactions to the environment and for not covering the range of emotions that may be experienced by shoppers (Richins, 1997; Machleit and Eroglu, 2000). For example, certain emotions may be elicited entirely independently of the environment that can be directly attributed to the consumers' interaction with a salesperson or other customers in store.

A second problem with the PAD scale is that it has been shown to be very mixed in its ability to predict outcome behaviours. For example in a replication of Mehrabian and Russell's (1974) model, Donovan and Rossiter (1982) found that of the three emotion dimensions, mixed support was found for the arousal dimension with it only relating to the affiliation approach behaviour. Similarly, it was also found that an interaction effect between pleasure and arousal existed whereby arousal emerged as a predictor of time and affiliation. The third emotion dimension, dominance was not significantly related to any of the approach and avoidance measures. This tenuous finding, along with the small amount of variance it accounted for (8.1% in a principal components analysis) led Donovan and Rossiter (1994) and many others (Russell and Pratt, 1980; Sherman et al,

1997; Chebat and Michon, 2003) to exclude this dimension in further studies. Further evidence for the unreliability of the Pleasure-Arousal measure can be seen in Donovan et al's (1994) follow up study. Here it was found that arousal no longer influenced approach and avoidance behaviours in pleasant environments.

A third study which also employed the Pleasure-Arousal measure is that of Sherman et al (1997). Two of the dependent measures in this study, 'money spent' and 'time spent' are very similar to the two dependent measures in Donovan et al's (1994) study ('extra time' and 'unplanned spending'), yet a number of contradictions were found. While Donovan et al (1994) found that pleasure had a significant impact on time spent in store for Sherman et al (1997) no support for this could be found. Sherman et al (1997) also included two additional dependent variables; store liking and number of items purchased with the predictive ability of the Arousal and Pleasure dimensions being very mixed.

In recent years, Foxall and colleagues (1999; 2001; 2006) have taken a renewed interest in the dominance dimension, proposing that dominance can be used to predict consumer behaviour in different environments. This is incorporated into their Behavioural Perspective Model which posits that emotional responses can be 'predicted from two elements of situational influence: the scope of the consumer behaviour setting and the pattern of utilitarian (functional benefits of consumption) and informational (symbolic benefits such as social status and self-esteem) reinforcement' (Yani-de-Soriano and Foxall, 2006, p. 409).

Despite the re-emergence of the dominance dimension by Foxall et al (1999; 1999; 2006), cogent arguments discounting the use of the PAD measure have been put forward by Richins (1997). While she accepts that the PAD measure may be suitable to measure emotional responses to environmental stimuli, she questions the validity of its use in instances where emotional responses to non-environmental cues are included in a study. As this dissertation will be measuring emotional responses to both the servicescape and as an antecedent to customer satisfaction, the PAD and derivatives of it (e.g. Russell and Pratt's (1980) excitement measure) may not be suitable in this instance. In addition, the PAD scale does not claim to measure individual emotions as such, but whether pleasure, arousal or dominance, are educed by the environment. Therefore, it cannot be concluded whether specific emotional states such as fear, delight or anger are experienced by individuals by their PAD scores. This research hopes to address this limitation. Further to this, Donovan et al (1994) note that anecdotal feedback from respondents suggested that some items on the scale were difficult to relate to experiences felt in store which may have caused some of the ambiguous findings found. A final problem with the PAD measure is that for those who have used the scale in full (Pleasure-Arousal-Dominance) and those who have used the shortened version (Pleasure-Arousal), very mixed results have been found with most authors reporting contradictory results thus questioning the applicability of the PAD measure.

An intuitively simple scale (Bagozzi et al, 1999) that overcomes the shortcomings of the PAD measure is the PANAS scale (Watson and Tellegen, 1988). This scale is based upon Watson and Tellegen's (1988) consensual two-factor (positive affect/negative

affect) model which is a 45º rotation of Russell's (1980) bi-dimensional circumplex model of affect (pleasantness-arousal). The advantage of this scale is that it is equally balanced, containing an equivalent number of both positive (ten) and negative (ten) emotions. The positive emotions reflect the extent to which a person feels enthusiastic, alert and in a state of pleasurable engagement. Negative affect on the other hand is a dimension that relates to distress and includes contemptuous and nervous feelings. Unlike the PAD measure which has been designed to measure emotional reactions to environments only, this scale is not context specific and can thus be applied to any situation. Consumer research applications of this scale include an assortment of services (Bloemer and de Ruyter, 1999), a recreational wildlife theme park and a symphony concert (Oliver et al, 1997), and a selection of products (Mano and Oliver, 1993). These studies have found that this scale is both reliable and stable. However, none of the above have used this scale in its entirety, instead it has either been reduced to just positive emotions (Bloemer and de Ruyter, 1999), or used in conjunction with a second scale (Mano and Oliver, 1993; Oliver et al, 1997).

Combining Scales

Combining two scales to create one emotions scale is also quite popular in the literature. What generally occurs is that items are taken from one scale and put together with some or all of the items from another scale (Mano and Oliver, 1993; Phillips and Baumgartner, 2002; Pullman and Gross, 2004). The rationale for amalgamating items is somewhat unclear with some providing no justification for their choice (Mano and Oliver, 1993; Phillips and Baumgartner, 2002). Pullman and Gross (2004) give reasons for

incorporating two scales, and cite Richins (1997) to justify their stance, saying that given the complexity of the context, established scales did not represent the diversity of emotions that would be experienced in the service. Blending two scales together in this way appears to be a rather unstable method of creating a scale and should be avoided if at all possible.

Rationale for Scale Choice

Out of the emotions scales developed by consumer researchers, the most superior is Richins (1997) Consumption Emotion Set (CES) as it underwent substantial testing and development procedures. However, lack of parsimony is a major problem with this scale as it consists of 64 items. In addition, its applicability has been questioned by Mudie et al (Gilboa and Rafaeli, 2003) who found that many of the emotion items in the scale were not experienced across many of the services included in their study. Other consumer scales have either been found to be poor measures of their antecedents (Yoo et al, 1998) or have rather arbitrarily constructed their content (Alford and Sherrell, 1996; Dube et al, 1996; Liljander and Strandvik, 1997; Yu and Dean, 2001; White and Yu, 2005). Ideally it would be preferable to use a scale that has been developed specifically for services that has proven to be both parsimonious and reliable. Unfortunately this was not the case with no service emotions scale available. This then leaves only those scales from the field of psychology as suitable candidates. Emotion scales that were possible contenders were scales discussed in the Psychology Emotion Scales section: Mehrabian and Russell's (1974) PAD, Izard's (1997) DES II, Plutchik's (1980) Emotion Set and Watson et al's (1988) PANAS.

A study by Machleit and Eroglu (2000) empirically compared Izard's DES II (1977), Plutchik's (1980) measure and the aforementioned Mehrabian and Russell (1974) PAD instrument. Using a mix of student and non-student respondents, individuals were asked to complete the questionnaire immediately after their next shopping trip. Shopping contexts included in the study were: malls, department, grocery and discount stores. Both canonical correlation analysis and regression analysis was used to compare the three measures. Using canonical correlation analysis it was found that the both Plutchik (1980) and Izard (1977) measures explained more variance than Mehrabian and Russell's (1974) PAD. Using customer satisfaction as the dependent variable, regression analysis also confirmed the superiority of Plutchik (1980) and Izard's (1977) scales over the PAD measure. However, the choice between Izard's (1977) and Plutchik's (1980) measures is not very clear cut as Izard's (1977) is slightly superior in the canonical correlation analysis and Plutchik's (1980) having the advantage in the regression analysis. For this reason Machleit and Eroglu (2000) recommend that the decision on whether to use one of these scales needs to made on different grounds.

Given the volume of atmospheric studies that have applied Mehrabian and Russell's PAD scale (Donovan and Rossiter, 1982; Hui and Bateson, 1991; Donovan et al, 1994; Gilboa and Rafaeli, 2003; Martin et al, 2008) it was initially thought that this measure would be the most suitable. However, in addition to the shortcomings found by Machleit and Eroglu (2000), a second problem with this scale is inherent in its very design. This scale was constructed with the intention of being used solely as an instrument which can gauge

individuals' emotional responses to environments. The structural model under scrutiny here not only views emotional responses as direct consequences of emotional responses but also hypothesises that emotional responses will lead directly to customer satisfaction evaluations. It is assumed that individuals arrive at a level of satisfaction upon evaluation of both environmental and non-environmental cues thus requiring a non-environmental emotions scale to be used. For this reason, this scale is no longer a viable option for this research problem.

As the relationship between emotion and satisfaction are to be assessed, it is key that the chosen scale contains a sufficient number of positive emotion items so that the depth and intensity of positive emotions experienced by customers is accurately captured. The reason for this is because it has been found in the past that when satisfaction is measured, the distribution of responses falls heavily on the satisfied end of the continuum, with satisfaction responses possessing 'a distribution in which a majority of responses indicate that customers are satisfied and the distribution itself is negatively skewed' (Peterson and Wilson, 1992, p. 62). Therefore, using an emotions scale that contains a sufficient number of positive emotions is essential so that insights into the full spectrum of emotions that are experienced by consumers can be garnered. Negatively biased scales such as Izards's (1977) DES II (six negative emotions, one neutral and three positive) would not capture the variety of positive emotions that may be experienced by consumers. In fact, if Izard's (1977) DES II contained more positive emotions than negative emotions it would not pose as much of a problem as the skewness of the scale would reflect the skewness of the satisfaction response. Unfortunately this is not the case

and due to this it was decided that Izard's (1977) DES II is not a suitable measure for this study.

With regards to Plutchik's (1980) measure, the cumbersome length of the scale would make it too unwieldy to administer (64 items). Although it has been used in the past by consumer researchers, there are key differences between their studies and the current research situation. For instance, the scale has either been reduced to a smaller number of items or else has been used in an experimental study with only two other scales (Machleit and Eroglu, 2000). Given the number of constructs examined here (service quality, servicescape, customer satisfaction, behavioural intentions and emotion), and that real-life consumers in a non-experimental context are to be used, using Plutchik's (1980) measure would be unmanageable as it would more than likely fatigue respondents.

Considering that all of the above scales have some serious drawbacks, it was decided to utilise Watson et al's (1998) twenty-item PANAS instrument (see Table 7.1 below). This scale was chosen for a number of key reasons: Firstly, it is a balanced scale containing an equal number of both positive and negative emotions which represents a full spectrum of emotions capturing each dimension of the circumplex model. The multitude of positive emotions (ten) will hopefully be a sufficient number of positive emotions so that a true reflection of the negatively skewed satisfaction response can be attained. Secondly, the authors adhered to strict methodological criteria in its development resulting in a scale that is both reliable and valid. Thirdly, this scale was designed for use in any situation

and has been demonstrated in the past to be a reliable measure in service contexts (Mano and Oliver, 1993; Oliver et al, 1997).

Table 7.1: Watson et al's (1988) PANAS

Please state **to what extent** you experienced the following **emotions while in this store**. Please circle a number beside **each** emotion.											
	Not at all	A little	Moder-ately	Quite a bit	A lot		Not at all	A little	Moder-ately	Quite a bit	A lot
Interested	1	2	3	4	5	**Distressed**	1	2	3	4	5
Nervous	1	2	3	4	5	**Guilty**	1	2	3	4	5
Excited	1	2	3	4	5	**Scared**	1	2	3	4	5
Upset	1	2	3	4	5	**Hostile**	1	2	3	4	5
Strong	1	2	3	4	5	**Enthusiastic**	1	2	3	4	5
Irritable	1	2	3	4	5	**Proud**	1	2	3	4	5
Alert	1	2	3	4	5	**Ashamed**	1	2	3	4	5
Inspired	1	2	3	4	5	**Determined**	1	2	3	4	5
Jittery	1	2	3	4	5	**Attentive**	1	2	3	4	5
Active	1	2	3	4	5	**Afraid**	1	2	3	4	5

7.2.2 Customer Satisfaction

Following the decision to use the PANAS scale to measure emotions the selection of a suitable customer satisfaction measure needed to be made. In order to select the most appropriate scale, the vast literature on customer satisfaction was reviewed. This process revealed a variety of different approaches to measuring customer satisfaction which will be discussed next.

Methods of Measurement

Measurement of customer satisfaction typically takes the form of direct, subjective estimation by consumers (Westbrook and Oliver, 1981; Yi, 1990; Peterson and Wilson, 1992). Indirect methods through customer complaints and repeat purchases is also

possible (Yi, 1990). These means of data collection are advantageous in that they involve no intrusion and are closely related to satisfaction. Two problems exist with this method. Firstly, this method does not provide an indication of how a broad spectrum of consumers feel about a particular brand or service. Instead, it provides businesses with feedback from a small minority who fall close or near to the end-points of the distribution. Furthermore, the indirect nature of this method does not permit structural modelling and as a result has not been chosen as a data collection method for this dissertation. This complication is not a problem per se, but renders this method inapplicable for the present study.

Direct Measures: Single-item vs. Multiple-Item Measures

For those choosing to use direct methods the choice primarily lies between either a single item or multiple item measure. The use of single-item measures is prevalent in the literature (1980; Tse and Wilton, 1988; Bitner, 1990; Alford and Sherrell, 1996; Liljander and Strandvik, 1997; Mittal et al, 1998; Bloemer and de Ruyter, 1999) and are chosen over alternative methods given their simplicity. Single-item scales are usually anchored by 'very satisfied' to 'very dissatisfied' with a number of levels between which tend to range from four to seven. Westbrook (1980) tested the reliability of a variety of single-item scales, including; the Delighted-Terrible Scale, Percentage Scale, Semantic-Differential Need Scale and a Content Analytic approach. Previous to Westbrook's (1980) study, the Delighted-Terrible scale, had only been utilised in sociology studies to assess individuals' job satisfaction and life satisfaction. Westbrook (1980) found that the Delighted-Terrible scale had reasonable reliabilities (α = .65-.84) and was convergent

with other measures. In addition to this, it was found that it reduced the skewness of the satisfaction response. Despite the respectable reliability statistics achieved, the use of single item measures oversimplifies a complicated, multi-faceted phenomenon (Churchill, 1979; Danaher and Haddrell, 1996). In employee satisfaction surveys, single-item scales have also been found to overestimate individuals' actual satisfaction levels which may account for the marked negative skewness found. Further to this, single-item scales are not appropriate for structural equation modelling as they do not allow for the assessment of goodness of fit, unidimensionality, or convergent and discriminant validity (Howell, 1987). On the other hand, multiple item measures reduce the ambiguity of constructs in structural equation modelling (Kenny and McCoach, 2003). Even for instances where structural equation modelling is not used, multiple-item scales are superior in that they reduce measurement error, thus enhancing reliability (Westbrook and Oliver, 1981). Given the many shortcomings of single-item scales, multi-item scales are by far the more desirable alternative for measuring customer satisfaction (Yi, 1990) and will thus be used here.

As with single-item approaches, an assortment of multiple-item customer satisfaction instruments have appeared in the literature. Using MTMM analysis, a selection of multi-item batteries were tested for reliability, convergent and discriminant validity by Westbrook and Oliver (1981). Out of a number of scales, it was found that Likert scales and semantic-differentials are the most reliable (alpha coefficients ranging between .90 and .96). It has been found by some that semantic-differentials fit data better than Likert scales (Friborg et al, 2006). However, similar studies conducted in a marketing context

have found only marginal differences in the results produced by each method (Kassarjian and Nakanishi, 1967; Menezes and Elbert, 1979; Flamer, 1983). Given this, the choice between semantic differential or Likert scales needs to be made on other grounds. After a careful review of the literature, Likert scales were chosen over semantic-differentials for the following reasons: Firstly, Irish respondents would be more familiar with the Likert format than the semantic-differential as many commercial marketing research programmes use this presentation. Secondly, it has been found that semantic-differentials increase cognitive effort, thus leading to high levels of random error (Stephens, 1991). Thirdly, it has been recommended that if the sample are likely to come from a poorer demographic, then Likert scales are more appropriate given their simplicity (Menezes and Elbert, 1979). It is highly likely that a diverse range of respondents from varying backgrounds will be surveyed here thus making Likert-type scales more suitable.

When Likert scales are used to measure customer satisfaction, they tend to take two forms; an attribute-specific approach or a generic approach. Attribute measurement techniques include items on key product or service dimensions and the responses to the individual attributes are then summed to produce an overall satisfaction score. A number of authors have adopted this mode of customer satisfaction measurement (Churchill and Suprenant, 1982; Jayanti and Jackson, 1991; Spreng et al, 1996). The alternative method does not include evaluative items on specific service facets, per se, instead it uses a broad measure that could be used for any product or service. Oliver's (1980; 1997) customer satisfaction scale is an example of such an instrument and it is the scale that appears most frequently in the literature. When originally published (Oliver, 1980), this scale consisted

of the six items in Table 7.2, but was later expanded to consist of twelve items (Oliver, 1997).

Table 7.2: Customer Satisfaction Items

Customer Satisfaction Items	Comment
I am satisfied with my decision to visit X	Satisfaction anchor
My choice to attend X was a wise one	Success attribution
If I could do it over again, I'd go to a different X	Regret
I feel bad about my decision to go to X	Failure attribution
I am not happy that I visited X	Negative affect
I am sure I did the right thing by coming to X	Success attribution

Each of these items reflects a different satisfaction dimension. Success attribution is represented though the 'My choice to attend this *x* was a wise one', while overall performance is expressed through, 'I am satisfied with my decision to visit X'. Although the expanded twelve item scale has been proven to be reliable (Wang and Wallendorf, 2006), it is felt that given the array of constructs to be studied in this dissertation the shorter version would be more suitable. The shorter, six-item instrument has also shown reliable Cronbach alphas (.82-.92) with studies using as few as three items showing reliable internal consistencies (Gotlieb et al, 1994). It has been recommended by Oliver (1997) that if less than twelve items are to be used, deletion of any item is acceptable as long as it is not the satisfaction anchor itself ('I am satisfied with my decision to visit to *x*'). The final customer satisfaction scale used here consisted of four of Oliver's (1980) items. Two items from the original subset were excluded as it was felt they were worded in a highly emotive manner which could possibly cause discriminant validity problems with the emotions scale. These items were, 'I feel bad about my decision to go to X', and 'I am not happy that I visited X'. Similarly as these items are negatively worded they may pose problems for the respondents. However, one negatively worded item was

retained as a check for reliability. The final customer satisfaction items can be viewed in Table 7.3 overleaf:

Table 7.3: Final Customer Satisfaction Items

Customer Satisfaction Items	Comment
I am satisfied with my decision to visit X	Satisfaction anchor
My choice to attend X was a wise one	Success attribution
If I could do it over again, I'd go to a different X	Regret
I am sure I did the right thing by coming to X	Success attribution

7.2.3 Behavioural Intentions

The final measurement decision was in choosing a behavioural intentions scale. In the literature, those who have measured behavioural intentions have tended to use either one of two scales. The first of these is Oliver and Swan's (1989) four item semantic differential. When this scale is used respondents are asked whether they would return to the service again by rating the following bipolar adjective scales, 'likely-unlikely', 'very probable-not probable', very 'possible-impossible', and 'certain-no chance'. While this scale has been used relatively widely in the services marketing literature (e.g. Wakefield and Baker, 1998; Jones et al, 2000), Zeithaml et al's (1996) thirteen-item scale has far surpassed Oliver and Swan's (1989) scale in terms of its popularity with a large number of authors using it in the literature (de Ruyter and Bloemer, 1999; Cronin et al, 2000; Yu and Dean, 2001; Hightower et al, 2002; Brady et al, 2005; White and Yu, 2005; White, 2006; Dean, 2007; Goode and Harris, 2007; Harris and Ezeh, 2008). This scale consists of five facets of behavioural intentions which are categorised as loyalty intentions, switching intentions, price sensitivity, external and internal responses. A complete list of these items is described in Table 7.4 below:

Table 7.4: Zeithaml et al's (1996) Behavioural Intentions Measure

Behavioural Intentions	Item Wording
Loyalty	Say positive things about XYZ to other people Recommend XYZ to someone who seeks your advice Encourage friends and relatives to do business with XYZ Consider XYZ your first choice to buy _____ services Do more business with XYZ in the next few years
Switch	Do less business with XYZ in the next few years Take some of your business to a competitor that offers better prices
Pay more	Continue to do business with XYZ if its prices increase somewhat Pay a higher price than competitors charge for the benefits you currently receive from XYZ
External response	Switch to a competitor if you experience a problem with XYZ's service Complain to other customers if you experience a problem with XYZ's service Complain to external agencies, such as the Better Business Bureau if you experience a problem with XYZ's service
Internal response	Complain to XYZ's employees if you experience a problem with XYZ's service

The internal consistency of each of the sub-scales varied considerably however, the loyalty scale, with its five items was found to have excellent reliabilities with alphas ranging from .93 to .94. Those who have employed this battery have varied greatly in the number of items they have used which ranges from three (Cronin et al, 2000; Brady et al, 2005), five (Bloemer and de Ruyter, 1999; Dean, 2007), six (Harris and Ezeh, 2008), seven (Goode and Harris, 2007), ten (Hightower, 2002), eleven (White and Yu, 2005), twelve (Yu and Dean, 2001) to all thirteen items (White, 2006).

Other authors have developed their own measures of behavioural intentions which are very similar in terms of content to Zeithaml et al's (1996) items except that the wording differs slightly (Boulding et al, 1993; Taylor and Baker, 1994; Andreassen and Lindestad, 1998; Dabholkar et al, 2000; Chandon et al, 2005; Martin et al, 2008). Generally these

scales contain quite a small number of items from two (Andreassen and Lindestad, 1998; Dabholkar et al, 2000) to three items (Taylor and Baker, 1994; Martin et al, 2008) and include an intention to return item and a word of mouth behaviour question.

Given that Zeithaml et al's (1996) has been proven to be both reliable and popular in the literature this scale was chosen. Although the more scale items the better (Kline, 2005) for structural equation modelling, it was decided to use this scale in its abbreviated format taking items from the loyalty subset only. Four of these items were used. The fifth item, 'Do more business with XYZ in the next few years' was felt to be irrelevant to the context of the current research problem and was thus not used.

7.3 Assessing the Measurement Model

According to Anderson and Gerbing (1988) a two step approach when undertaking structural equation modelling is the most suitable. This involves ensuring that constructs in the model meet criteria for convergent and discriminant validity. This next section of the chapter will begin by first assessing the constructs for convergent validity. Following this discriminant validity was checked.

7.3.1 Confirmatory Factor Analysis for Emotions

Two separate confirmatory factor analyses were run for emotions; one for positive emotions and the second for negative emotions. As per Watson et al's (1988) PANAS it was hypothesised that all ten positive emotions would load onto the one construct. The initial CFA on which the ten items loaded onto the positive emotion construct was found

to be a poor fitting model with the following fit statistics: Chi-Square = 144.37, df = 35, P-value = 0.00000, RMSEA=0.094, NFI=0.87, CFI=0.90. Inspection of the output indicated that a number of items were of dubious reliability and these are highlighted in red in Table 7.5 below.

In addition to having very low reliabilities many of these items had high residuals which also points towards instability. Therefore, the five items highlighted in red were removed from further analyses. Inspection of the refined five-item (active, determined, enthusiastic, proud, attentive) positive emotions construct finds an acceptable fit (Chi-Square = 8.69, df = 5, P-value = 0.12189, RMSEA = 0.046, NFI=.98, CFI=.99).

Table 7.5: Initial CFA for Positive Emotions

	Estimates	**Standardised Loading**	**t-value**	**R^2**
Interest	0.58	0.51	9.63*	0.26
Excited	0.20	0.35	6.31*	0.12
Strong	0.42	0.47	8.78*	0.22
Alert	0.60	0.48	8.87*	0.23
Inspired	0.38	0.53	9.94*	0.28
Active	0.77	0.68	13.75*	0.47
Determined	0.71	0.71	14.57*	0.52
Enthusiastic	0.73	0.72	14.72*	0.51
Proud	0.63	0.72	14.61*	0.51
Attentive	0.68	0.62	12.00*	0.38
Model fit: Chi-Square=144.37, df=35, P-value=0.00000, RMSEA=0.094 * = significant at the $p = 0.01$ level				

The negative emotions items proved much more unstable with the initial confirmatory factor analysis having a fit of Chi-Square = 558.57, df = 35, P-value = 0.00000, RMSEA = 0.206, NFI=0.57, CFI=0.59.

Table 7.6: Initial CFA for Negative Emotions

	Estimates	Standardised Loading	t-value	R^2
Nervous	0.08	0.34	6.18	0.12
Upset	0.24	0.61	11.89*	0.37
Irritable	0.41	0.79	16.58*	0.62
Jittery	0.17	0.50	9.32*	0.25
Distressed	0.32	0.81	17.28*	0.66
Guilty	0.13	0.41	7.46*	0.17
Scared	0.14	0.65	12.74*	0.42
Hostile	0.09	0.54	10.12*	0.29
Ashamed	0.00	0.02	0.28	0.00
Afraid	0.07	0.28	4.90*	0.08
Model fit: Chi-Square = 558.57, df = 35, P-value = 0.00000, RMSEA = 0.206 * = significant at the p = 0.01 level				

The items 'nervous', 'guilty', 'ashamed' and 'afraid' were all subsequently removed and the CFA was re-run. Unusually, this disimproved the fit substantially with the following fit statistics found: Chi-Square = 271.55, df = 9, P-value = 0.00000, RMSEA = 0.287, NFI = 0.72, CFI = 0.73. With this configuration it was found that the 'jittery' item had very large Theta-Delta modification indices and thus removed. Fit slightly improved with this six item solution, Chi-Square=82.78, df=5, P-value=0.00000, RMSEA=0.210, NFI=0.87, CFI=0.88 however further modifications were necessary. Inspection of the output revealed the 'hostile' item was a candidate for deletion with a low r^2 (.22) and with high theta-delta modification indices across three items. This final four item formation (upset, irritable, distressed, scared) produced model fit statistics that were markedly improved however, they remained outside acceptable levels (Chi-Square = 14.96, df = 2, P-value = 0.00056, RMSEA = 0.135, NFI = 0.97, CFI = 0.97).

Despite the unacceptable fit of the negative emotions confirmatory factor analysis, it was decided to proceed to the next stage of the analysis. At this stage, the four-item negative

emotion construct and the five-item positive emotion were run in a confirmatory factor analysis model to determine whether there were any additional problems with the data. Overall fit for this measurement model was very good, Chi-Square = 35.48, df = 18, P-value = 0.00821, RMSEA = 0.052, NFI = 0.96, CFI = 0.98. The correlation between positive and negative emotions was 0.18 indicating that the two construct are wholly independent of one another which is consistent with previous literature in this area (Watson et al, 1988). Given the acceptability of the fit for this model it was decided that this final solution would represent the emotions construct.

7.3.2 Confirmatory Factor Analysis for Customer Satisfaction and Behavioural Intentions

The initial confirmatory factor analysis of the customer satisfaction items revealed a very poorly fitting model, Chi-Square = 53.09, d f= 2, P-value = 0.00000, RMSEA = 0.269, NFI = 0.90, CFI = 0.90. The underlying cause of this was unclear from the LISREL output. Deletion of any one of the four items would lead to a fully saturated, overidentified model therefore further obfuscating which of the satisfaction items were causing problems. It was decided to leave the satisfaction items temporarily and to examine the behavioural intentions items which would then be run in a confirmatory factor analysis in conjunction with the satisfaction items. A similar problem arose with the behavioural intentions items with the four items providing a fit of Chi-Square = 8.03, df = 2, P-value = 0.01801, RMSEA = 0.092, NFI = 0.99, CFI = 0.99. While the fit of this model is not excessively poor, in order to determine whether any modifications needed to be made the behavioural intentions construct was modelled alongside the satisfaction

items. This model produced a poor fit, Chi-Square = 147.03, df = 19, P-value = 0.00000, RMSEA = 0.138, NFI = 0.91, CFI = 0.92, however modification indices suggested that three items were lacking discriminant validity and wanted to cross-load onto the other dimension. These items were, 'My choice to go to this service station was a wise one', 'I think I did the right thing when I decided to go to this service station' and 'I would say positive things about this service station'. Deletion of these items revealed a satisfactory fit of Chi-Square = 10.83, df = 4, P-value = 0.02858, RMSEA = 0.069, NFI = 0.99, CFI = 0.99. The details of this model can be found in Table 7.7 below.

Table 7.7: Confirmatory Factor Analysis for Customer Satisfaction and Behavioural Intentions

	Estimates	**Standardised Loading**	**t-value**	**R^2**
Customer Satisfaction				
I am satisfied with my decision to go to this service station	1.09	0.99	7.32*	0.97
If I had to do it all over again, I would feel differently about this service station**	0.75	0.47	5.91*	0.22
Behavioural Intentions				
I would recommend this service station to someone else	1.56	0.96	23.30*	0.92
I would encourage friends and family to go to this service station	1.61	0.96	23.04*	0.91
I would consider this service station as my first choice in the future	1.09	0.58	11.70*	0.33
Model fit: Chi-Square = 10.83, df = 4, P-value = 0.02858, RMSEA = 0.069 Intercorrelation (φ) = 0.37 * = significant at the p = 0.01 level **= reverse scored				

The items that remain in the factor solutions while reduced somewhat in terms of the number of items, fortunately the content validity of the constructs has not been jeopardised. According to Oliver (1997, p. 344), 'any item but the satisfaction anchor can be considered for deletion', as the satisfaction anchor remains ('I am satisfied with

my decision to go to this service station'), the two item satisfaction construct sufficiently captures this important construct. With regards to behavioural intentions, although one word-of-mouth communication item was deleted ('I would say positive things about this service station') two word-of-mouth behaviours remained therefore adequately capturing this aspect of loyalty intentions. Most importantly, the behavioural loyalty item remained 'I would consider this service station as my first choice in the future' which captures the economic aspect of consumers' behavioural intentions.

7.3.3 Assessing the Discriminant Validity of the Structural Model

Discriminant validity refers to testing statistically whether two constructs differ. In previous chapters this has been dealt with as part of assessing the measurement model. As new constructs have been introduced in this chapter, discriminant validity will be assessed for the entire model, including both the servicescape and the service quality constructs from previous chapters. Using Bagozzi and Yi's (1991) recommendations, the Phi (φ) values between all constructs were examined to determine whether they were significantly less than 1.00. All constructs had relatively low inter-correlations apart from the Design and Behavioural Intentions constructs which had a Phi (φ) of 0.82 (0.12) therefore violating Bagozzi and Yi's (1991) cut-off point for discriminant validity. Unusually the standardised correlation between the two constructs was 0.48, a relatively low correlation that would normally not lead to discriminatory problems as it is considerably below the 0.85 correlation cut off used by many researchers. Given this finding, further tests of discriminant validity were conducted. Fornell and Larcker's (Fornell and Larcker, 1981) more stringent test of discriminant validity was applied. This

formula takes the Average Variance Extracted (AVE) for the construct and compares it to the square of the construct inter-correlations. If the AVE is greater than the inter-correlation then discriminant validity holds. Using the AVE of .72 for behavioural intentions from Table 7.9 it is found that this is greater than the squared inter-correlation (see Table 7.10) of these two variables ($.21^2 = .044$) therefore finding that discriminant validity between these two constructs exists.

7.4 Analysis of the Full Structural Research Model

In addition to the assessment of the measures using the above tests, further discriminant validity checks were conducted using Netemeyer et al's (2003) guidelines. This involved running a full measurement model of all constructs in the model to assess whether violations of discriminant validity existed. Through inspection of the phi matrix it was found that no violations were made (see Appendix 7A) and analysis of the full model proceeded. Common methods variance was also tested using Harman's single factor test (Podsakoff et al, 2003) and an exploratory factor analysis of all variables produced a seven factor solution thus confirming that common methods variance was not at play here. Results from this can be viewed in Appendix 7B.

Since all measurement of all constructs have been sufficiently assessed using the guidelines recommended in the extant literature the next stage involves testing the various hypotheses outlined in the previous chapter. All hypotheses culminate to form the research model depicted in Figure 7.1 below:

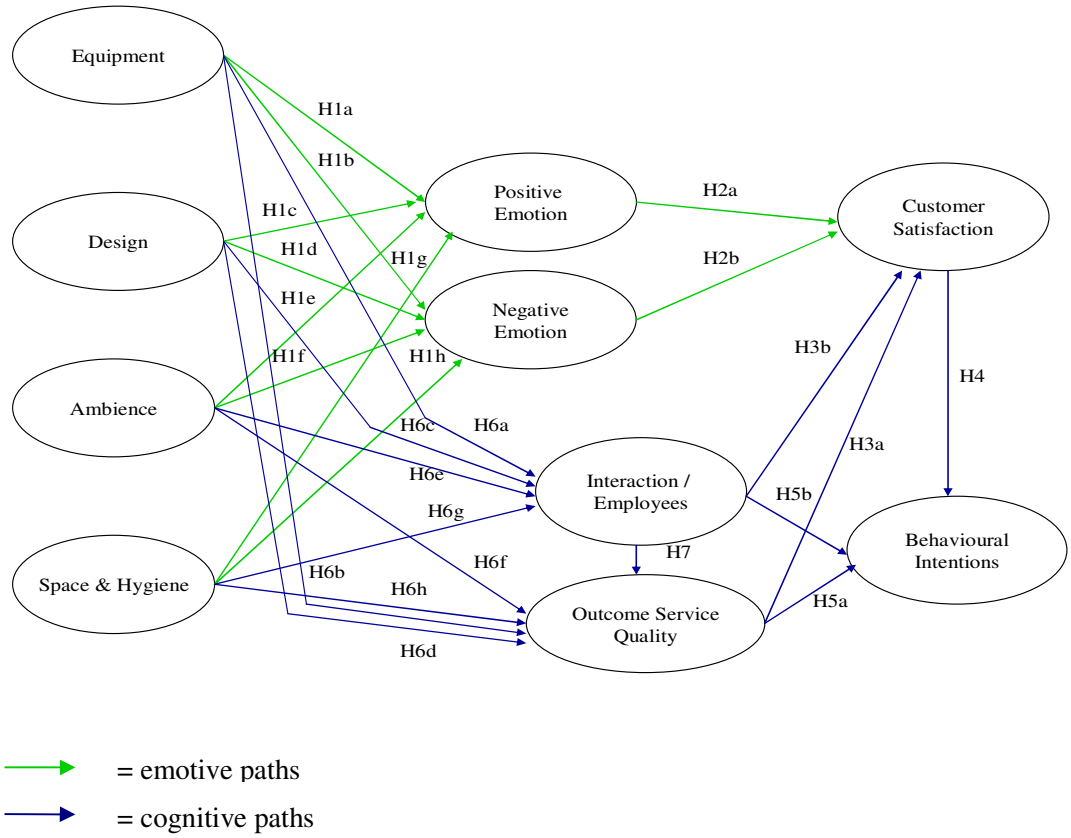


Figure 7.1: Proposed Research Model

While it is not conventional to examine structural models in stages, it was decided to examine the paths leading to and from the emotions construct prior to running the entire model. The reasons for this are twofold, firstly, at the measurement model stage described above it was found that many of the emotions items were quite unstable. The cause of this is not entirely clear and it is hoped that by scrutinising this construct in closer detail it may reveal further anomalies in the data. Secondly, the sheer complexity of the above model means that by analysing the subsets in isolation prior to their amalgamation makes both analysis and interpretation more manageable.

7.4.1 Testing Hypotheses H1 and H2

These hypotheses proposed that emotion would moderate the relationship between servicescape evaluations and overall customer satisfaction. The formal hypotheses are reiterated here:

H1a: Consumers' perceptions of Servicescape Design will have a positive effect on positive emotions.

H1b: Consumers' perceptions of Servicescape Design will have a negative effect on negative emotions.

H1c: Consumers' perceptions of Servicescape Ambience will have a positive effect on positive emotions.

H1d: Consumers' perceptions of Servicescape Ambience will have a negative effect on negative emotions.

H1e: Consumers' perceptions of Servicescape Equipment will have a positive effect on positive emotions.

H1f: Consumers' perceptions of Servicescape Equipment will have a negative effect on negative emotions.

H1g: Consumers' perceptions of Servicescape Hygiene & Space will have a positive effect on positive emotions.

H1h: Consumers' perceptions of Servicescape Hygiene & Space will have a negative effect on negative emotions.

Structural Model

The global fit of the model was good (Chi-Square = 508.88, df = 259, P-value = 0.00000, RMSEA = 0.052, NFI = 0.84, CFI = 0.91) indicating good fit of the model (Hu and Bentler, 1999). While the overall fit for the model is satisfactory the results cannot wholly confirm that the servicescape causes emotional responses and that these emotional responses in turn effect customer satisfaction evaluations. In Figure 7.2 overleaf it is clear that many of the hypothesised paths were found to have insignificant t-values therefore leading to the conclusion that the proposed hypothesised links must be rejected. The paths from the Design construct to both the emotions constructs were found to be significant, however the direction of the path between Design and Negative Emotion was not in the direction expected. Both Equipment and Ambience were found to have no effect on both Positive and Negative Emotions, thus disconfirming the ability of these aspects of the servicescape to elicit emotional responses. Mixed findings were found for the Space and Hygiene dimension with it having no effect on Positive Emotions, however it was found to have a negative effect on Negative Emotions, which is in the direction expected.

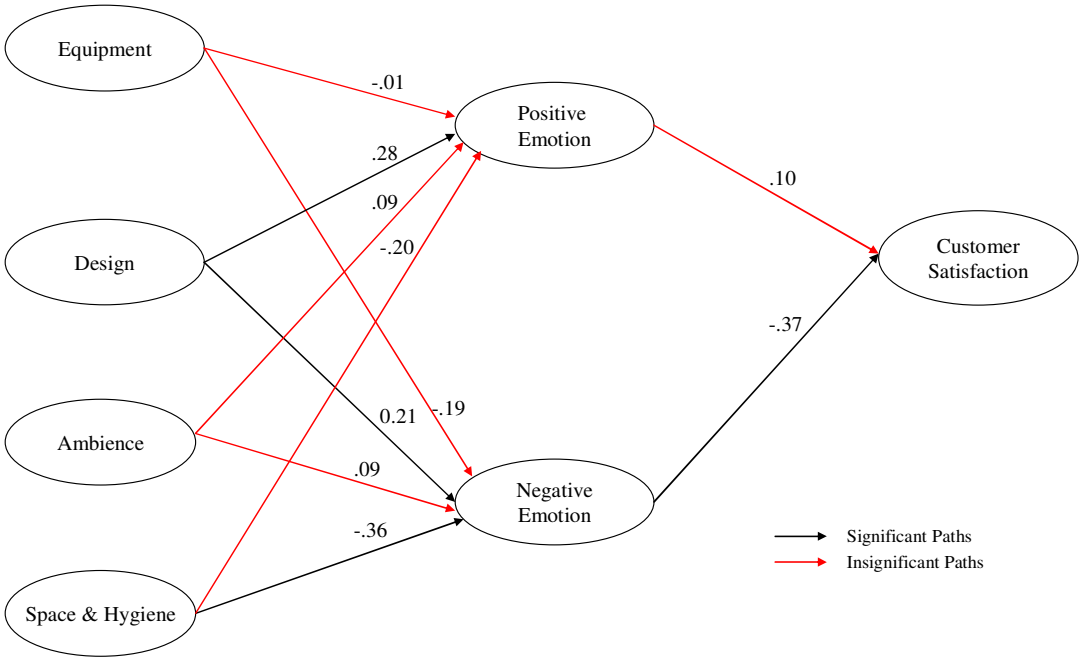


Figure 7.2: Results of Model Estimation

Mixed results were also found for Hypotheses 2a and 2b with positive emotions contributing only marginally and insignificantly to the variance in customer satisfaction. Interestingly, it was found that the presence of negative emotions have a significant negative effect on customer satisfaction.

Further Evidence for the Insignificant Role that Emotions Play in Utilitarian Service Encounters

The above results cast doubt over the ability of servicescapes to elicit emotional responses. Similarly, for this service, emotions were found to account for only a small percentage of the variance in customer satisfaction. These findings question the preconceived notion that emotions play an important role in the consumer experience. In

Section 7.3.1 above it was found that many of the items used in Watson et al's (1988) PANAS were highly unstable and were thus removed from further analysis.

Further inspection of all twenty of the PANAS items provides additional evidence that suggests the role of emotions in this service context is nearly entirely irrelevant. Basic descriptive statistics such as the mean and the standard deviation (See Table 7.8 below) of the items shows that the responses were highly skewed towards the lower end of the scale which represents that these respondents reported that they did not experience these emotions while in the service station (1 = not at all, 2 = a little, 3 = moderately, 4 = quite a bit, 5 = a lot).

Table 7.8: Descriptive Statistics for Emotions Items

	Mean	Standard Deviation
Interested	2.17	1.104
Nervous	1.03	.242
Excited	1.22	.579
Upset	1.06	.397
Strong	1.45	.896
Irritable	1.14	.525
Alert	2.44	1.253
Inspired	1.30	.726
Jittery	1.08	.348
Active	1.96	1.124
Distressed	1.06	.397
Guilty	1.04	.312
Scared	1.03	.218
Hostile	1.02	.167
Enthusiastic	1.75	1.007
Proud	1.41	.876
Ashamed	1.03	.190
Determined	1.57	.993
Attentive	1.86	1.103
Afraid	1.03	.254

These findings add further credence to the structural model presented in Figure 7.2 therefore consolidating the notion that emotions are experienced at only a negligible level for this particular service.

Model Discussion

The servicescape literature strongly advocates that environments elicit emotional responses in individuals (Mehrabian and Russell, 1974; Bitner, 1990; Wakefield and Blodgett; 1999). Although previous servicescape studies have demonstrated this relationship in service contexts (Sherman et al, 1997; Wakefield and Blodgett, 1999) for the current context this relationship could not be shown. The likely cause of this is the utilitarian nature of the service. Customers of this service station firstly do not spend a significant amount of time in the store and therefore are less likely to have experienced a range of emotions. Secondly, the purposes for which respondents attend this store are likely to be laden with utilitarian needs. For example, individuals frequent service stations when they need to purchase petrol for their car, to buy a newspaper or to get a take away coffee if they are on the road. These purchases are evidently more cognitively primed therefore leaving little room for emotional responses. In fact, it is entirely likely that the respondents do not attend service stations because of the experiential aspects of the service offering but instead are drawn to the service either due to convenience or price.

In the customer satisfaction literature it has become generally accepted that satisfaction has both a cognitive and an affective content (Westbrook and Oliver, 1991; Mano and

Oliver, 1993; Oliver, 1997; Oliver et al, 1997). The findings here only partly support this with only negative emotions having a negative effect on customer satisfaction. The four negative emotions which were used in the structural model (upset, irritable, distressed, scared) if experienced by customers will have a negative effect on satisfaction that is over and above the influence of positive emotions. This finding is in line with Liljander and Strandvik's (1997) study where it was found that positive emotions contributed insignificantly to the satisfaction response. Again it is likely that this finding is linked to the service under scrutiny here with individuals arriving at a level of customer satisfaction based on their *a priori* expectations. For utilitarian services, consumers are likely to have utilitarian expectations, therefore the evaluative process will not involve strong emotional responses.

This section has demonstrated that emotions are only tenuously linked to both servicescape and customer satisfaction evaluations. This is an important contribution as up until this point, the literature has wholly supported that emotional responses are a key moderator of the relationship between environmental stimuli and cognitive post-consumption processes such as satisfaction. The findings in this section also question whether emotional responses should be included at the next stage of the analysis when the full model is tested. Given that only limited support could be found for the role of emotion in this service context it seems unnecessary to proceed with further analysis in which the emotions construct is included. Therefore, emotions are excluded from all subsequent tests and the remainder of the hypotheses will be tested (Hypothesis 3 through to 7) as part of one structural model

7.4.2 Testing Hypotheses H3 to H7

In this section Hypotheses 3 through to 7 will be analysed. This final stage of analysis focuses solely on the cognitive paths in the model and tests whether the hypothesised relationships exist. As the discriminant validity of the constructs has previously been tested the analysis can advance directly to analysing the full structural model. The hypothesised paths are shown in Figure 7.3 below.

As with previous models, the testing of the structural model was done through the use of structural equation modelling with LISREL 8.51 and Maximum Likelihood Estimation. All hypothesised parameters were entered into the LISREL program and the model was run. Overall fit statistics for the structural model were very good (Chi-Square = 548.41, df = 280, P-value = 0.00000, RMSEA = 0.052, SRMR = 0.066 CFI = .93, χ^2/df = 1.96) with Hu and Bentler's (1999) two combinational rule of a RMSEA less than .06 and a SRMR less than .09 found here. Examination of the Q-Plot of Standardized Residuals also corroborates acceptable fit with the residuals lying approximately along the diagonal (see Appendix 7C).

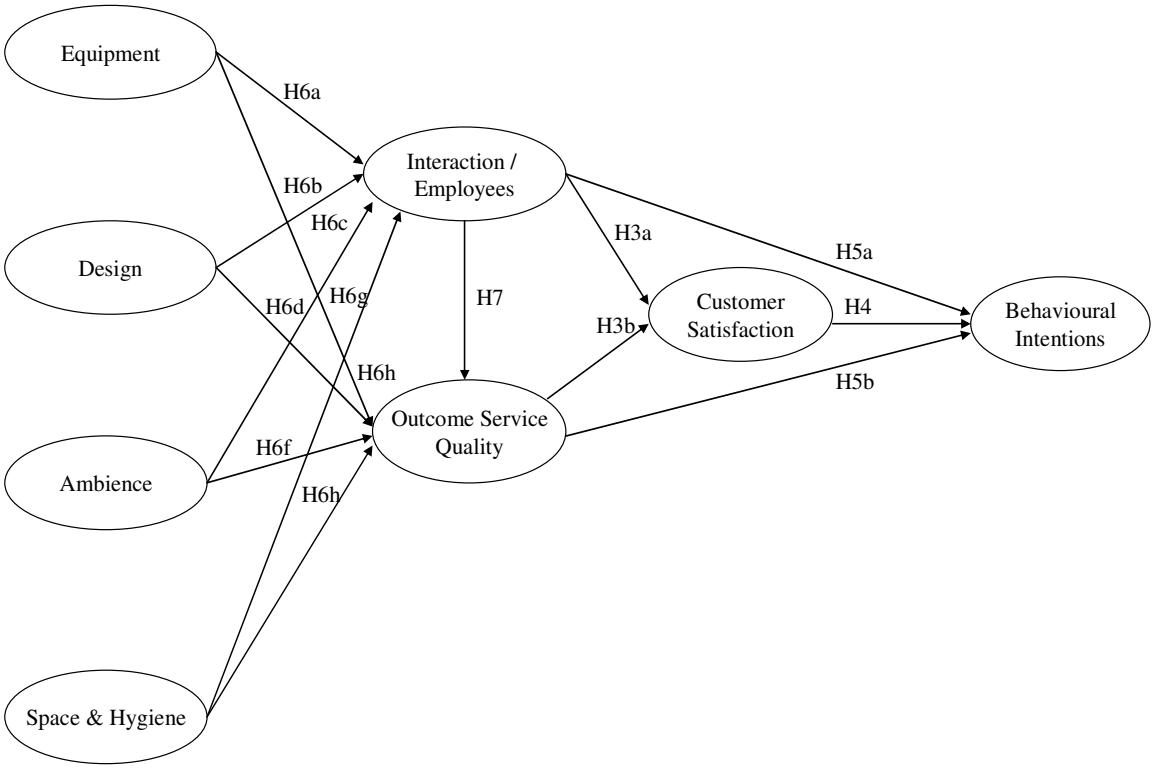


Figure 7.3: Hypothesised Cognitive Model

Convergent validity of the measures was assessed by calculating the average variance extracted, composite reliabilities and alpha coefficients. Further to this, the path coefficients (loadings) for each latent factor on their manifest indicators were also examined. All composite reliabilities were satisfactory and above the .60 level recommended by Bagozzi and Yi (1991). Average variance extracted statistics were also all above the standard accepted cut off of .50 (Fornell and Larcker, 1981) apart from the Space and Hygiene and Equipment constructs which were just below the cut off level. All results from these statistics can be viewed in Table 7.9 overleaf. Analysis of the factor loadings indicated that all items on both exogenous and endogenous constructs

were significant at the 0.05 level. Full details of all these loadings are provided in Table 7.9 below. Construct Inter-correlations are also provided in Table 7.10.

Table 7.9: Measures of Internal and Convergent Reliability

Construct	**Cronbach's ά**	**Average Variance Extracted (AVE)**	**Composite Reliability (CR)**
Ambience	.76	.68	.80
Design	.78	.50	.80
Equipment	.69	.43	.69
Hygiene and Space	.75	.44	.80
Interaction/Employee Service Quality	.79	.57	.84
Outcome Service Quality	.79	.57	.80
Customer Satisfaction	.60	.55	.69
Behavioural Intentions	.85	.72	.88

Table 7.10: Construct Inter-Correlations

	ESQ	**OSQ**	**CS**	**BI**	**Ambience**	**H&S**	**Design**	**Equip**
ESQ	1.00							
OSQ	.71	1.00						
CS	.49	.60	1.00					
BI	.28	.48	.39	1.00				
Ambience	.35	.29	.19	.12	1.00			
H&S	.82	.65	.44	.27	.46	1.00		
Design	.42	.45	.28	.21	.48	.66	1.00	
Equip	.53	.65	.40	.30	.38	.72	.75	1.00

ESQ = Employee Service Quality, OSQ = Outcome Service Quality, CS = Customer Satisfaction, BI = Behavioural Intentions, H&S = Hygiene & Space, Equip = Equipment.

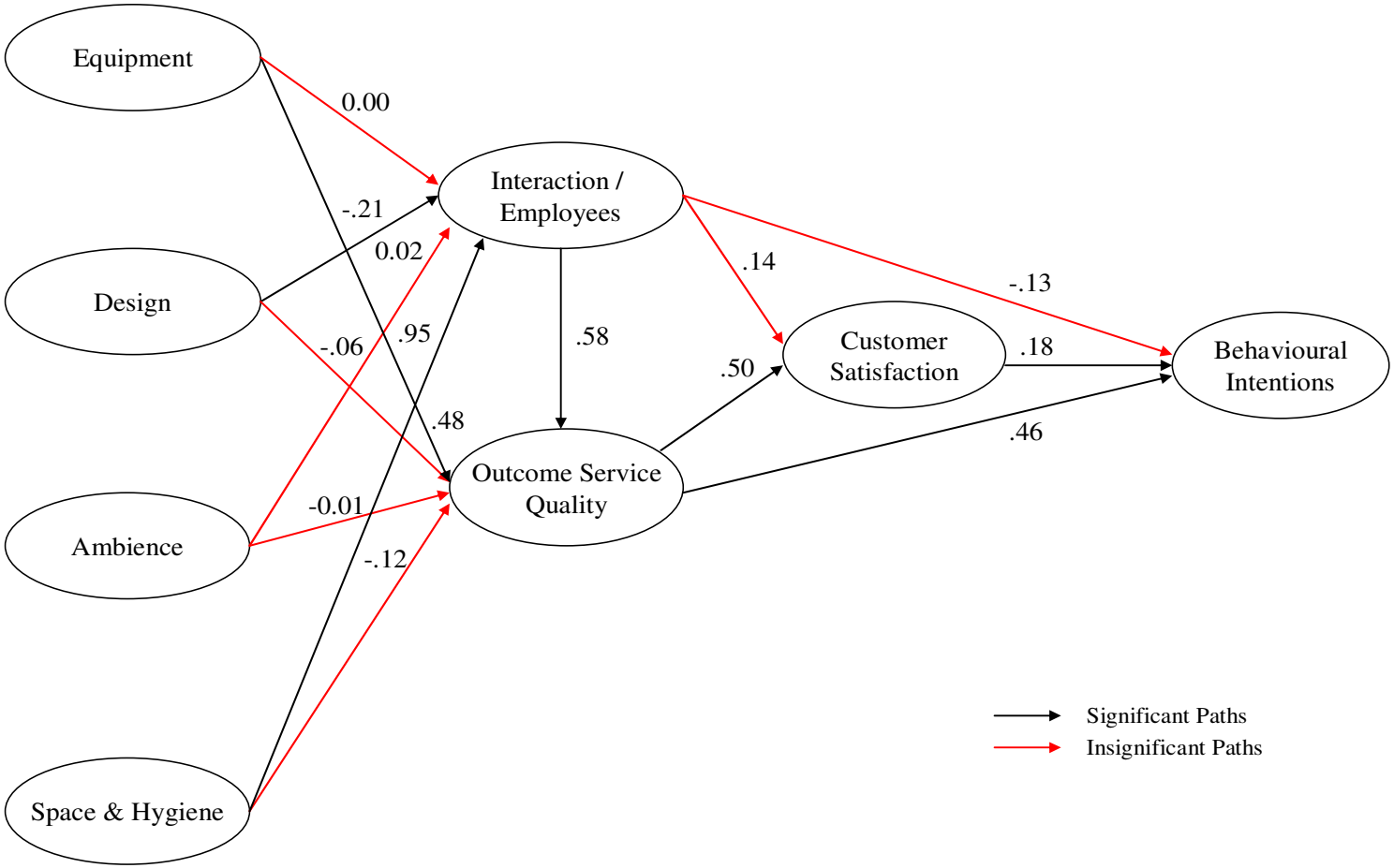


Figure 7.4: Research Model

Table 7.11: Factor Loadings, *t* Values, Means and Standard Deviations for Measurement Items

Construct	Factor Loading	*t* Value	Mean	Standard Deviation
Ambience				
The background music was pleasant	.63	7.26 [b]	4.58	1.25
The background music was appropriate	.98	---- [a]	4.81	1.24
Design				
I found the interior design visually appealing	.81	---- [a]	4.64	1.36
The interior design was attractive	.81	15.16 [b]	4.73	1.36
The colour schemes were pleasant	.59	10.87 [b]	4.98	1.31
The architecture was attractive	.58	10.69 [b]	4.40	1.55
Equipment				
The equipment was modern looking	.66	8.65 [b]	5.15	1.17
The electronic equipment was excellent	.55	---- [a]	5.16	1.23
The equipment was of high quality	.75	9.17 [b]	5.32	1.28
Hygiene and Space				
The store was very clean	.61	---- [a]	5.83	.92
The service station appeared to be hygienic	.69	10.23 [b]	6.08	.98
The flooring was appropriate	.60	9.26 [b]	5.44	1.20
I found my way around quite easily	.65	9.84 [b]	5.99	.942
The employees were neat and tidy in appearance	.55	8.55 [b]	6.09	.83
Interaction/Employees				
The staff were polite	.83	13.02 [b]	6.49	.63
The staff were helpful	.68	---- [a]	6.29	.96
The staff were trustworthy	.73	11.82 [b]	6.15	1.06
The service in store was delivered promptly	.65	10.80 [b]	6.36	.82
Outcome Service Quality				
The service here suited my needs	.66	12.50 [b]	6.16	1.06
The service here is reliable	.83	---- [a]	6.33	.893
The service station provided quality service	.76	14.74 [b]	6.19	1.04
Customer Satisfaction				
I am satisfied with my decision to go to this service station	.94	---- [a]	6.08	1.10
If I had to do it all over again, I would feel differently about this service station	.49	5.99 [b]	5.99	1.6
Behavioural Intentions				
I would recommend this service station to someone else	.96	---- [a]	5.26	1.62
I would encourage friends and family to go to this service station	.95	29.26 [b]	5.12	1.68
I would consider this service station as my first choice in the future	.58	12.55 [b]	4.95	1.89

[a] = The corresponding coefficient was fixed to set the metric of the latent construct.
[b] = Significant at the $p < 0.05$ level

Figure 7.4 showing the structural equation modelling results for the hypotheses testing finds that some of the paths are not supported. Each of these paths now be discussed in sequence. The first hypotheses were numbers 6a and 6b and they proposed that the Equipment construct would have a positive effect on both elements of service quality. This was only shown between the Equipment and Outcome Service Quality dimension. Hypotheses 6c and 6d between Design and Service Quality only found support for a significant path between Design and Employee/Interaction Service Quality with this effect being in the negative direction. As was found in the Service Quality chapter, no evidence of Ambience having a significant effect on either service quality dimensions was found. Hypotheses 6g and 6h which linked the Space and Hygiene construct to the two service quality dimensions was also found to predict only one of the service quality dimensions, this being the Employee dimension. Employee Service Quality was found to have a strong positive effect on Outcome Service Quality (H7). Of the two service quality dimensions, only the Outcome Service Quality dimension was found to have a significant effect on overall Customer Satisfaction (H3a and H3b). As was expected, Customer Satisfaction also had a direct effect on Behavioural Intentions (H4). Finally, H5a and H5b which hypothesised paths that led directly from both service quality constructs to Behavioural Intentions was found to be only supported for the Outcome Service Quality construct.

Modification indices suggested by the program were very few. However, it was interesting to find that the largest modification index was the addition of a direct path

between Design and Behavioural Intentions. The model was therefore re-run including only those hypotheses supported above so as to provide a clearer picture of the significant paths. Model fit is hugely improved to Chi-Square = 468.43, df = 241, P-value = 0.00000, RMSEA = 0.051, SRMR = 0.054, CFI = .94, χ^2/df = 1.94, however it is not possible to calculate a Chi-Square difference test as the models are not nested (the Ambience dimension was wholly removed as both paths leading from it were insignificant). The second advantage of removing these insignificant paths is that it allows a more parsimonious model to be tested (Diamamtopoulos and Siguaw, 2006). Indeed the PNFI in the more parsimonious model increases .02 points to 0.77. The removal of these paths has some minor changes to the values of the structural paths, however it did not effect the significance of the paths in any way nor were there any large changes in the standardized values. By removing the insignificant paths it also allows the total effects of all predictor variables on the final dependent variable of Behavioural Intentions. These total effects can be seen in Table 7.13 overleaf. The parameter estimates of this final model can be found in Table 7.12 below.

Table 7.12: Standardized Gamma and Beta Parameter Estimates

Path	**Standardized Path Coefficient**	***t* Value**
Hygiene → Employee/Interaction	.93	7.76*
Design → Employee/Interaction	-.19	-.2.34*
Design → Behavioural Intentions	.33	5.42*
Equipment → Service	.33	4.57*
Employee/Interaction → Outcome Service	.53	7.33*
Outcome Service → Customer Satisfaction	.61	10.38*
Customer Satisfaction → Behavioural Intentions	.16	2.27*

* = Significant at the $p < 0.05$ level

Table 7.13: Total Effects and Indirect Effects

DIRECT EFFECTS						
	Hygiene	**Design**	**Equipment**	**Employee/ Interaction Service Quality**	**Outcome Service Quality**	**Customer Satisfaction**
Behavioural Intentions	-----[1]	.33	-----[1]	-----[1]	.20	.16
INDIRECT EFFECTS						
	Hygiene	**Design**	**Equipment Employee/**	**Interaction Service Quality**	**Outcome Service Quality**	**Customer Satisfaction**
Employee	-----[2]					
Service	0.49	-.10*	-----[2]	-----[2]		
Satisfaction	0.30	-.06*	.20	.32	-----[2]	
Behavioural Intentions	0.15	-.03*	.10	.16	.10	-----[2]
TOTAL EFFECTS						
	Hygiene	**Design**	**Equipment**	**Employee/ Interaction Service Quality**	**Outcome Service Quality**	**Customer Satisfaction**
Behavioural Intentions	0.15	0.31	0.10	0.16	0.30	0.16

* = not significant, -----[1] = no direct effects were specified between these constructs, -----[2] = direct effects were specified between these constructs

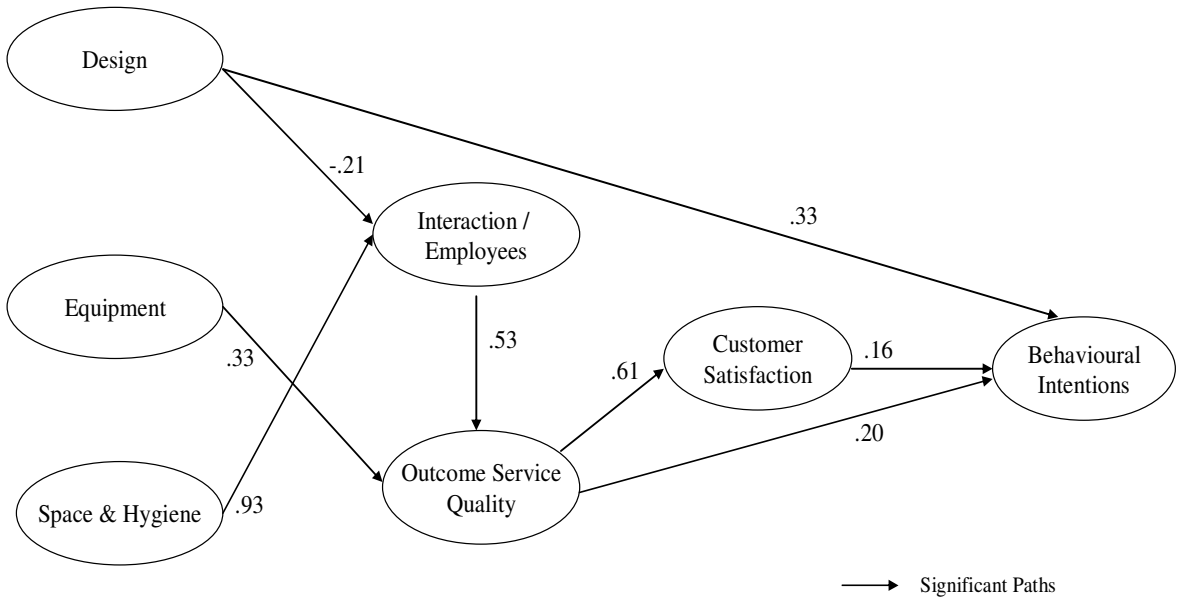


Figure 7.5: Final Structural Model

Model Implications

In this section the implications of all findings from the structural model are discussed. The discussion will begin by discussing the significant paths and their theoretical implications which will be followed by a discussion on the presence of the insignificant paths.

The Space and Hygiene construct had a strong and significant effect on respondents' perceptions of the interaction with employees. Why this effect was significant and not its effect on Outcome Service Quality seems at first difficult to explain. However, it is a possibility that respondents attributed the cleanliness and upkeep of the store environment directly to the employees, therefore resulting in this direct effect. For the Design construct, initial examination of its relationship to the Employee construct shows that it has a negative effect. One can only surmise as to why this is the case; one possible reason for this effect is that in service stations in Ireland, invariably the only time that respondents are in contact with employees is when they are paying for their petrol or goods. In the service stations surveyed, the design of the stores is such that the employees are located behind a high desk and shatterproof glass. It is likely that the conditions inherent in this particular design have a negative effect on respondents as they may find it difficult to hear the employee or they believe that the relatively cramped conditions in which the employees work is unsatisfactory. However, despite this negative effect on respondents' perceptions of employees it was found that the design of the store has a direct and positive effect on behavioural intentions. The cause of this maybe because although there were negatives inherent in the design of the store, on the

whole, respondents viewed the design favourably, so much so that it would encourage them to engage in positive behavioural activities. What is also interesting is that the design of the store contributed more of the variance in behavioural intentions than satisfaction. This is of significant importance to service stations as it demonstrates that the design of the store is more salient than the actual service quality that is received by customers.

The cognitive paths are also supported in the model with the direct and positive effect of Employee perceptions on Outcome Service Quality, with Outcome Service Quality being a strong determinant of customer satisfaction and finally customer satisfaction influencing whether respondents intend to return or engage in positive word-of-mouth activities. It is found that customers seem to engage in a set of very specific cognitive stages whereby they evaluate the level of service received from staff which leads them to arrive at a perceived level of Outcome Service Quality. This then had a direct effect on customer satisfaction which mediated the relationship between Outcome Service Quality and Customer Satisfaction.

It was also found that a number of the paths were insignificant. The most notable of these unsupported paths was the inability of the ambience dimension to impact the evaluative constructs. As ambient cues are more likely to affect the subconscious there is a strong possibility that the customers of the service stations paid more attention to the physical aspects of the servicescape than the musical cues therefore leading to their rather paltry effect on evaluative constructs. Equipment having no effect on perceptions of the

employee service is an interesting finding as it is highly likely to be attributable to the very service under study. The majority of the equipment seen by customers in service stations is that which involves self-service, examples of which are the petrol pumps, the coffee dispensing machines and the in-store ATMs. The tills operated by the staff are a form of equipment that is not used by customers, however, it is likely that the tills have no influence on equipment perceptions given the relatively high positioning of the sales assistant, with customers only getting a view of the back of the till.

It was also found that the Space and Hygiene dimension does not directly affect Outcome Service Quality, however there were indirect effects through the Employee/Interaction Service Quality dimension (indirect and total effects are shown in Table 7.12). This finding signifies that if perceptions of the cleanliness or tidiness of both the physical environment and the staff is deemed to be poor by potential customers then it will have a knock-on effect on Outcome Service Quality perceptions.

It is also promising to find that both the total and indirect effects demonstrate the positive effects on behavioural intentions (See Table 7.12). Although consumers' self-reported purchase intentions are not perfectly correlated with actual purchase behaviour, nor is this relationship very clear, however what remains is that behavioural intentions remain an important metric for firms who do not have access to loyalty behaviours. This is particularly meaningful given that behavioural intentions have been linked to profits, market share and customer lifetime value (Anderson et al, 1994; Jones and Sasser, 1995; Szymanski and Henard, 2001; Gupta and Zeithaml, 2006; Lariviere, 2008).

Table 7.14: Results of Hypothesis Testing

Hypothesis	**Rejected/ Accepted**
H1a: Consumers' perceptions of Servicescape Design will have a positive effect on positive emotions.	Accepted
H1b: Consumers' perceptions of Servicescape Design will have a negative effect on negative emotions.	Rejected
H1c: Consumers' perceptions of Servicescape Ambience will have a positive effect on positive emotions.	Rejected
H1d: Consumers' perceptions of Servicescape Ambience will have a negative effect on negative emotions	Rejected
H1e: Consumers' perceptions of Servicescape Equipment will have a positive effect on positive emotions.	Rejected
H1f: Consumers' perceptions of Servicescape Equipment will have a negative effect on negative emotions.	Rejected
H1g: Consumers' perceptions of Servicescape Hygiene & Space will have a positive effect on positive emotions.	Rejected
H1h: Consumers' perceptions of Servicescape Hygiene & Space will have a negative effect on negative emotions.	Accepted
H2a: Positive emotional responses will have a positive influence on satisfaction.	Rejected
H2b: Negative emotional reactions will have a negative influence on satisfaction.	Accepted
H3a: Outcome service quality evaluations will have a direct effect on customer satisfaction.	Accepted
H3b: Employee service quality evaluations will have a direct effect on customer satisfaction.	Rejected
H4: Customer satisfaction will have a positive, direct effect on behavioural intentions.	Accepted
H5a: Outcome service quality will have a positive, direct effect on behavioural intentions.	Accepted
H5b: Employee service quality will have a positive, direct effect on behavioural intentions	Rejected
H6a: Consumers perceptions of Servicescape Equipment will have a positive effect on perceptions of Employees	Rejected
H6b: Consumers perceptions of Servicescape Equipment will have a positive effect on perceptions of Outcome Service Quality	Accepted
H6c: Consumers perceptions of Servicescape Design will have a positive effect on perceptions of Employees	Accepted
H6d: Consumers perceptions of Servicescape Design will have a positive effect on perceptions of Outcome Service Quality	Rejected
H6e: Consumers perceptions of Servicescape Ambience will have a positive effect on perceptions of Employees	Rejected
H6f: Consumers perceptions of Servicescape Ambience will have a positive effect on perceptions of Outcome Service Quality	Rejected
H6g: Consumers perceptions of Servicescape Space & Hygiene will have a positive effect on perceptions of Employees	Accepted

Hypothesis	Rejected/ Accepted
H6h: Consumers perceptions of Servicescape Space & Hygiene will have a positive effect on perceptions of Outcome Service Quality	Rejected
H7: Employee/Interaction Service Quality will have a positive and direct effect on Outcome Service Quality	Accepted

Model Discussion

The first major finding this chapter demonstrates is that the multi-dimensional construct developed in earlier chapters has provided nomological validity in that each of the constructs (apart from the Ambience dimension) has been shown to predict one, if not more of the outcome constructs. While the final three factor structure of the servicescape (Design, Equipment and Space & Hygiene) is conceptually different to previous authors such as Bitner (1992) and Baker's (1987) frameworks it adds credence to the notion that store environments can be broken into discrete dimensions. Conceptually, there are two major differences between the servicescape constructs included in the final model presented in Figure 7.4 and Bitner (1992) and Baker's (1987) work. Firstly, no evidence could be found for the inclusion of any ambience cues here. This contrasts starkly with experimental studies where music has been demonstrated to have a direct effect on behavioural responses (Milliman, 1982; Areni and Kim, 1993; Dube et al, 1995; Chebat et al, 2001). Given that this study was conducted in the field with actual customers wherein no manipulations of the ambience cues were possible, the ability to detect the effect of ambient cues radically diminishes. Experimental studies offer the advantage of allowing the researcher to vary the stimuli under study and then rather than directly measuring outcome variables through self-reports, observational techniques can be used to determine whether significant changes in behaviour occurred. The disadvantage of this however, is that it is very difficult to effectively examine more than four stimuli

effectively as interpretation becomes unwieldy, therefore rendering this technique unworkable when a more comprehensive view of stimuli is required. Unfortunately non-experimental studies on the servicescape tend to have ambience cues embedded in constructs alongside non-ambience dimensions such as items relating to hygiene (Wakefield and Blodgett, 1999) therefore making it difficult to determine whether the effect was purely ambience based. Interestingly, a recent study by Harris and Ezeh (2008) in which multiple regression was used, found that both aroma and music produced standardised *beta* coefficients of less than .07 a finding which almost entirely agrees with the propositions put forward here. From this it seems that one can conclude that ambient cues are only a tangential part of the utilitarian servicescape design. However this is a presumption that cannot be assumed as it is unknown how respondents would react if music was missing from the servicescape design. Given how the plethora of studies on music has shown how music has the ability to tap into behavioural responses it is therefore a strong possibility that music influences our behaviour on a very abstract, subconscious level.

This model has also demonstrated that tangible cues lie outside the realm of service quality and serve as an antecedent, rather than an element of the construct. Notable authors on service quality such as Parasuraman et al (1985; 1988; 1994) as well as many others (Cronin et al, 1997; Brady and Cronin, 2001) have published studies in this area in which tangibles are subsumed under service quality. The difference between these studies and the current project is that the majority of these studies have not sought to delineate between the servicescape and service quality. One such study where somewhat

similar findings to this have been found is that of Reimer and Kuehn (2005). Using path analysis, it was demonstrated how an aggregated measure of the servicescape has an effect on intangible service quality cues. The model depicted in Figure 7.4 adds to the work of Reimer and Kuehn's (2005) by modelling the servicescape as a multidimensional construct and therefore allowing for the relationships between individual servicescape sub-dimensions to be related to service quality which allows one to discern if certain dimensions are more influential. The model tested here also authenticates that Baker's (1987) conceptualisation is theoretically unclear. While the service quality literature has generally included servicescape elements, Baker (1987) has advocated the antithesis of this where service quality-type items are included in a store environment model alongside predominantly tangible or servicescape items.

It is difficult to make direct comparisons between this study and the few who have also taken a dimensional approach to measuring the servicescape, as no authors have employed the same dimensions as those utilised here. However, Wakefield and Blodgett (1996) used structural equation modelling to model how a number of servicescape constructs relate to overall servicescape quality and in this study it was found that aesthetics were the primary determinant of perceived service quality. While differences exist between this study and theirs it is important to note that design, a construct which is in terms of content almost identical to Wakefield and Blodgett's (1996) aesthetics items, over and above *all* other constructs was the most significant antecedent to behavioural intentions. A very similar finding was found by Wakefield and Blodgett (1996) where their Aesthetics dimension contributed a substantially greater amount of variance than the

remaining four servicescape constructs (layout accessibility, seating comfort, electronic displays, cleanliness). It was also found by Wakefield and Blodgett (1996) that of the servicescape constructs studied, cleanliness contributed the least amount of variance. For the current study, it was in fact the equipment construct that contributed the least variance out of the three servicescape dimensions. The reason for this difference between the two studies probably lies in the difference of service and elucidates how modelling the servicescape as a multidimensional construct allows comparisons to be made across different service industries.

Despite other studies in the past having examined the direct effects of the environment on behavioural and loyalty intentions (Harris and Ezeh, 2008), the hypotheses set out in this chapter did not specify a direct link between the servicescape and behavioural intentions. Instead it was expected that all servicescape constructs would affect behavioural intentions through service quality and customer satisfaction. All but one of the constructs, namely design, were found to adhere to an indirect route thus demonstrating that while conceptually similar, each of the servicescape sub-dimensions behave in a unique way influencing perceptions and behaviour either directly or indirectly.

Numerous studies have specified relationships between service quality, customer satisfaction and behavioural intentions. Those studies that emanate from the customer satisfaction literature define customer satisfaction as the primary and direct link to outcome measures such as behavioural intentions (Anderson et al, 1994; Bolton and Lemon, 1999). On the other hand, some authors from the service quality literature

propose that service quality has a direct effect on behavioural intentions (Boulding et al, 1993; Taylor and Baker, 1994; Zeithaml et al, 1996). From the results found here both schools of thought are supported as direct effects between both Outcome Service Quality and Behavioural Intentions and between Customer Satisfaction and Behavioural Intentions. Of these two constructs it is Outcome Service Quality that has the greatest effect on behavioural intentions which emphasises that it would be insufficient to substitute a customer satisfaction measure for a service quality measure as certain relationships which are of imperative importance would remain untapped. A final finding was that a distinction was found between the service quality constructs and customer satisfaction. This concurs with previous research in the area (Cronin et al, 1996; Dabholkar et al, 2000) and further highlights the necessity of measuring both constructs in customer surveys.

7.5 Conclusions

The analysis presented in this chapter has made a number of key contributions. Firstly, up until this point no study in the extant literature has attempted to integrate all the constructs included in the proposed research model tested in this chapter. This answers the call for more comprehensive studies in which a greater number of constructs are included which was made in recent years by Gupta and Zeithaml (2006).

The services marketing literature is awash with studies citing the importance of emotional responses as part of the consumption process. These authors hail from studies which have shown how the servicescape is an essential determinant of affect and also from

those studies on customer satisfaction where it has become almost *de rigueur* to find that satisfaction is synonymous with emotion. For this service station context this was not the case. The standard transactional approach of this service leads to a consumption process that is almost entirely absent of any kind of emotional response, be them positive or negative. It is shown here that respondents make their evaluations on purely cognitive processes where their utilitarian needs are satisfied in a utilitarian manner. This finding adds a new dimension to the literature by demonstrating that in not all service contexts do emotions play a role of such magnitude.

A third contribution this chapter makes lies in the multidimensional conceptualisation of the servicescape presented here. While multidimensional approaches have appeared in the literature they have suffered from conceptual and methodological limitations. Here it is demonstrated how the servicescape impacts on a range of post-consumption processes and shows how certain aspects of the servicescape are of greater importance than others. As is demonstrated in the sections above, by using structural equation modelling it was possible to determine which of the servicescape elements are of greater importance. Indeed, it became clear that ambience, in the form of music has only a very minimal and insignificant effect on individuals experience with the service, a finding that could not be detected if servicescape items are aggregated. Further to this, given the flexibility and superiority of structural equation modelling over other methods, it was also possible to test multiple interdependent links between the constructs and through this it was possible to find that the store design had a greater effect on behavioural intentions than customer satisfaction. This demonstrates the potency of the servicescape in influencing how

consumers behave and is an important insight as it shows that the more manageable tangible cues can be manipulated in such a way as to increase repatronage intentions.

8.0 CONCLUSIONS

8.1 Introduction

This final chapter concludes the thesis by firstly examining the contributions it has made to theory. Through this discussion, emphasis will be placed upon how this dissertation has built upon previous work in the area. Secondly, the chapter will outline the limits of the study and from these limitations recommendations will be given for future research. The chapter will conclude with the managerial and practical implications of the study and will advise on how management of the service stations can use this research to improve their service offering.

8.2 Concluding Comments and Contributions to Theory

From the literature reviewed in Chapter One, it became evident that while there is a relatively substantial body of literature on the servicescape, it is severely lacking in a number of respects. Firstly, the work to date on the servicescape has focussed heavily upon service situations that are quite unusual in that individuals would only patronise these services very sporadically (Wakefield and Blodgett, 1994; Wakefield and Blodgett, 1996; Wakefield and Blodgett, 1999; Hightower et al, 2002). Secondly, the purposes for which individuals would use these services would be to gratify a hedonic need therefore classifying these services as hedonic in nature (Hirschman and Holbrook, 1982; Okada, 2005; Chitturi et al, 2008). This research sought to fill this gap by exploring how the servicescape impacts consumer responses in a non-hedonistic, utilitarian service. Previous to this Reimer and Kuehn (2005) have attempted to compare the role of the

servicescape in a utilitarian service (bank) to a hedonic service (restaurant), the current project makes additional advances on this study. In terms of context, this research adds to the body of knowledge by choosing a retail service in which customers make low-involvement purchases. The service stations used in the current study can be classified as a truly utilitarian service, unlike Reimer and Kuehn's (2005) banking context. Whether banks can accurately be described as utilitarian in nature is highly debatable as large amounts of money can change hands and with high levels of privacy and trust required in the relationship.

In the frameworks presented by theorists such as Bitner (1992) and Baker (1987) they are relatively clear in what they regard as the stimuli that should be included under the servicescape umbrella. Applications of the theory are less lucid, with a rather arbitrary approach taken to how the servicescape is conceptualised in empirical research (Wakefield and Blodgett, 1996; Sherman et al, 1997; Wakefield and Baker, 1998). For instance, some have taken to measuring the servicescape as a single construct either using a single indicator in its measurement or using multiple items modelled onto one construct (Wakefield and Blodgett, 1994; Wakefield and Blodgett, 1999; Reimer and Kuehn, 2005). This is clearly incongruent with the theories advocated by Bitner (1992) and Baker (1987) who both propose multidimensional frameworks of the servicescape. Therefore, measuring the servicescape as a single construct is too basic a method as the servicescape is a rich, multifaceted construct that requires a more sophisticated analytical approach. There are a limited number who have recognised this fact and have moved beyond the one dimensional view by conceiving the servicescape as a system of sub-

constructs that individually contribute to overall servicescape evaluations (Wakefield and Blodgett, 1996; Lee et al, 2008). Again these studies suffer from some serious limitations. Firstly, the manner with which the servicescape sub-dimensions are chosen seems to be rather haphazard with no clear rationale given as to why the particular dimensions were chosen. Secondly, each of these studies chose servicescape dimensions that were highly context-specific thus making replications and generalisations very difficult. Thirdly, some of these studies have incorporated dubious dimensions that possess items that do not relate to the tangible cues of the service but instead seem to be more in keeping with intangible service quality cues (Lee et al, 2008). To overcome these limitations, this dissertation integrated the extant literature to date and in doing so hypothesised that the servicescape consists of a six factor structure where each of the dimensions were hypothesised to inter-correlate yet remain distinct from one another. This led the servicescape to be represented as a second-order factor structure, which is a form of structural model that accounts for the relationships among the factors by specifying a single higher order factor (the servicescape) which has a number of lower first order factors (each servicescape sub-dimension) which are each linked to this higher level of abstraction. Previous to this, no author has sought to model the servicescape in this way. Although multidimensional approaches to the servicescape have appeared in the literature, a second order model has not yet been specified. Second-order factor structures are statistically much more rigorous and parsimonious than measurement models therefore applying a much more stringent test of a models' applicability (Rindskopf and Rose, 1988).

To meet these objectives, the servicescape literature was further reviewed to determine whether any scales in the studies previously published would be suitable to meet the instrument needs for this dissertation. From this review, it became obvious that no scale existed in the literature that contained a sufficient number of items so that all hypothesised servicescape constructs could be measured adequately. For instance, the majority of scales to date contain items that are very specific to the service under study thus making them inappropriate for the service station context. Some scales do not suffer from this disadvantage, however, in these context-free scales there were an insufficient number of items with either only one item per construct (Reimer and Kuehn, 2005) or if the scale did contain multiple items per construct omitted key servicescape constructs hypothesised in the current research's model (Wakefield and Blodgett, 1999; Hightower et al, 2002). Given that no scale in the servicescape literature met the requirements of this study, a scale development process was embarked upon. The resultant scale was developed using a student sample and exploratory factor analysis was used as the redundancy technique. The final solution adhered to a remarkably easy to interpret structure and unlike previous servicescape scales, the scale developed here captured multiple facets of the servicescape therefore providing a more rounded view of the concept than previously done. This servicescape battery was thus used in the empirical data collection stage alongside other validated measurement instruments.

This data collection stage met a number of key objectives for this dissertation, however the first of these was to determine whether the servicescape structure initially hypothesised is an accurate reflection of how servicescape sub-dimensions interact. Data

collection yielded a satisfactory sample size and using this data a series of confirmatory factor analyses and second-order factor models were run. An initial one factor model provided very poor fit therefore providing anecdotal evidence that perhaps a multidimensional structure is a more suitable depiction of this complicated construct. A second series of confirmatory factor analyses tested the measurement model and once all measurement issues were dealt with the second-order factor structure was derived. This model was an improvement on the one factor model with regards to fit, however a number of items and constructs were found to be problematic. The most notable of these was the ambience dimension. Severe discriminant and convergent problems were found with three of the items demonstrating particularly poor results ('The lighting was comfortable', 'The store had a pleasant smell' and 'The atmosphere was comfortable'). An explanation for this could be that respondents were not aware of the presence of these items in the environment which led them to give highly variable answers to these questions as they were unsure how to answer them. A second issue which arose at this stage was that the Space and Layout and Hygiene dimensions were found to lack discriminant validity. Up until this point, these two constructs have always remained distinct, however it seems reasonable that these two constructs should go hand in hand as clean environments provide a greater feeling of space therefore making these two constructs interdependent and conceptually very similar. This chapter also sought to resolve whether the proposed servicescape construct is a more accurate description of the servicescape than those presented by Baker (1987) and Bitner (1992). Structural equation modelling was used to meet this end and it was found that Bitner's (1992) model far exceeded the current model and Baker's (1987) model in terms of fit. While

statistically Bitner's (1992) model was found to be superior, its superiority ends here with its content lacking as many important environmental features such as hygiene and equipment entirely omitted. The model being put forward here, performs slightly worse statistically, yet contains a more comprehensive servicescape structure which would serve practitioners and managers better than the overly parsimonious Bitner (1992) model.

This comparative analysis also highlighted how the two primary theories in this area (Baker's and Bitner's) conflict with regards to where they place employees in their servicescape frameworks. Bitner (1992) entirely excludes the role employees have on servicescape perceptions but instead views the servicescape as a stimulus which affects employees in the same way it does customers. Baker (1987) on the other hand contends that employees serve as an environmental stimulus which contributes to our overall environmental perceptions. Theoretically, categorising the behaviour of employees as a servicescape dimension causes difficulties, particularly since the service quality literature is replete with papers emphasising employee behaviour as a central tenet of this theory. So as to avoid construct and theoretical confusion, employee behaviour items were removed from the second-order factor structure of the servicescape. This final four factor servicescape representation resulted in an even greater improvement to fit thus adding credence to the hypothesis that employees do not theoretically and conceptually fit well with the remaining tangible items. On the whole, Chapter Four illustrated how the servicescape is comprised of a discrete number of dimensions which work together to form consumers' overarching perceptions of this construct. In addition, this chapter also

highlights the strength of the servicescape measure developed in Chapter Two as it was possible to discriminate statistically between the four constructs while at the same time showed how consumers perceive them to be linked by the higher order, all encompassing servicescape construct. A final contribution made by this chapter was that it compared the current model with those proposed by Bitner (1992) and Baker (1987) and from this it was possible to ascertain that while the current model exhibits slightly worse fit, it is preferable as it captures a more complete view of this involute construct.

Within Chapter Five the focus then turned to the service quality literature. It becomes apparent from the literature reviewed that the boundaries between the service quality and servicescape literature are very blurred. Theories and empirical research on service quality have invariably included tangible, servicescape-type items as part of service quality (Parasuraman et al, 1985; Parasuraman et al, 1988; Cronin and Taylor, 1992). When Bitner's (1992) seminal work on servicescapes is examined it is evident that no conscious or explicit effort was made by her to delineate or conceptually remove the servicescape from service quality. It was proposed here that a distinction needs to be made between these two constructs if theory is to advance further. Moreover, it was hypothesised that the servicescape, rather than being regarded as another service quality dimension, would instead serve as a cue which would impact service quality evaluations and act as an antecedent to service quality. To meet this end, service quality was modelled as a two-dimensional structure with the four servicescape constructs seen as direct predictors. The findings from this stage of analysis were mixed with some of the paths between the constructs found to be insignificant. However, those paths for which

support was found were easily interpreted and logical. Firstly, equipment had a direct and positive effect on Outcome Service Quality but not on perceptions of employees which clearly reflects how the equipment in this service context substitutes employee service. Secondly, as with the ambience items measured in Chapter Four, it was found that music was entirely insignificant and it was deduced from this that this occurred because respondents found the music imperceptible. A final finding here was that Hygiene and Space had a direct effect on employee service quality evaluations which suggests that customers attribute the maintenance of the environment to the service staff. In sum, the contributions made in this chapter were that firstly, service quality is best viewed as a parsimonious, two dimensional structure which consists of an employee/interaction dimension and an outcome service quality dimension. Secondly, this chapter also showed how the servicescape is a separate entity to the service quality construct and that elements of the servicescape serve as cues which have direct effects on service quality perceptions.

It is almost ubiquitously assumed that emotional reactions are central to our understanding of how environments shape behaviour. Similarly, the customer satisfaction literature strongly advocates that emotional, or affective responses play an important role in the evaluative response. The literature underpinning emotion and customer satisfaction were therefore extensively reviewed in Chapter Six. This review culminated in the development of a full research model in which the servicescape is a primary antecedent of emotions and service quality evaluations and an indirect antecedent of customer satisfaction and behavioural intentions. This model served three purposes.

Firstly, it allowed the researcher to determine whether consumers in this context rely on a more emotional processing system or a more logical, cognitive evaluative process. Secondly, through the conceptualisation of the servicescape as a number of discrete dimensions it permits one to determine which of the dimensions has the ability to elicit emotional responses. Finally, this model also makes a contribution in that no study to date has investigated the multiple direct and indirect links between the servicescape, emotion, service quality, customer satisfaction and behavioural intentions.

The hypotheses proposed in Chapter Six were subsequently tested in Chapter Seven. The first notable finding was that very few of the servicescape → emotion → customer satisfaction paths were significant. In addition, many of the emotion items were found to be unreliable with the pool of twenty emotions radically reduced to seven. This sheds serious doubt over the role that emotions play in this service. Through the inspection of descriptive statistics this uncertainty was confirmed with only two of the emotions having mean responses greater than two, therefore consolidating the notion that emotions are experienced at only a negligible level for this service. This contrasts starkly with the literature to date, however when one reflects upon the context of the study it seems a rather obvious and reasonable finding. Firstly, as customers do not spend a significant amount of time in service stations it reduces the probability that they would experience a broad range of emotions. Secondly, the purpose for which customers use this service are likely to be satisfying utilitarian, rather than hedonic needs, thus leading to a more cognitively primed experience.

Given that only limited support could be found for the emotions paths, the data analysis proceeded to examine whether a cognitively oriented approach is more appropriate. The first finding from this analysis is that nomological validity of the servicescape measure was found. The ambience/music dimension is the exception to this, as no paths between this dimension and any other included in the final model were significant. Denouncing music and other atmospheric cues as irrelevant would be hasty, particularly since single stimulus atmospheric studies have consistently shown how experimental manipulations of variables such as music can cause changes in behaviour. The likelihood is that the respondents were completely oblivious to the music and other atmospheric variables present in the store, therefore when questioned about their perceptions of these items respondents were unsure how to answer leading to highly variable responses.

As was found in Chapter Five, it was again confirmed that the servicescape is best described as an antecedent of service quality. Structural equation modelling also showed that the servicescape design influenced behavioural intentions over and above any of the other constructs in the model. This is of significant importance as it displays how powerful the servicescape is in adding to the service experience. Indeed, it signifies that customers of this store would return to the store primarily because the design is both enticing and appealing. Of the three servicescape constructs, equipment had the weakest effect on behavioural intentions. These findings demonstrate how conceptualising and modelling the servicescape as discrete dimensions allows one to determine which aspects of the store design contribute most to the evaluative constructs. Moreover, it also shows how each of the servicescape dimensions behave in a unique way influencing perceptions

and behaviour either directly or indirectly. Thus the servicescape and other tangible cues can be seen as service facilitators wherein they act as an indication of the quality of the service and are thereby an integral part of the service product.

8.3 Limitations to the Study and Recommendations for Further Research

As does any research project, this study has some limitations. The service tested accounts for only a small portion of utilitarian services which makes generalising the results difficult. The present findings suggest that Outcome Service Quality better predicts behavioural intentions than customer satisfaction. Given that the current study is limited to one organisation in one service industry, this assertion, along with others found in the study should be expanded and validated by further research. Further research could extend the current findings by examining a range of other utilitarian services to determine whether the findings presented here hold true. Similarly, a comparison to a hedonic service using a methodology such as multi-group structural equation modelling would allow one to assess whether certain services have either a different servicescape structure or to see if ambient cues play a role in alternative services. As previous studies have conducted research on hedonic services such as casinos, sports stadiums and recreation centres, future research should choose a hedonic service that is different to those already researched. Therefore, an interesting line of inquiry could be to choose a hedonic retail service such as high-end boutiques or haute couture contexts as multidimensional servicescape research seems to be lacking in this area. In researching hedonic services it may also be found that emotions are indeed a significant mediator of the servicescape-satisfaction relationship as the more experiential nature of these services

may trigger emotional responses. Although the measures scored well within the service stations, further analyses in different contexts will also allow for further research of the servicescape items and will provide more definitive proof of its reliability and validity.

The employment of a cross-sectional design also poses as a limitation as all measures were collected simultaneously. Due to this, inferences should only be made as associations between constructs rather than assuming causal relationships. The cross-sectional nature of this study also means that the study represents a static model of service evaluation where the findings are representative of a single point in time. To overcome this, further research should seek to employ a longitudinal design to see whether consumers' perceptions of the service stations changes over time. This is particularly pertinent to the service stations included in the current study as they are presently undergoing refurbishment and being re-branded under a different name. By conducting a longitudinal study, it may provide very interesting insights into the role of the servicescape and how consumers' perceptions evolve over time and whether the chosen servicescape design is viewed more favourably than the old one.

This study has presented all results on an aggregate level which may cloud whether differences exist between groups. Therefore a promising avenue for future research would be to determine whether differences exist between customers of varying usage levels, customer type, gender, social status or age. This would be worthwhile to test as the data provided from these studies could be used for segmentation purposes and also because it may indicate that different groups have different perceptions and responses to

the service provided by the service stations. Researchers should examine other moderators of the relationships tested in the model such as whether consumers' level of involvement with the service has an effect on their evaluation of the service. Previous research has shown that when involvement is high the relationship between customer satisfaction and purchase intentions is stronger (Bolton, 1998). In the current research the correlation between customer satisfaction and behavioural intentions was not particularly strong at 0.16 and perhaps if involvement mediated this relationship it would elucidate the relationship further. Further research could also pursue the role of value and how it moderates the relationships in the model. Price and value are likely to be of significant importance in the current research situation where it is possible that customers seek out service stations that offer the cheapest petrol.

The measurement of behavioural intentions also poses two limitations. Firstly, because an individual states that they intend to repatronage the service stations does not mean that this will come to fruition (Morwitz, 1997). Situational and normative constraints are two such phenomena that may intervene the process that cause individuals to either purchase from an alternative service or to not purchase at all (Morwitz, 1997). Similarly, self-generated validity effects may actually strengthen the relationship between customer satisfaction and behavioural consequences (Chandon et al, 2005). These disadvantages could be overcome through the measurement of actual behaviour which could be an objective of future research in this area. For example employing alternative measures such as time spent in store, actual spending or the number of items purchased such as those used by Donovan et al (1994).

The second problem with the measurement of behavioural intentions here was that merely three items were used to capture the construct. Although previous research has employed the same number of items and has demonstrated internal reliability (Cronin et al, 2000; Brady et al, 2005) this research could have benefited from the inclusion of additional behavioural intentions items that captured switching behaviour. It is likely that for the service industry under scrutiny here that consumers are highly price sensitive. With oil prices on the rise, price sensitivity is likely to be a significant reason for people to switch, therefore including items such as, 'Take some of my business to a competitor that offers better prices' (Zeithaml et al, 1996) may have further reduced the relationship between satisfaction and behavioural intentions. Convenience may also be an issue in the current context with individuals patronising a service station only when their petrol tank reaches a certain level. This may then cause respondents to frequent a competitor service station that is closest to them at this time. Unfortunately Zeithaml et al's (1996) measure does not have an item that captures this phenomenon therefore future research could seek to design an item that captures this issue.

Interviewer bias (Boyd and Westfall, 1970) may also have been a source of error in the current research design. To reduce the influence of interviewer bias, pains were taken to ensure that the interviewers were sufficiently trained and selected. Alternative research designs which entirely eliminate the presence of the researcher such as internet methods or mail methods would not have been suitable for the present study as it would have

involved asking respondents to reflect on a previous experience with the service provider which would ultimately have led to different types of error being introduced.

The purely quantitative nature of this study could also be viewed as a drawback. It is possible that further insights could be gleaned and the relationships explained in more detail if qualitative research was conducted as a form of triangulation procedure. For example in the discussion section that accompanies each stage of the analysis, qualitative research in the form of interviews could have been used to explain the findings. The procedures adhered to at the servicescape scale development stage would also have benefited from qualitative research in the form of focus groups and in-depth interviews. Unfortunately time and monetary constraints limited the ability for the researcher to engage in these techniques and perhaps further research should seek to develop a servicescape scale which incorporates this important stage. A second limitation associated with the servicescape scale is that when the scale was being devised respondents were asked to reflect on a different service to the one eventually used in the study. The reason for this is that originally the focus of this PhD was to be professional services, in particular dentists. For reasons outside of the researchers' control, it was not possible to gain access to such a service with all those contacted and asked to participate refusing to take part. Therefore, the scale that was devised was modified very slightly and used in the final empirical analyses. Fortunately this did not cause any problems at the measurement stage with all items loading according to a meaningful structure.

Since the research that is presented in this dissertation has been conducted, an emotions scale that has been specifically designed for service contexts has been published (Diamantopoulos and Schoefer, 2008). This scale, known as Emotions During Service Recovery Encounters (ESRE), although initially devised for use when service failures occur, has been recommended to be used in customer satisfaction research. Given that Diamantopoulos and Schoefer's (2008) scale is shorter than the PANAS (Watson et al, 1988) used here, and was designed with services in mind, perhaps further research should attempt to employ this scale rather than PANAS. This is particularly important as many of the items in the PANAS scale were found to be highly unreliable.

8.4 Managerial Implications

As implied by the discussion in the preceding sections, this research offers new and significant insights and emphasises the need for continuing research to examine the generalisability of the findings. Despite the limitations of the study, the research presented here offers some guidelines for the managers of the service stations that can be used to improve their service offering. These will be discussed within the next section.

It is important that managers recognise the influence of the servicescape as a marketing tool. Up until this point, the extant literature has provided little insight into the linkages between various servicescape elements and quality, customer satisfaction and behavioural intention inferences. Much of the literature to date has tended to focus on how the store environment affects consumers' perceptions of merchandise quality and store image with no study to date examining how the servicescape impacts quality, customer satisfaction

and behavioural intentions. For store managers of this service station, the results of this study suggest that ambient cues are a servicescape element that has little impact on overall experience. This is not to say that managers can ignore this facet of the servicescape, as it is highly likely that if music were to be removed from the store environment respondents would have a negative reaction to it. Although the generalisability of these findings to other types of retail environments has yet to be established, the ability for service stations to change the design of the store is relatively easy.

Of the antecedents to behavioural intentions studies, it is interesting to find that the design construct contributes the greatest amount to behavioural intentions (see Table 7.13). This is a very promising insight as it means that more manageable and static elements of the service experience have the greatest influence on consumers' behaviour. In this service context, it is very difficult for managers to both retain and attract the 'right' kind of staff. In fact, out of all aspects of the service offering it is the most complex part to manage and control and is likely to be highly variable from one service provider to the next. Therefore, it is very reassuring to see that the physical aspects of the service, which are not variable, are the facet of the service experience that holds the most value for customers. Because of this, the servicescape should be under constant evaluation and observation by management, while a major re-haul of the design might be both costly and difficult to do, repainting the store and adding decorative features will maintain its attractiveness. If the upkeep of the servicescape is let slide, the likelihood is that this will cause customers to patronise a competitor service station.

The cleanliness and hygiene of the store was also found to be an important determinant of the outcome variables. In terms of control, this element of the service design is even easier to manage than the design construct. All flooring and outdoor and indoor space should at all times be free from litter and dirt and management must oversee regular cleaning of the environment. It is also essential that managers understand that perceptions of the cleanliness of the environment has a direct effect on how consumers perceive the service staff. The reasoning behind this is that it seems that customers attribute the upkeep of the store as the responsibility of the staff and therefore it should be realised by management that if they wish customers to perceive staff in a positive light then care should be taken to maintain a hygienic and clean store environment. The equipment dimension had quite a modest impact on behavioural intentions. Despite this, managers should attempt to ensure that all self-service equipment is in working order and without any mechanical defects. Also, as this dimension currently contributes the least to behavioural intentions, perhaps improvements to the equipment need to be made. One such improvement could be to provide payment facilities at the petrol pump which would make the purchase of petrol much more convenient for consumers. On the whole, constant attention needs to be paid to the servicescape and innovations and ways to improve it should constantly be considered by management.

This research also sought to clarify the role of emotions in the service experience. From the results found here, it is clear that emotions do not play an integral part in the servicescape-satisfaction response. Much of the previous work on consumer emotions

has found support that the consumption process contains some affective reactions. Generally the recommendations provided by these authors are to either maximise positive emotional reactions (which will ultimately have favourable consequences on customer satisfaction and behavioural intentions) or to minimise the negative ones. The problem with this is that knowing how to minimise or maximise emotional responses is quite a difficult task. This is even more of a problem for managers when they are given vague recommendations that state they should train staff to gauge emotional responses from facial reactions. Fortunately for the management of this service station, stimulating or reading emotions is not a priority. The thought process that consumers engage in for this service is entirely cognitive, whereby they methodologically determine their level of satisfaction. For the managers of service stations this makes improvements to the service easier to make as their customers can be viewed as rational beings who adhere to logical thought processes.

As well as striving towards a servicescape design that optimises repatronage intentions, managers should also consider the importance of service quality and customer satisfaction as strategic goals. Training programmes may need to be introduced to capitalise on the delivery of service quality and to ensure high levels of customer satisfaction are achieved. The benefits of such training will not only increase customer satisfaction and perceived levels of service quality but will in turn provide a competitive advantage through increasing levels of customer retention.

REFERENCES

Acock, Alan C. (2005), "Working with Missing Values," *Journal of Marriage and Family*, 67 (4), 1012-28.

Ajzen, Icek (2001), "Nature and Operation of Attitudes," *Annual Review of Psychology*, 52 (1), 27-58.

Ajzen, Icek and Martin Fishbein (1980), *Understanding Attitudes and Predicting Social Behavior*. Englewood Cliffs, NJ: Prentice Hall.

Albers-Miller, Nancy D. and Marla Royne Stafford (1999), "International Services Advertising: An Examination of Variation in Appeal Use for Experiential and Utilitarian Services," *Journal of Services Marketing*, 13 (4/5), 390-406.

Alford, Bruce L. and Daniel L. Sherrell (1996), "The role of affect in consumer satisfaction: judgment of credence-based services," *Journal of Business Research*, 37 (1), 71-84.

Anderson, Eugene W. and Claes Fornell (1994), "A Customer Satisfaction Research Prospectus," in *Service Quality: New Directions in Theory and Practice*, Roland T. Rust and Richard L. Oliver, Eds. Thousand Oaks, CA: Sage Publications, 241-68.

Anderson, Eugene W., Claes Fornell, and Donald R. Lehmann (1994), "Customer Satisfaction, Market Share, and Profitability: Findings from Sweden," *Journal of Marketing*, 58 (3), 53-66.

Anderson, Eugene W., Claes Fornell, and Sanal K. Mazvancheryl (2004), "Customer Satisfaction and Shareholder Value," *Journal of Marketing*, 68 (4), 172-85.

Anderson, Eugene W. and Vikas Mittal (2000), "Strengthening the Satisfaction-Profit Chain," *Journal of Service Research*, 3 (2), 107-20.

Anderson, Eugene W. and Mary W. Sullivan (1993), "The Antecedents and Consequences of Customer Satisfaction for Firms," *Marketing Science*, 12 (2), 125-43.

Anderson, James C. and David W. Gerbing (1988), "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach," *Psychological Bulletin*, 103 (3), 411-23.

Andreassen, Tor Wallin and Bodil Lindestad (1998), "Customer Loyalty and Complex Services: The Impact of Corporate Image on Quality, Customer Satisfaction and Loyalty for Customers with Varying Degrees of Service Expertise," *International Journal of Service Industry Management*, 9 (1), 7-23.

Areni, C. and D. Kim (1993), "The Influence of Background Music on Shopping Behavior: Classical Versus Top-Forty Music in a Wine Store," *Advances in Consumer Research*, 20 (1), 336-40.

Areni, C. and D. Kim (1994), "The Influence of In-Store Lighting on Consumers' Examination of Merchandise in a Wine Store," *International Journal of Research in Marketing*, 11 (2), 117-25.

Arndt, Johan (1967), "Role of Product-Related Conversations in the Diffusion of a New Product," *Journal of Marketing Research*, 4 (3), 291-95.

Arnould, Eric J. and Linda L. Price (1993), "River Magic: Extraordinary experience and the extended service encounter," *Journal of Consumer Research*, 20 (1), 24-45.

Arnould, Eric J., Linda L. Price, and P. Tierney (1998), "Communicative Staging of the Wilderness Servicescape," *The Service Industries Journal*, 18 (3), 90-115.

Atkinson, Rita (2000), *Hilgard's Introduction to Psychology*. London: Harcourt College Publishers.

Aubert-Gamet, Veronique (1999), "Twisting Servicescapes: Diversion of the Physical Environment in a Re-Appropriation Process " *International Journal of Service Industry Management*, 8 (1), 26-41.

Aydin, Serkan and Gökhan Özer (2005), "The Analysis of Antecedents of Customer Loyalty in the Turkish Mobile Telecommunication Market," *European Journal of Marketing*, 39 (7/8), 910-25.

Aylott, Russell and Vincent-Wayne Mitchell (1998), "An Exploratory Study of Grocery Shopping Stressors," *International Journal of Retail & Distribution Management*, 26 (9), 362-73.

Babakus, Emin and Gregory W. Boller (1992), "An Empirical Assessment of the SERVQUAL Scale," *Journal of Business Research*, 24 (2), 253-68.

Bagozzi, Richard P. (1980), *Causal Models in Marketing*. New York: John Wiley and Sons.

Bagozzi, Richard P., Mahesh Gopinath, and Prashanth U. Nyer (1999), "The Role of Emotions in Marketing," *Journal of the Academy of Marketing Science*, 27 (2), 184-206.

Bagozzi, Richard P. and Youjae Yi (1991), "Multitrait-Multimethod Matrices in Consumer Research," *Journal of Consumer Research*, 17 (4), 426-39.

Bagozzi, Richard P., Youjae Yi, and Lynn W. Phillips (1991), "Assessing Construct Validity in Organizational Research," *Adminstrative Science Quarterly*, 36 (3), 421-58.

Baker, Julie (1987), "The Role of the Environment in Marketing Services: The Consumer Perspective," in *The Services Challenge: Integrating for Competitive Advantage*, John A. Czepeil and Caroline A. Congram and James Shanahan, Eds. Chicago: American Marketing Association, 74-84.

Baker, Julie, Dhruv Grewal, and A. Parasuraman (1994), "The Influence of Store Environment on Quality Inferences and Store Image," *Journal of the Academy of Marketing Science*, 22 (4), 328-29.

Baker, Julie, A. Parasuraman, Dhruv Grewal, and Glenn B. Voss (2002), "The Influence of Multiple Store Environment Cues on Perceived Merchandise Value and Patronage Intentions," *Journal of Marketing*, 66 (2), 120-41.

Bandalos, Deborah L. and Craig K. Enders (1996), "The Effects of Nonnormality and Number of Response Categories on Reliability," *Applied Measurement in Education*, 9 (2), 151-60.

Barrett, Paul (2007), "Structural Equation Modelling: Adjudging Model Fit," *Personality and Individual Differences*, 42 (5), 815-24.

Bebko, Charlene Pleger (2000), "Service Intangibility and its Impact on Consumer Expectations of Service Quality," *Journal of Services Marketing*, 14 (1), 9-26.

Bellizzi, Joseph A., Ayn E. Crowley, and Ronald W. Hasty (1983), "The Effects of Color in Store Design," *Journal of Retailing*, 59 (1), 21-45.

Bellizzi, Joseph A. and Robert E. Hite (1992), "Environmental Color, Consumer Feelings, and Purchase Likelihood," *Psychology and Marketing*, 9 (5), 347-63.

Bendall-Lyon, Dawn and Thomas L. Powers (2004), "The impact of structure and process attributes on satisfaction and behavioural intentions," *Journal of Services Marketing*, 18 (2), 114-21.

Benson, Jeri and John A. Fleishman (1994), "The robustness of maximum likelihood and distribution-free estimators to non-normality in confirmatory factor analysis," *Quality and Quantity*, 28 (2), 117-36.

Bentler, Peter M. (1980), "Multivariate Analysis with Latent Variables: Causal Modeling," *Annual Review of Psychology*, 31 (1), 419-56.

Bentler, Peter M. (1990), "Comparative Fit Indexes in Structural Models," *Psychological Bulletin*, 107 (2), 238-46.

Bentler, Peter M. and Douglas C. Bonnet (1980), "Significance Tests and Goodness of Fit in the Analysis of Covariance Structures," *Psychological Bulletin*, 88 (3), 588-606.

Berry, Leonard L., Lewis P. Carbone, and Stephan H. Haeckel (2002), "Managing the Total Customer Experience," *Sloan Management Review*, 43 (3), 85-89.

Bigné, J. Enrique, Luisa Andreu, and Gnoth Juergen (2005), "The theme park experience: An analysis of pleasure, arousal and satisfaction," *Tourism Management*, 26 (6), 833-44.

Bigné, J. Enrique, Anna S. Mattila, and Luisa Andreu (2008), "The Impact of Experiential Consumption Cognitions and Emotions on Behavioral Intentions," *Journal of Services Marketing*, 22 (4), 303-15.

Bitner, Mary Jo (1990), "Evaluating Service Encounters: The Effect of Physical Surroundings and Employee Responses," *Journal of Marketing*, 54 (2), 69-84.

Bitner, Mary Jo (1992), "Servicescapes: The Impact of Physical Surroundings on Customers and Employees," *Journal of Marketing*, 56 (2), 57-71.

Bitner, Mary Jo (2000), "The Servicescape," in *Handbook of Services Marketing and Management*, Teresa A. Swartz and Dawn Iacobucci, Eds. Thousand Oaks, California: Sage Publications, 37-50.

Bitner, Mary Jo and Amy R. Hubbert (1994), "Encounter Satisfaction Versus Overall Satisfaction Versus Quality," in *Service Quality: New Directions in Theory and Practice*, Roland T. Rust and Richard L. Oliver, Eds. Thousand Oaks, California: Sage Publications, 72-94.

Blackwell, Louise (1995), "Visual Cues and Their Effects on Odour Assessment," *Nutrition and Food Science*, 5 (5), 24-28.

Bloch, Peter H., Nancy M. Ridgway, and Daniel L. Sherrell (1989), "Extending the Concept of Shopping: An Investigation of Browsing Activity," *Journal of the Academy of Marketing Science*, 17 (1), 13-21.

Bloemer, José and Ko de Ruyter (1999), "Customer Loyalty in High and Low Involvement Service Settings: The Moderating Impact of Positive Emotions," *Journal of Marketing Management*, 15 (4), 315-30.

Bollen, Kenneth A. (1989), *Structural Equations with Latent Variables*. New York: Wiley.

Bollen, Kenneth A. (1990), "Overall Fit in Covariance Structure Models: Two Types of Sample Size Effects," *Psychological Bulletin*, 107 (2), 256-59.

Bolton, Ruth N. (1998), "A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: The Role of Satisfaction," *Marketing Science*, 17 (1), 45-65.

Bolton, Ruth N. and James H. Drew (1991), "A Multistage Model of Customers' Assessments of Service Quality and Value," *Journal of Consumer Research*, 17 (4), 375-84.

Bolton, Ruth N. and Katherine N. Lemon (1999), "A Dynamic Model of Customers' Usage of Services: Usage as an Antecedent and Consequence of Satisfaction," *Journal of Marketing Research*, 36 (2), 171-86.

Bone, Paula Fitzgerald and Pam Scholder Ellen (1999), "Scents in the Marketplace: Explaining a Fraction of Olfaction," *Journal of Retailing*, 75 (2), 243-62.

Boulding, William, Ajay Kalra, Richard Staelin, and Valerie A. Zeithaml (1993), "A Dynamic Process of Service Quality: From Expectations to Behavioral Intentions," *Journal of Marketing Research*, 30 (1), 7-27.

Bowen, David E. (1990), "Development of a Taxonomy of Services to Gain Strategic Marketing Insights," *Journal of the Academy of Marketing Science*, 18 (1), 43-9.

Boyd, Harper W. and Ralph Westfall (1970), "Interviewer Bias Once More Revisited," *Journal of Marketing Research*, 7 (2), 249-53.

Brady, Michael K. and J. Joseph Cronin (2001), "Some New Thoughts on Conceptualizing Perceived Service Quality: A Hierarchical Approach," *Journal of Marketing*, 65 (3), 34-49.

Brady, Michael K., Gary A. Knight, J. Joseph Cronin, G. Tomas M. Hult, and Bruce D. Keillor (2005), "Removing the Contextual Lens: A Multinational, Multi-setting Comparison of Service Evaluation Models," *Journal of Business Research*, 81 (3), 215-30.

Brady, Michael K. and Christopher J. Robertson (2001), "Searching for a Consensus on the Antecedent Role of Service Quality and Satisfaction: An Exploratory Cross-National Study," *Journal of Business Research*, 51 (1), 53-60.

Briggs, Stephen R. (1992), "Assessing the Five-Factor Model of Personality Description," *Journal of Personality*, 60 (2), 253-93.

Brogowicz, Andrew A., Linda M. Delene, and David M. Lyth (1990), "A Synthesised Service Quality Model with Managerial Implications," *International Journal of Service Industry Management*, 1 (1), 27-45.

Brown, Tom J., Gilbert A. Churchill, and J. Paul Peter (1993), "Improving the Measurement of Service Quality," *Journal of Retailing*, 69 (1), 127-39.

Bruner, Gordon C. (1990), "Music, Mood, and Marketing," *Journal of Marketing*, 54 (4), 94-104.

Buck, Ross and Mats Georgson (1997), "Consumer Emotion Space: An Investigation of Semantic Space and Context Effects in Self-Reported Emotion Elicitation," *Advances in Consumer Research*, 24 (1), 431-37.

Burger, Jerry M. (1999), "The Foot-in-the-Door Compliance Procedure: A Multiple-Process Analysis and Review," *Personality and Social Psychology Review*, 3 (4), 303-25.

Burns, David J. and Lewis Neisner (2006), "Customer satisfaction in a retail setting: the contribution of emotion," *International Journal of Retail and Distribution Management*, 34 (1), 49-66.

Bush, Alan J. and Joseph S. Hair (1985), "An Assessment of the Mall Intercept as a Data Collection Method," *Journal of Marketing Research*, 22 (3), 158-67.

Buttle, Francis (1996), "SERVQUAL: review, critique, research agenda," *European Journal of Marketing*, 30 (1), 8-32.

Buttle, Francis A. (1998), "Word of Mouth: Understanding and Managing Referral Marketing," *Journal of Strategic Marketing*, 6 (3), 241-54.

Byrne, Barbara M. (1998), *Structural Equation Modeling with LISREL, PRELIS and SIMPLIS: Basic Concepts, Applications and Programming*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Caceres, Ruben Chumpitaz and Nicholas G. Paparoidamis (2005), "Service Quality, Relationship Satisfaction, Trust, Commitment and Business-to-Business Loyalty," *European Journal of Marketing*, 41 (7/8), 836-67.

Cadotte, Ernest R., Robert B. Woodruff, and Roger L. Jenkins (1987), "Expectations and Norms in Models of Consumer Satisfaction," *Journal of Marketing Research*, 24 (3), 305-14.

Carbone, Lewis P. and Stephan H. Haeckel (1994), "Engineering Customer Experiences," *Marketing Management*, 3 (3), 8-19.

Carman, James M. (1990), "Consumer Perceptions of Service Quality: An Assessment of the SERVQUAL Dimensions," *Journal of Retailing*, 66 (1), 33-55.

Carrillat, Francois A., Fernando Jaramillo, and Jay P. Mulki (2007), "The Validity of the SERVQUAL and SERVPERF Scales: A Meta-Analytic View of 17 Years Research

Across Five Countries," *International Journal of Service Industry Management*, 18 (5), 472-90.

Carroll, James M., Michelle S.M. Yik, James A. Russell, and Lisa Feldman Barrett (1999), "On the Psychometric Principles of Affect," *Review of General Psychology*, 3 (1), 14-22.

Caruana, Albert (2002), "Service loyalty: The effects of service quality and the mediating role of customer satisfaction," *European Journal of Marketing*, 36 (7/8), 811-28.

Cattell, Raymond B. and S. Vogelmann (1977), "A Comprehensive Trial of the Scree and KG Criteria for Determining the Number of Factors," *Multivariate Behavioral Research*, 12 (3), 289.

Chandon, Pierre, Vicki G. Morwitz, and Werner J. Reinartz (2005), "Do Intentions Really Predict Behavior? Self Generated Validity Effects in Survey Research," *Journal of Marketing*, 69 (2), 1-14.

Chebat, Jean-Charles, C.G. Chebat, and D. Vaillant (2001), "Environmental background music and in-store selling," *Journal of Business Research*, 54 (2), 115-23.

Chebat, Jean-Charles, Pierre Filiatrault, Claire Gelinas-Chebat, and Alexander Vaninsky (1995), "Impact of waiting attribution and consumer's mood on perceived quality," *Journal of Business Research*, 34 (3), 191-96.

Chebat, Jean-Charles and Richard Michon (2003), "Impact of Ambient Odors on Mall Shoppers' Emotions, Cognition, and Spending: A Test of Competitive Causal Theories," *Journal of Business Research*, 56 (7), 529-39.

Chen, Fang Fang, Karen H. Sousa, and Stephen G. West (2005), "Teacher's Corner: Testing Measurement Invariance of Second-Order Factor Models," *Structural Equation Modeling*, 12 (3), 471-92.

Chevalier, Michel (1975), "Increase in Sales Due to In-Store Display," *Journal of Marketing Research*, 12 (4), 426-31.

Childers, Terry L., Christopher L. Carr, Joann Peck, and Stephen Carson (2001), "Hedonic and Utilitarian Shopping Motivations for Online Retail Behavior," *Journal of Retailing*, 77 (4), 511-35.

Chin, Wynne W. (1998), "The Partial Least Squares Approach to Structural Equation Modeling," in *Modern Methods for Business Research*, George A. Marcoulides, Ed. New Jersey: Lawrence Erlbaum Associates Inc., 295-336.

Chitturi, Ravindra, Rajagopal Raghunathan, and Vijay Mahajan (2008), "Delight by Design: The Role of Hedonic Versus Utilitarian Brands," *Journal of Marketing*, 72 (3), 48-63.

Choi, Kui-Son, Woo-Hyun Cho, Sunhee Lee, Hanjoon Lee, and Chankon Kim (2004), "The relationships among quality, value, satisfaction and behavioral intention in health care provider choice: A South Korean study," *Journal of Business Research*, 57 (8), 913-21.

Churchill, Gilbert A. (1979), "A Paradigm for Developing Better Measures of Marketing Constructs," *Journal of Marketing Research*, 16 (1), 64-73.

Churchill, Gilbert A. and Carol Suprenant (1982), "An Investigation into the Determinants of Customer Satisfaction," *Journal of Marketing Research*, 19 (4), 491-5-4.

Clark, Lee Anna and David Watson (1995), "Constructing Validity: Basic Issues in Objective Scale Development," *Psychological Assessment*, 7 (3), 309-19.

Clow, Kenneth E. and Douglas W. Vorhies (1993), "Building Competitive Advantage for Service Firms," *Journal of Services Marketing*, 7 (1), 22-32.

Comrey, Andrew L. (1988), "Factor-Analytic Methods of Scale Development in Personality and Clinical Psychology," *Journal of Consulting and Clinical Psychology*, 56 (5), 754-61.

Cortina, Jose M. (1993), "What is Coefficient Alpha? An Examination of Theory and Applications," *Journal of Applied Psychology*, 78 (1), 98-104.

Costello, Anna B. and Jason W. Osborne (2005), "Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis," *Practical Assessment, Research & Evaluation*, 10 (7), 1-9.

Cox, Eli P. (1980), "The Optimal Number of Response Alternatives for a Scale: A Review," *Journal of Marketing Research*, 17 (4), 407-22.

Cronbach, Lee J. and Paul E. Meehl (1955), "Construct Validity in Psychological Tests," *Psychological Bulletin*, 52 (4), 281-302.

Cronin, J. Joseph, Michael K. Brady, R.R. Brand, Roscoe Hightower, and D.J Shemwell (1997), "A Cross-Sectional Test of the Effect and Conceptualisation of Service Value," *Journal of Services Marketing*, 11 (6), 375-91.

Cronin, J. Joseph, Michael K. Brady, and G. Tomas M. Hult (2000), "Assessing the Effects of Quality, Value, and Customer Satisfaction on Consumer Behavioural Intentions in Service Environments," *Journal of Retailing*, 76 (2), 193-218.

Cronin, J. Joseph and Steven A. Taylor (1992), "Measuring Service Quality: A Reexamination and Extension," *Journal of Marketing*, 56 (3), 55-68.

Cronin, J. Joseph and Steven A. Taylor (1994), "SERVPERF Versus SERVQUAL: Reconciling Performance-Based and Perceptions-Minus-Expectations Measurement of Service Quality," *Journal of Marketing*, 58 (1), 125-31.

Dabholkar, Pratibha A., C. David Shepherd, and Dayle I. Thorpe (2000), "A Comprehenisve Framework for Service Quality: An Investigation of Critical Conceptual and Measurement Issues Through a Longitudinal Study," *Journal of Retailing*, 76 (2), 139-73.

Danaher, Peter J. and Vanessa Haddrell (1996), "A Comparison of Question Scales Used for Measuring Customer Satisfaction," *International Journal of Service Industry Management*, 7 (4), 4-26.

Datamonitor (2007), "Retailing in Ireland 2007." London.

Davies, Barry J., Dion Kooijman, and Philippa Ward (2003), "The Sweet Smell of Success: Olfaction in Retailing," *Journal of Marketing Management*, 19 (5/6), 611-27.

Dawes, John (2008), "Do data characteristics change according to the number of scale points used?," *International Journal of Market Research*, 50 (1), 61-77.

Day, Ellen (1992), "Conveying Service Quality Through Advertising," *Journal of Services Marketing*, 6 (4), 53-61.

Day, George S. (1969), "A Two-Dimensional Concept of Brand Loyalty," *Journal of Advertising Research*, 9 (3), 55-68.

Day, George S. (1971), "Attitude Change, Media and Word of Mouth," *Journal of Advertising Research*, 11 (6), 31-40.

de Ruyter, Ko and José Bloemer (1999), "Customer Loyalty in Extended Service Settings: The Interaction Between Satisfaction, Value Attainment and Positive Mood," *International Journal of Service Industry Management*, 10 (3), 320-36.

Dean, Alison M. (2007), "The Impact of the Customer Orientation of Call Center Employees on Customers' Affective Commitment and Loyalty," *Journal of Service Research*, 10 (2), 161-73.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977), "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, 39 (1), 1-38.

DeVellis, Robert F. (2003), *Scale Development: Theory and Applications* (2nd ed.). Thousand Oaks, California: Sage Publications.

Dhar, Ravi and Klaus Wertenbroch (2000), "Consumer Choice Between Hedonic and Utilitarian Goods," *Journal of Marketing Research*, 37 (1), 60-71.

Diamantopoulos, Adamantios and Klaus Schoefer (2008), "Measuring Experienced Emotions During Service Recovery Encounters: Construction and Assessment of the ESRE Scale," *Service Business: An International Journal*, 2 (1), 65-81.

Diamantopoulos, Adamantios and Judy A. Siguaw (2000), *Introducing LISREL*. London: Sage Publications.

Diamantopoulos, Adamantios and Judy Siguaw, A. (2006), "Formative Versus Reflective Indicators in Organizational Measure Development: A Comparison and Empirical Illustration," *British Academy of Management*, 17, 263-82.

Dickson, John and Gerald Albaum (1977), "A Method of Developing Tailor Made Semantic Differentials for Specific Marketing Content Areas," *Journal of Marketing Research*, 14 (1), 87-91.

Donovan, Robert J. and John R. Rossiter (1982), "Store Atmosphere: An Environmental Psychology Approach," *Journal of Retailing*, 58 (1), 34-57.

Donovan, Robert J., John R. Rossiter, Gilian Marcoolyn, and Andrew Nesdale (1994), "Store Atmosphere and Purchasing Behavior," *Journal of Retailing*, 70 (3), 283-94.

Dubé, Laurette, C.G. Chebat, and S. Morin (1995), "The Effects of Background Music on Consumers' Desire to Affiliate in Buyer-Seller Interactions," *Psychology and Marketing*, 12 (4), 305-19.

Dubé, Laurette and Kalyani Menon (2000), "Multiple roles of consumption emotions in post-purchase satisfaction with extended service transactions," *International Journal of Service Industry Management*, 11 (3), 287-304.

Dubé, Laurette and S. Morin (2001), "Background music pleasure and store evaluation: Intensity effects and psychological mechanisms," *Journal of Business Research*, 54 (2), 107-13.

Durvasula, Srinivas, Steven Lysonski, and Subhash Mehta (1999), "Testing the SERVQUAL Scale in the Business-to-Business Sector: The Case of Ocean Freight Shipping Service," *Journal of Services Marketing*, 13 (2), 132-50.

Edell, Julie A. and Marian Chapman Burke (1987), "The Power of Feelings in Understanding Advertising Effects," *Journal of Consumer Research*, 14 (3), 421-33.

Edvardsson, Bo, Anders Gustafsson, and Inger Roos (2005), "Service Portraits in Service Research: A Critical Review," *International Journal of Service Industry Management*, 16 (1), 107-21.

Edwards, Jeffrey R. (2001), "Multidimensional Constructs in Organizational Behavior Research: An Integrative Analytical Framework," *Organizational Research Methods*, 4 (2), 144-92.

Edwards, Jeffrey R. and Richard P. Bagozzi (2000), "On the Nature and Direction of Relationships Between Constructs and Measures," *Psychological Methods*, 5 (2), 155-74.

Eisinger, Richard A., W. Peter Janicki, Robert L. Stevenson, and Wendel L. Thompson (1974), "Increasing Returns in International Mail Surveys," *Public Opinion Quarterly*, 38 (1), 124-30.

Ekinci, Yuksel (2001), "The validation of the generic service quality dimensions: an alternative approach," *Journal of Retailing and Consumer Services*, 8 (6), 311-24.

Ekinci, Yuksel, Michael Riley, and Chris Fife-Schaw (1998), "Which School of Thought? The Dimensions of Resort Hotel Quality," *International Journal of Contemporary Hospitality Management*, 10 (2), 63-67.

El Sayed, Ismail M., Dalia A. Farrag, and Russell W. Belk (2004), "The Effects of Physical Surroundings on Egyptian Consumers' Emotional States and Buying Intentions," *Journal of International Consumer Marketing*, 16 (1), 5-27.

Ellen, Pam Scholder and Paula Fitzgerald Bone (1998), "Does It Matter If It Smells? Olfactory Stimuli As Advertising Executional Cues," *Journal of Advertising*, 27 (4), 29-39.

Enders, Craig K. and Deborah L. Bandalos (2001), "The Relative Performance of Full Information Maximum Likelihood Estimation for Missing Data in Structural Equation Models," *Structural Equation Modeling*, 8 (3), 430-57.

Eroglu, Sevgin and Gilbert D. Harrell (1986), "Retail Crowding: Theoretical and Strategic Implications," *Journal of Retailing*, 62 (4), 346.

Evans, Phil (1988), *Motivation and Emotion*. London: Routledge.

Ezeh, Chris and Lloyd Harris (2007), "Servicescape Research: A Review and a Research Agenda," *The Marketing Review*, 7 (1), 59-78.

Fabrigar, Leandre R., Robert C. MacCallum, Duane T. Wegener, and Erin Strahan (1999), "Evaluating the Use of Exploratory Factor Analysis in Psychological Research," *Psychological Methods*, 4 (3), 272-99.

Fan, Xitao, Bruce Thompson, and Lin Wang (1999), "Effects of Sample Size, Estimation Methods, and Model Specification on Structural Equation Modeling Fit Indexes," *Structural Equation Modeling*, 6 (1), 56-83.

Finn, David W. and Charles W. Lamb (1991), "An Evaluation of the SERVQUAL Scales in a Retailing Setting," *Advances in Consumer Research*, 18 (1), 483-90.

Flamer, Stephen (1983), "Assessment of the Multitrait-Multimethod Matrix Validity of Likert Scales Via Confirmatory Factor Analysis," *Multivariate Behavioral Research*, 18 (3), 275-308.

Flipo, Jean Paul (1988), "On the Intangibility of Services," *The Service Industries Journal*, 8 (3), 286-98.

Floyd, Frank J. and Keith F. Widaman (1995), "Factor Analysis in the Development and Refinement of Clinical Assessment Instruments," *Psychological Assessment*, 7 (3), 286-99.

Folkes, Valerie S. (1984), "Consumer reactions to product failure: An attributional approach," *Journal of Consumer Research*, 10 (4), 398-409.

Folkes, Valerie S. (1988), "Recent Attribution Research in Consumer Behaviour: A Review and New Directions," *Journal of Consumer Research*, 14 (4), 548-65.

Folkes, Valerie S. and Barbara Kostos (1986), "Buyers' and Sellers' Explanations for Product Failure: Who Done It," *Journal of Marketing*, 50 (2), 74-80.

Ford, J. Kevin, Robert C. MacCallum, and Marianne Tait (1986), "The Application of Exploratory Factor Analysis in Applied Psychology: A Critical Review and Analysis," *Personnel Psychology*, 39 (2), 291-314.

Fornell, Claes and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 28 (1), 39-50.

Fornier, Susan and David Glen Mick (1999), "Rediscovering Satisfaction," *Journal of Marketing*, 63 (4), 5-23.

Foxall, Gordon (1984), "Marketing is Service Marketing," *The Service Industries Journal*, 4 (3), 1-6.

Foxall, Gordon R. (1999), "The behavioural perspective model: Consensibility and consensuality," *European Journal of Marketing*, 33 (5/6), 570-96.

Foxall, Gordon R. and Gordon E. Greenley (1999), "Consumers' Emotional Responses to Service Environments," *Journal of Business Research*, 46 (2), 149-58.

Freedman, Jonathan L. and Scott C. Fraser (1966), "Compliance Without Pressure: The Foot-in-the-Door Technique," *Journal of Personality and Social Psychology*, 4 (2), 195-202.

Friborg, Oddgeir, Monica Martinussen, and Jan H. Rosenvinge (2006), "Likert-based vs. semantic differential-based scorings of positive psychological constructs: A psychometric comparison of two versions of a scale measuring resilience," *Personality and Individual Differences*, 40 (5), 873-84.

Frijda, Nico, Peter Kuipers, and Elisabeth ter Shure (1989), "Relations Among Emotion, Appraisal, and Emotional Action Readiness," *Journal of Personality and Social Psychology*, 57 (2), 212-28.

Frost-Norton, Tammie (2005), "The Future of Mall Research: Current Trends Affecting the Future of Marketing Research in Malls," *Journal of Consumer Behaviour*, 4 (4), 293-301.

Gagliano, Kathryn Bishop and Jan Hathcote (1994), "Customer Expectations and Perceptions of Service Quality in Retail Apparel Specialty Stores," *Journal of Services Marketing*, 8 (1), 60-69.

Garlin, Francine V. and Katherine Owen (2006), "Setting the tone with the tune: A meta-analytic review of the effects of background music in retail settings," *Journal of Business Research*, 59 (6), 755-64.

Gerbing, David W. and James C. Anderson (1985), "The Effects of Sampling Error and Model Characteristics on Parameter Estimation for Maximum Likelihood Confirmatory Factor Analysis," *Multivariate Behavioral Research*, 20 (3), 255.

Gerbing, David W. and James C. Anderson (1988), "An Updated Paradigm for Scale Development: Incorporating Unidimensionality and Its Assessment," *Journal of Marketing Research*, 25 (2), 186-92.

Gilboa, Shaked and Anat Rafaeli (2003), "Store environment, emotions and approach behaviour: applying environmental aesthetics to retailing," *International Review of Retail Distribution and Consumer Research*, 13 (2), 195-211.

González, María Elisa Alén, Lorenzo Rodríguez Comesaña, and José Antonio Fraiz Brea (2007), "Assessing tourist behavioral intentions through perceived service quality and customer satisfaction," *Journal of Business Research*, 60 (2), 153-60.

Goode, Mark, M.H. and Lloyd C. Harris (2007), "Online Behavioural Intentions: An Empirical Investigations of Antecedents and Moderators," *European Journal of Marketing*, 41 (5/6), 512-36.

Gotlieb, Jerry B., Dhruv Grewal, and Stephen W. Brown (1994), "Consumer Satisfaction and Perceived Quality: Complementary or Divergent Constructs?," *Journal of Applied Psychology*, 79 (6), 875-85.

Gounaris, Spiros (2005), "An Alternative Measure for Assessing Perceived Quality of Software House Services," *Service Industries Journal*, 25 (6), 803-23.

Grace, Debra (2005), "Consumer Disposition Toward Satisfaction (CDS): Scale Development and Validation," *Journal of Marketing Theory and Practice*, 13 (2), 20-31.

Grace, Debra and Aron O'Cass (2004), "Examining Service Experiences and Post-Consumption Evaluations," *Journal of Services Marketing*, 18 (6), 450-61.

Grace, Debra and Aron O'Cass (2005), "Service Branding: Consumer Verdicts on Service Brands," *Journal of Retailing & Consumer Services*, 12 (2), 125-39.

Grapentine, Terry (2000), "Path Analysis vs. Structural Equation Modeling," *Marketing Research*, 12 (3), 13-20.

Grönroos, Christian (1984), "A Service Quality Model and its Marketing Implications," *European Journal of Marketing*, 18 (4), 36-44.

Guadagnoli, Edward and Wayne F. Velicer (1988), "Relation of Sample Size to the Stability of Component Patterns," *Psychological Bulletin*, 103 (2), 265-75.

Gupta, Sunil and Valarie Zeithaml (2006), "Customer Metrics and Their Impact on Financial Performance," *Marketing Science*, 25 (6), 718-39.

Hair, Joseph S., William C. Black, Barry J. Babin, Ralph E. Anderson, and Ronald L. Tatham (2006), *Multivariate Data Analysis* (6th ed.). New Jersey: Prentice-Hall.

Hallowell, Robert (1996), "The relationships of customer satisfaction, customer loyalty, and profitability: an empirical study," *International Journal of Service Industry Management*, 7 (4), 27-42.

Hansen, Torben (2002), "The Effect of Physical Surroundings in Usage Situations on Consumer Perception of Food Quality and on Consumer Emotions," *Journal of International Consumer Marketing*, 15 (1), 31-51.

Harrell, G.D., M.D. Hutt, and J.C. Anderson (1980), "Path Analysis of Buyer Behaviour Under Conditions of Crowding," *Journal of Marketing Research*, 17 (1), 45-51.

Harris, Lloyd C. and Chris Ezeh (2008), "Servicescape and Loyalty Intentions: An Empirical Investigation," *European Journal of Marketing*, 42 (3/4), 390-422.

Hartline, Michael D. and O.C. Ferrell (1996), "The Management of Customer-Contact Service Employees: An Empirical Investigation," *Journal of Marketing*, 60 (3), 52-70.

Hebb, Donald Olding (1949), *The Organization of Behavior*. New York: Wiley.

Hellier, Phillip K., Gus M. Geursen, Rodney A. Carr, and John A. Rickard (2003), "Customer repurchase intention: A general structural equation model," *European Journal of Marketing*, 37 (11/12), 1762-800.

Herrington, J. Duncan and Louis M. Capella (1994), "Practical application of music in service settings," *Journal of Services Marketing*, 8 (3), 50-65.

Herrington, J. Duncan and Louis M. Capella (1996), "Effects of music in service environments: a field study," *Journal of Services Marketing*, 10 (2), 26-41.

Hightower, Roscoe (2003), "Framework for Managing the Servicescape: A Sustainable Competitive Advantage," *Marketing Management Journal*, 13 (2), 84-95.

Hightower, Roscoe, Michael K. Brady, and Thomas L. Baker (2002), "Investigating the Role of the Physical Environment in Hedonic Service Consumption: An Exploratory Study of Sporting Events," *Journal of Business Research*, 55 (9), 697-707.

Hill, Donna J. (1986), "Satisfaction and Consumer Services," *Advances in Consumer Research*, 13 (1), 311-5.

Hinkin, Timothy R. (1995), "A Review of Scale Development Practices in the Study of Organizations," *Journal of Management*, 21 (5), 967-88.

Hirsch, Alan R. (1995), "Effects of Ambients Odors on Slot-Machine Usage in a Las Vegas Casino," *Psychology and Marketing*, 12 (7), 585-94.

Hirschman, Elizabeth C. and Morris B. Holbrook (1982), "Hedonic Consumption: Emerging Concepts, Methods and Propositions," *Journal of Marketing*, 46 (3), 92-101.

Hoffman, Douglas K., Scott W. Kelley, and Beth C. Chung (2003), "A CIT investigation of servicescape failures and associated recovery strategies," *Journal of Services Marketing*, 17 (4), 322-40.

Holbrook, Morris B. and R. Batra (1987), "Assessing the role of emotions as mediatiors of consumer responses to advertising," *Journal of Consumer Research*, 14 (3), 404-20.

Holbrook, Morris B. and R.A. Westwood (1989), "The Role of Emotion in Advertising Revisited: Testing a Typology of Emotional Responses," in *Cognitive and Affective Responses to Advertising*, Patricia Cafferata and Alice M. Tybout, Eds. MA, Lexington Books, 353-71.

Homburg, Christian, Nicole Koschate, and Wayne D. Hoyer (2006), "The Role of Cognition and Affect in the Formation of Customer Satisfaction: A Dynamic Perspective," *Journal of Marketing*, 70 (3), 21-31.

Hooper, Daire, Joseph Coughlan, and Michael R. Mullen (2008), "Overall Fit Indices for Structural Equation Modelling: State of Current Practice," in 7th European Conference on Research Methodology for Business and Management Studies. Regent's College, London.

Horn, John L. (1965), “A rationale and test for the number of factors in factor analysis,” *Psychometrika*, 30 (2) 179-185.

Hornik, Jacob and Shmuel Ellis (1988), "Strategies to Secure Compliance for a Mall Intercept Interview," *Public Opinion Quarterly*, 52 (3), 539-51.

Howell, Roy D. (1987), "Covariance Structure Modeling and Measurement Issues: A Note on "Interrelations Among a Channel Entity's Power Sources"," *Journal of Marketing Research*, 24 (1), 119-26.

Hu, Li Tze and Peter M. Bentler (1999), "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives," *Structural Equation Modeling*, 6 (1), 1-55.

Hui, Michael K. and John E.G. Bateson (1991), "Perceived Control and the Effects of Crowding and Consumer Choice on the Service Experience," *Journal of Consumer Research*, 18 (2), 174-84.

Hunt, H. Keith (1977), "CS/D - Overview and Future Research Directions," in *Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction*, H. Keith Hunt, Ed. Cambridge, MA: Marketing Science Institute, 455-88.

Ibanez, Vanessa Apaolaza, Patrick Hartmann, and Pilar Zorrilla Calvo (2006), "Antecedents of customer loyalty in residential energy markets: Service quality, satisfaction, trust and switching costs," *Service Industries Journal*, 26 (6), 633-50.

Izard, Carroll E. (1977), *Human Emotions*. New York: Plenum Press.

Izard, Carroll E. (1991), *The Psychology of Emotions*. New York: Plenum Press.

Izard, Carroll E. (2007), "Basic Emotions, Natural Kinds, Emotion Schemas, and a New Paradigm," *Perspectives on Psychological Science*, 2 (3), 260-80.

Izard, Carroll E., Brian P. Ackerman, Kristen M. Schoff, and Sarah E. Fine (2000), "Self-Organization of Discrete Emotions, Emotion Patterns, and Emotion Cognition Relations," in *Emotion, Development, and Self-Organization*, Mark D. Lewis and Isabella Granic, Eds. Cambridge, UK: Cambridge University Press, 15-36.

Jackson, Dennis L. (2003), "Revisiting Sample Size and Number of Parameter Estimates: Some Support for the N:q Hypothesis," *Structural Equation Modeling*, 10 (1), 128-41.

Jacob, Céline (2006), "Styles of background music and consumption in a bar: An empirical evaluation," *International Journal of Hospitality Management*, 25 (4), 716-20.

Jacoby, Jacob (1971), "A Model of Multi-Brand Loyalty," *Journal of Advertising Research*, 11 (3), 25-31.

Jacoby, Jacob and D.B. Kyner (1973), "Brand Loyalty vs. Repeat Purchasing Behaviour," *Journal of Marketing Research*, 10 (1), 1-9.

Jayanti, Rama and Anita Jackson (1991), "Service Satisfaction: An Exploratory Investigation of Three Models," *Advances in Consumer Research*, 18, 603-10.

Johnson, Allison R. and David W. Stewart (2005), "A Reappraisal of the Role of Emotion in Consumer Behavior: Traditional and Contemporary Approaches," in *Review of Marketing Research*, Naresh K. Malhotra, Ed. Vol. 1. Armonk, New Jersey: ME Sharpe, 3-33.

Johnson, Lesley, Karl J. Mayer, and Elena Champaner (2004), "Casino Atmospherics from a Customer's Perspective: A Re-Examination," *UNLV Gaming Research and Review Journal*, 8 (2), 1-10.

Jones, Michael A., David L. Mothersbaugh, and Sharon E. Beatty (2000), "Switching Barriers and Repurchase Intentions in Services," *Journal of Retailing*, 76 (2), 259-74.

Jones, Michael A., Kristy E. Reynolds, and Mark J. Arnold (2006), "Hedonic and Utilitarian Shopping Value: Investigating Differential Effects on Retail Outcomes," *Journal of Business Research*, 59 (9), 974-81.

Jones, Thomas O. and W. Earl Sasser (1995), "Why Satisfied Customers Defect," *Harvard Business Review*, 73 (6), 88-99.

Jöreskog, Karl and Dag Sörbom (1993), *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*. Chicago, IL: Scientific Software International Inc.

Jöreskog, Karl and Dag Sörbom (1996), *LISREL 8: User's Reference Guide*. Chicago, IL: Scientific Software International Inc.

Kassarjian, Harold H. and Masao Nakanishi (1967), "A Study of Selected Opinion Measurement Techniques," *Journal of Marketing Research*, 4 (2), 148-53.

Keillor, Bruce D. and Robert J. Sutton (1993), "An Investigation of Avoidance Behavior in Mall-Intercept Market Research," *Journal of Marketing Management*, 3 (1), 50-56.

Kellaris, James J. and Robert J. Kent (1992), "The Influence of Music on Consumers' Temporal Perceptions: Does Time Fly When You're Having Fun?," *Journal of Consumer Psychology*, 1 (4), 365-76.

Kellaris, James J. and Robert J. Kent (1993), "An Exploratory Investigation of Responses Elicited by Music Varying in Tempo, Tonality, and Texture," *Journal of Consumer Psychology*, 2 (4), 381-401.

Kelley, Harold H. (1967), "Attribution Theory in Social Psychology," in *Nebraska Symposium on Motivation*, David Levine, Ed. Lincoln: University of Nebraska Press, 192-242.

Kelley, Harold H. and John L. Michela (1980), "Attribution Theory and Research," *Annual Review of Psychology*, 31, 457-501.

Kenny, David A. and D. Betsy McCoach (2003), "Effect of the Number of Variables on Measures of Fit in Structural Equation Modeling," *Structural Equation Modeling*, 10 (3), 333-51.

Kline, Rex B. (2005), *Principles and Practice of Structural Equation Modeling* (2nd ed.). New York: The Guilford Press.

Knowles, Patricia A., Stephen J. Grove, and Gregory M. Pickett (1999), "Mood Versus Service Quality Effects on Customers' Responses to Service Organizations and Service Encounters," *Journal of Service Research*, 2 (2), 187-99.

Kotler, Philip (1973-1974), "Atmospherics as a Marketing Tool," *Journal of Retailing*, 49 (4), 48-64.

Kotzan, Jeffrey A. and Robert V. Evanson (1969), "Responsiveness of Drug Store Sales to Shelf Space Allocations," *Journal of Marketing Research*, 6 (4), 465-9.

Krampf, Robert, Linda Ueltschy, and Michael d'Amico (2003), "The Contribution of Emotion to Consumer Satisfaction in the Service Setting," *Marketing Management Journal*, 13 (1), 32-52.

Ladhari, Riadh (2007), "The Effect of Consumption Emotions on Satisfaction and Word-of-Mouth Communications," *Psychology and Marketing*, 24 (12), 1085-108.

Lariviere, Bart (2008), "Linking Perceptual and Behavioral Customer Metrics to Multiperiod Customer Profitability: A Comprehensive Service-Profit Chain Application," *Journal of Service Research*, 11 (3), 3-21.

Laroche, Michel, J. Bergeron, and C. Goutaland (2001), "A Three-Dimensional Scale of Intangibility," *Journal of Service Research*, 4 (1), 26-38.

Laroche, Michel, J. Bergeron, and C. Goutaland (2003), "How intangibility affects perceived risk: the moderating role of knowledge and involvement," *Journal of Services Marketing*, 17 (2), 122-40.

Laroche, Michel, Maria Kalamas, Soumaya Cheikhrouhou, and Adélaide Cézard (2004), "An Assessment of the Dimensionality of Should and Will Service Expectations," *Canadian Journal of Administrative Sciences*, 21 (4), 361-75.

Larsen, Randy J. and Ed Diener (1992), "Promises and Problems With the Circumplex Model of Emotion," in *Review of Personality and Social Psychology* Vol. 13. Newbury Park, CA: Sage, 25-59.

Lassar, Walfried M., Chris Manolis, and Robert Winsor, D. (2000a), "Service Quality Perspectives and Satisfaction in Private Banking," *Journal of Services Marketing*, 14 (3), 244-71.

Lassar, Walfried M., Chris Manolis, and Robert Winsor, D. (2000b), "Service Quality Perspectives and Satisfaction in Private Banking," *International Journal of Bank Marketing*, 18 (4), 181-99.

Lawley, D.N. and A.E. Maxwell (1962), "Factor Analysis as a Statistical Method," *The Statistician*, 12 (3), 209-29.

Lazarus, Richard (1982), "Thoughts on the relations between emotion and cognition," *American Psychologist*, 37 (9), 1019-24.

Lazarus, Richard (1984), "On the primacy of cognition," *American Psychologist*, 39, 124-9.

Ledesma, Ruben Daniel and Valero-Mora Pedro (2007), "Determining the Number of Factors to Retain in EFA: an easy-to-use computer program for carrying out Parallel Analysis," *Practical Assessment, Research & Evaluation*, 12 (2), 1-11.

Lee, Yong-Ki, Choong-Ki Lee, Seung-Kon Lee, and Barry J. Babin (2008), "Festivalscapes and patrons' emotions, satisfaction and loyalty," *Journal of Business Research*, 61 (1), 56-64.

Lehtinen, Uolevi and Jarmo Lehtinen (1991), "Two Approaches to Service Quality Dimensions," *Service Industries Journal*, 11 (3), 287-303.

Levitt, Theodore (1972), "Production-line Approach to Service," *Harvard Business Review*, 50 (5), 41-52.

Liden, Robert C. and John M. Maslyn (1998), "Multidimensionality of Leader-Member Exchange: An Empirical Assessment through Scale Development," *Journal of Management*, 24 (1), 43-72.

Liljander, Veronica and Tore Strandvik (1997), "Emotions in Service Satisfaction," *International Journal of Service Industry Management*, 8 (2), 148-69.

Lim, Elison Ai Ching and Swee Hoon Ang (2008), "Hedonic vs. Utilitarian Consumption: A Cross-Cultural Perspective Based on Cultural Conditioning," *Journal of Business Research*, 61 (3), 225-32.

Lip, Evelyn (1997), *What is Feng Shui?* New York: Academy Editions.

Lucas, Anthony F. (2003), "The Determinants and Effects of Slot Servicescape Satisfaction in a Las Vegas Hotel Casino," *UNLV Gaming Research and Review Journal*, 7 (1), 1-19.

MacCallum, Robert C., Michael W. Browne, and Hazuki Sugawara, M. (1996), "Power Analysis and Determination of Sample Size for Covariance Structure Modeling," *Psychological Methods*, 1 (2), 130-49.

Machleit, Karen A. and S.A. Eroglu (2000), "Describing and Measuring Emotional Response to Shopping Experience," *Journal of Business Research*, 49 (2), 101-11.

Machleit, Karen A., Sevgin A. Eroglu, and Susan Powell Mantel (2000), "Perceived Retail Crowding and Shopping Satisfaction: What Modifies This Relationship?," *Journal of Consumer Psychology*, 9 (1), 29-42.

MacInnis, Deborah J. and Chan W. Park (1991), "The Differential Role of Characteristics of Music on High and Low Involvement Consumers' Processing of Ads," *Journal of Consumer Research*, 18, 161-73.

MacKenzie, Scott B. (2001), "Opportunities for Improving Consumer Research through Latent Variable Structural Equation Modeling," *Journal of Consumer Research*, 28 (1), 159-66.

Malhotra, Naresh K. (1987), "Analyzing Marketing Research Data with Incomplete Information on the Dependent Variable," *Journal of Marketing Research*, 24 (1), 74-84.

Malhotra, Naresh K. (2006), *Marketing Research: An Applied Orientation*: Prentice-Hall.

Mano, Haim and Richard L. Oliver (1993), "Assessing the Dimensionality and Structure of the Consumption Experience: Evaluation, Feeling and Satisfaction," *Journal of Consumer Research*, 20 (3), 451-66.

Martin, David, Martin O'Neill, Susan Hubbard, and Adrian Palmer (2008), "The Role of Emotion in Explaining Consumer Satisfaction and Future Behavioural Intention," *Journal of Services Marketing*, 22 (3), 224-36.

Maslow, Abraham H. (1943), "A Theory of Human Motivation," *Psychological Review*, 50 (4), 370-96.

Mattila, Anna S. and Jochen Wirtz (2001), "Congruency of scent and music as a driver of in-store evaluations of behavior," *Journal of Retailing*, 77 (2), 273-89.

Maute, Manfred F. and Laurette Dubé (1999), "Patterns of Emotional Responses and Behavioural Consequences of Dissatisfaction," *Applied Psychology: An International Review*, 48 (3), 349-66.

Maxwell-Hudson, C. (1999), *Aromatherapy Massage* (2nd ed.). London: Dorling Kindersley.

Mayer, Karl J. and Lesley Johnson (2003), "A Customer-based Assessment of Casino Atmospherics," *UNLV Gaming Research and Review Journal*, 7 (1), 21-31.

McDonald, Roderick P. and Moon-Ho Ringo Ho (2002), "Principles and Practice in Reporting Statistical Equation Analyses," *Psychological Methods*, 7 (1), 64-82.

McDougall, G.H.G. and D. Snetsinger (1990), "The intangibility of services: measurement and competitive perspectives," *Journal of Services Marketing*, 4 (4), 27-40.

McIntosh, Cameron (2006), "Rethinking fit assessment in structural equation modelling: A commentary and elaboration on Barrett (2007)," *Personality and Individual Differences*, 42 (5), 859-67.

McIntyre, Shelby H. and Sherry D. F. G. Bender (1986), "The Purchase Intercept Technique (PIT) in Comparison to Telephone and Mail Surveys," *Journal of Retailing*, 62 (4), 364.

Mehrabian, Albert and James A. Russell (1974), *An Approach to Environmental Psychology*. Cambridge, MA: MIT Press.

Mels, Gerhard, Christo Boshoff, and Deon Nel (1997), "The Dimensions of Service Quality: The Original European Perspective Revisited," *Service Industries Journal*, 17 (1), 173-89.

Menezes, Dennis and Norbert F. Elbert (1979), "Alternative Semantic Scaling Formats for Measuring Store Image: An Evaluation," *Journal of Marketing Research*, 16 (1), 80-87.

Michon, Richard, Jean-Charles Chebat, and Lou W. Turley (2005), "Mall atmospherics: the interaction effects of the mall environment on shopping behavior," *Journal of Business Research*, 58 (5), 576-83.

Miles, Jeremy and Mark Shevlin (1998), "Effects of Sample Size, Model Specification and Factor Loadings on the GFI in Confirmatory Factor Analysis," *Personality and Individual Differences*, 25 (1), 85-90.

Miles, Jeremy and Mark Shevlin (2007), "A time and a place for incremental fit indices," *Personality and Individual Differences*, 42 (5), 869-74.

Milliman, Ronald E. (1982), "Using Background Music to Affect the Behavior of Supermarket Shoppers," *Journal of Marketing*, 46 (3), 86-91.

Milliman, Ronald E. (1986), "The Influence of Background Music on the Behavior of Restaurant Patrons," *Journal of Consumer Research*, 13 (2), 286-9.

Mitchell, Deborah J., Barbara .E. Kahn, and Susan C. Knasko (1995), "There's Something in the Air: Effects of Congruent or Incongruent Ambient Odor on Consumer Decision Making," *Journal of Consumer Research*, 22 (2), 229-38.

Mittal, Banwari and Walfried M. Lassar (1998), "Why do customers switch? The dynamics of satisfaction versus loyalty," *Journal of Services Marketing*, 12 (3), 177-94.

Mittal, Vikas and Wagner A. Kamakura (2001), "Satisfaction, Repurchase Intent, and Repurchase Behavior: Investigating the Moderating Effect of Customer Characteristics," *Journal of Marketing Research*, 38 (1), 131-42.

Mittal, Vikas, William T. Ross, and Patrick M. Baldasare (1998), "The Assymmetric Impact of Negative and Positive Attribute-Level Performance on Overall Satisfaction and Repurchase Intentions," *Journal of Marketing*, 62 (1), 33-47.

Morwitz, Vicki (1997), "Why Consumers Don't Always Accurately Predict Their Own Future Behavior," *Marketing Letters*, 8 (1), 57-70.

Mudie, Peter, Angela Cottam, and Robert Raeside (2003), "An Exploratory Study of Consumption Emotion in Services," *The Service Industries Journal*, 23 (5), 84-106.

Mulaik, Stanley A., Larry R. James, Judith Van Alstine, Nathan Bennet, Sherri Lind, and C. Dean Stilwell (1989), "Evaluation of Goodness-of-Fit Indices for Structural Equation Models," *Psychological Bulletin*, 105 (3), 430-45.

Mullen, Michael R., G.R. Milne, and Patricia M. Doney (1995), "An International Marketing Application of Outlier Analysis for Structural Equations: A Methodological Note," *Journal of International Marketing*, 3 (1), 45-62.

Neo, M. and A.J. Murrell (1993), "Valenced Emotions in Satisfaction: A Look at Affect in Shopping," *Advances in Consumer Research*, 20 (1), 667-72.

Netemeyer, Richard G., William O. Bearden, and Subhash Sharma (2003), *Scaling Procedures: Issues and Applications*. California: Sage Publications.

Newman, Andrew J. (2007), "Uncovering Dimensionality in the Servicescape: Towards Legibility," *The Service Industries Journal*, 27 (1), 15-28.

Nguyen, Nina (2006), "The Collective Impact of Service Workers and Servicescape on the Corporate Image Formation," *International Journal of Hospitality Management*, 25 (2), 227-44.

Nunnally, Jum C. (1978), *Psychometric Theory* (2nd ed.). New York: McGraw-Hill.

Nunnally, Jum C. and Ira Bernstein (2006), *Psychometric Theory* (3rd ed.). New York: Mc-Graw Hill.

Nyer, Prashanth U. (1997), "A Study of the Relationships Beteen Cognitive Appraisals and Consumption Emotions," *Journal of the Academy of Marketing Science*, 25 (4), 296-304.

O'Connor, Brian P. (2000), "SPSS and SAS Programs for Determining the Number of Components Using Parallel Analysis and Velicer's MAP Test," *Behavior Research Methods, Instruments and Computers*, 32 (3), 396-402.

O'Muircheartaigh, Colm, George Gaskell, and Daniel B. Wright (1995), "Weighing Anchors: Verbal and Numeric Labels for Response Scales," *Journal of Official Statistics*, 11 (3), 295-307.

Oakes, Steve (2000), "The influence of the musicscape within the servicescape," *Journal of Services Marketing*, 14 (7), 539-56.

Okada, Erica Mina (2005), "Justification Effects on Consumer Choice of Hedonic and Utilitarian Goods," *Journal of Marketing Research*, 42 (1), 43-53.

Oliver, Richard L. (1980), "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research*, 17 (4), 460-9.

Oliver, Richard L. (1993), "Cognitive, Affective and Attribute Bases of the Satisfaction Response," *Journal of Consumer Research*, 20 (3), 418-30.

Oliver, Richard L. (1997), *Satisfaction: A Behavioral Perspective on the Consumer*. New York: Mc-Graw-Hill International Editions.

Oliver, Richard L. and Wayne S. DeSarbo (1988), "Response Determinants in Satisfaction Judgments," *Journal of Consumer Research*, 14 (4), 495-507.

Oliver, Richard L., Roland T. Rust, and Sajeev Varki (1997), "Customer delight: foundations, findings and managerial insight," *Journal of Retailing*, 73 (3), 311-36.

Oliver, Richard L. and John E. Swan (1989), "Consumer Perceptions of Interpersonal Equity and Satisfaction in Transactions: A Field Survey Approach," *Journal of Marketing*, 53 (2), 21-35.

Olsson, Ulf, Tron Foss, and Einar Breivik (2004), "Two Equivalent Discrepancy Functions for Maximum Likelihood Estimation: Do Their Test Statistics Follow a Non-Central Chi-Square Distribution under Model Misspecification?," *Sociological Methods amd Research*, 32 (4), 453-510.

Olsson, Ulf, Tron Foss, Sigurd Troye, and Roy D. Howell (2000), "The Performance of ML, GLS, and WLS Estimation in Structural Equation Modeling Under Conditions of Misspecification and Nonnormality," *Structural Equation Modeling*, 7 (4), 557-95.

Overby, Jeffrey and Eun-Ju Lee (2006), "The Effects of Utilitarian and Hedonic Online Shopping Value on Consumer Preference and Intentions," *Journal of Business Research*, 59 (10-11), 1160-66.

Pallant, Julie (2004), *SPSS Survival Manual* (2nd ed.): Open University Press.

Parasuraman, A., Leonard L. Berry, and Valerie A. Zeithaml (1991a), "Refinement and Reassessment of the SERVQUAL Scale," *Journal of Retailing*, 67 (4), 420-50.

Parasuraman, A., Leonard L. Berry, and Valerie A. Zeithaml (1991b), "Understanding Customer Expectations of Service," *Sloan Management Review*, 32 (3), 39-48.

Parasuraman, A., Valerie A. Zeithaml, and Leonard L. Berry (1985), "A Conceptual Model of Service Quality and Its Implications For Future Research," *Journal of Marketing*, 49 (4), 441-50.

Parasuraman, A., Valerie A. Zeithaml, and Leonard L. Berry (1988), "SERVQUAL: A Multiple Item Scale for Measuring Customer Perceptions of Service Quality," *Journal of Retailing*, 64 (1), 12-40.

Parasuraman, A., Valerie A. Zeithaml, and Leonard L. Berry (1994), "Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Further Research," *Journal of Marketing*, 58 (1), 111-24.

Parasuraman, A., Valerie A. Zeithaml, and Arvind Malhotra (2005), "E-S-QUAL: A Multiple-Item Scale for Assessing Electronic Service Quality," *Journal of Service Research*, 7 (3), 213-33.

Park, Hee Dun, René Dailey, and Daisy Lemus (2002), "The Use of Exploratory Factor Analysis and Principal Components Analysis in Communication Research," *Human Communication Research*, 28 (4), 562-77.

Patterson, Paul G. and Richard A. Spreng (1997), "Modelling the Relationship Between Perceived Value, Satisfaction and Repurchase Intentions in a Business-to-Business, Services Context: An Empirical Examination," *International Journal of Service Industry Management*, 8 (5), 414-34.

Pei, Chao (2008), "Exploring the Nature of the Relationships Between Service Quality and Customer Loyalty: An Attribute-Level Analysis," *Service Industries Journal*, 28 (1), 95-116.

Perrien, Jean (1997), "Repositioning Demand Artifacts in Consumer Research," *Advances in Consumer Research*, 24 (1), 267-71.

Peterson, Robert A. and William R. Wilson (1992), "Measuring Customer Satisfaction: Fact and Artifact," *Journal of the Academy of Marketing Science*, 20 (1), 61-71.

Phillips, Diane M. and Hans Baumgartner (2002), "The Role of Consumption Emotions in the Satisfaction Response," *Journal of Consumer Psychology*, 12 (3), 243-52.

Plutchik, Robert (1980), *Emotion: A Psychoevolutionary Synthesis*. New York: Harper and Row.

Podsakoff, Philip M., Scott B. MacKenzie, Lee Jeong-Yeon, and Nathan P. Podsakoff (2003), "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies," *Journal of Applied Psychology*, 88 (5), 879.

Poels, Karolien and Siegfried Dewitte (2006), "How to Capture the Heart? Reviewing 20 Years of Emotion Measurement in Advertising," *Journal of Advertising Research*, 46 (1), 18-37.

Price, Linda L., Eric J. Arnould, and Sheila L. Deibler (1995), "Consumers' emotional responses to service encounters," *International Journal of Service Industry Management*, 6 (3), 34-63.

Pullman, Madeleine E. and Michael A. Gross (2004), "Ability of Experience Design Elements to Elicit Emotions and Loyalty Behaviours," *Decision Sciences*, 35 (3), 551-77.

Reeve, John (1997), *Understanding Motivation and Emotion* (2nd ed.). Florida: Harcourt Brace.

Reeves, Carol A. and David A. Bednar (1994), "Defining Quality: Alternatives and Implications," *Academy of Management Review*, 19 (3), 419-45.

Reeves, Carol A. and David A. Bednar (1995), "Quality as Symphony," *The Cornell Hotel and Restaurant Administration Quarterly*, 36 (3), 72-79.

Reimer, Anja and Richard Kuehn (2005), "The impact of servicescape on quality perception," *European Journal of Marketing*, 39 (7/8), 785-808.

Richins, Marsha L. (1997), "Measuring Emotions in the Consumption Experience," *Journal of Consumer Research*, 24 (2), 127-46.

Rigdon, Edward E. (1998), "Structural Equation Modeling," in *Modern Methods for Business Research*, George A. Marcoulides, Ed. New Jersey: Lawrence Erlbaum Associates Inc., 251-94.

Rindskopf, David and Tedd Rose (1988), "Some Theory and Applications of Confirmatory Second-Order Factor Analysis," *Multivariate Behavioral Research*, 23 (1), 51-67.

Robinson, Stewart (1999), "Measuring Service Quality: Current Thinking and Future Requirements," *Marketing Intelligence and Planning*, 17 (1), 21-32.

Russell, James A. (1980), "A Circumplex Model of Affect," *Journal of Personality and Social Psychology*, 39 (6), 1161-78.

Russell, James A. (2003), "Core Affect and the Psychological Construction of Emotion," *Psychological Review*, 110 (1), 145-72.

Russell, James A. and Albert Mehrabian (1976), "Environmental Variables in Consumer Research," *Journal of Consumer Research*, 3 (1), 62-63.

Russell, James A. and Geraldine Pratt (1980), "A Description of the Affective Quality Attributed to Environments," *Journal of Personality and Social Psychology*, 38 (2), 311-22.

Rust, Roland T. and Richard L. Oliver (1994), "Service Quality: Insights and Managerial Implications from the Frontier," in *Service Quality: New Directions in Theory and Practice*, Roland T. Rust and Richard L. Oliver, Eds. Thousand Oaks, California: Sage Publications, 1-19.

Rust, Roland T. and Richard L. Oliver (2000), "Should We Delight the Customer?," *Journal of the Academy of Marketing Science*, 28 (1), 86-94.

Ryu, Kisang and Soo Cheong Jang (2007), "The Effect of Environmental Perceptions on Behavioural Intentions Through Emotions: The Case of Upscale Restaurants," *Journal of Hospitality and Tourism Research*, 31 (1), 56-72.

Sangster, Roberta L., Fern K. Willit, John Saltiel, Fred O. Lorenz, and Todd H. Rockwood (2001), "Evaluating Numeric Rating Scales: Replicated Results," United States Department of Labor, Bureau of Labor Statistics.

Schachter, Stanley and Jerome Singer (1962), "Cognitive, Social and Physiological Determinants of Emotional State," *Psychological Review*, 69 (5), 379-99.

Schafer, Joseph L. and John W. Graham (2002), "Missing Data: Our View of the State of the Art," *Psychological Methods*, 7 (2), 147-77.

Schafer, Joseph L. and Maren K. Olsen (1998), "Multiple Imputation for Multivariate Missing-Data Problems: A Data Analyst's Perspective," *Multivariate Behavioral Research*, 33 (4), 545-71.

Schmid, John and John M. Leiman (1957), "The Development of Hierarchical Factor Solutions," *Psychometrika*, 22 (1), 53-61.

Schwarz, Norbert, Barbel Knauper, Hans-J. Hippler, Elisabeth Noelle-Nuemann, and Leslie Clark (1991), "Rating Scales: Numeric Values May Change the Meaning of Scale Labels," *Public Opinion Quarterly*, 55 (4), 570-81.

Sherman, Elaine, Anil Mathur, and Ruth Belk Smith (1997), "Store Environment and Consumer Purchase Behavior: Mediating Role of Consumer Emotions," *Psychology and Marketing*, 14 (4), 361-78.

Shimp, Terence A., Eva M. Hyatt, and David J. Snyder (1991), "A Critical Appraisal of Demand Artifacts in Consumer Research," *Journal of Consumer Research*, 18 (3), 273-83.

Shostack, G. Lynn (1977), "Breaking Free from Product Marketing," *Journal of Marketing*, 41 (Spring), 73-80.

Smith, Anne M. (1999), "Some Problems When Adopting Churchill's Paradigm for the Development of Service Quality Measurement Scales," *Journal of Business Research*, 46 (1), 109-20.

Smith, Patricia C. and Ross Curnow (1966), "Arousal Hypothesis and the Effects of Music on Purchasing Behavior," *Journal of Applied Psychology*, 50, 255-6.

Söderlund, Magnus and Niclas Öhman (2005), "Assessing Behavior Before it Becomes Behavior: An Examination of the Role of Intentions as a Link Between Satisfaction and Repatronizing Behavior," *International Journal of Service Industry Management*, 16 (2), 169-85.

Soriano, Mirella Y., Gordan R. Foxall, and Gordon J. Pearson (2001), "Emotion and environment: A test of the behavioural perspective model in a Latin American context," *Journal of Consumer Behaviour*, 2 (2), 138-54.

Spangenberg, Eric R., Ayn E. Crowley, and P.W. Henderson (1996), "Improving Store Environment: Do Olfactory Cues Affect Evaluations and Behaviors?," *Journal of Marketing*, 60 (2), 67-80.

Spearman, Charles (1904), "The proof and measurement of association between two things," *American Journal of Psychology*, 15 (1), 72-101.

Spector, Paul E. (1992), *Summated Rating Scale Construction*. Thousand Oaks, California: Sage Publications.

Spreng, Richard A., Scott B. MacKenzie, and Richard W. Olshavsky (1996), "A Reexamination of the Determinants of Consumer Satisfaction," *Journal of Marketing*, 60 (3), 15-32.

Stafford, Marla Royne and Ellen Day (1995), "Retail Services Advertising: The Effects of Appeal, Medium and Service," *Journal of Advertising*, 24 (1), 57-71.

Steenkamp, Jan-Benedict E.M. and Hans C.M. Trijp (1991), "The use of LISREL in validating marketing constructs," *International Journal of Research in Marketing*, 8, 283-99.

Steiger, James H. (2007), "Understanding the limitations of global fit assessment in structural equation modeling," *Personality and Individual Differences*, 42 (5), 893-98.

Stephens, Nancy (1991), "Cognitive Age: A Useful Concept for Advertising?," *Journal of Advertising*, 20 (4), 37-48.

Stewart, David W. (1981), "The Application and Misapplication of Factor Analysis in Marketing Research," *Journal of Marketing Research*, 18 (1), 51-62.

Strongman, Kenneth T. (2003), *The Psychology of Emotion: From Everyday Life to Theory* (5th ed.). Chichester, West Sussex: John Wiley and Sons.

Sulek, Joanne M. and Mary R. Lind (1995), "The Impact of a Customer Service Intervention and Facility Design on Firm Performance," *Management Science*, 41 (11), 1763-74.

Summers, Teresa A. and Paulette R. Hebert (2001), "Shedding some light on store atmospherics: Influence of illumination on consumer behavior," *Journal of Business Research*, 54 (2), 145-50.

Sureshchandar, G.S., Rajendran Chandrasekharan, and R.N. Anantharaman (2002), "Determinants of Customer-Perceived Service Quality: A Confirmatory Factor Analysis Approach," *Journal of Services Marketing*, 16 (1), 2002.

Sweeney, Jillian C. and Fiona Wyber (2002), "The role of cognitions and emotions in the music-approach-avoidance behaviour relationship," *Journal of Services Marketing*, 16 (1), 51-69.

Szymanski, David M. and David H. Henard (2001), "Customer Satisfaction: A Meta-Analysis of the Empirical Evidence," *Journal of the Academy of Marketing Science*, 29 (1), 16-35.

Tabachnick, Barbara G. and Linda S. Fidell (2007), *Using Multivariate Statistics* (5th ed.). New York: Allyn and Bacon.

Tai, Susan H.C. and Agnes M.C. Fung (1997), "Application of an Environmental Psychology Model to In-Store Buying Behaviour," *International Review of Retail Distribution and Consumer Research*, 7 (4), 311-37.

Tang, Esther P.Y., Ricky Y.K. Chan, and Susan H.C. Tai (2001), "Emotional Influences of Environmental Cues on Chinese Consumers in a Leisure Service Setting," *Journal of International Consumer Marketing*, 14 (1), 67-87.

Taylor, Steven A. and Thomas L. Baker (1994), "An Assessment of the Relationship Between Service Quality and Customer Satisfaction in the Formation of Consumers' Purchase Intentions," *Journal of Retailing*, 70 (2), 163-78.

To, Pui-Lai, Chechen Lioa, and Tzu-Hua Lin (2007), "Shopping Motivations on Internet: A Study Based on Utilitarian and Hedonic Value," *Technovation*, 27 (12), 774-87.

Tourangeau, Roger, Mick P. Cooper, and Frederick Conrad (2004), "Spacing, Positioning and Order: Interpretive Heuristics for Visual Features of Survey Questions," *Public Opinion Quarterly*, 68 (3), 368-93.

Tombs, Alastair and Janet R. McColl-Kennedy (2003), "Social-servicescape conceptual model," *Marketing Theory*, 3 (4), 447-75.

Tripp, Carolyn and John T. Drea (2002), "Selecting and Promoting Service Encounter Elements in Passenger Rail Transportation," *Journal of Services Marketing*, 16 (5), 432-42.

Tse, David K. and Peter C. Wilton (1988), "Models of Consumer Satisfaction Formation: An Extension," *Journal of Marketing Research*, 25 (2), 204-12.

Turley, Lou W. and Ronald E. Milliman (2000), "Atmospheric Effects on Shopping Behavior: A Review of the Experimental Evidence," *Journal of Business Research*, 49 (2), 193-211.

Ullman, Jodie B. (2007), "Structural Equation Modeling," in *Using Multivariate Statistics*, Barbara G. Tabachnick and Linda S. Fidell, Eds. 5th ed. Boston: Allyn and Bacon, 676-780.

van Dolen, Willemijn, Ko de Ruyter, and Jos Lemmink (2004), "An Empirical Assessment of the Influence of Customer Emotions and Contact Employee Performance on Encounter and Relationship Satisfaction," *Journal of Business Research*, 57 (4), 437-44.

Vanhamme, Joelle (2000), "The Link between Surprise and Satisfaction: An Exploratory Research on how to best Measure Surprise," *Journal of Marketing Management*, 16 (6), 565-82.

Vargo, Stephen L. and Robert F. Lusch (2004), "The Four Service Marketing Myths: Remnants of a Goods-Based, Manufacturing Model," *Journal of Service Research*, 6 (4), 324-35.

Velicer, Wayne F. (1976), “Determining the number of components from the matrix of partial correlations,” *Psychometrika*, 41 (3), 321-327.

Velicer, Wayne F. and Douglas N. Jackson (1990), "Component Analysis versus Common Factor Analysis: Some Issues in Selecting an Appropriate Procedure," *Multivariate Behavioral Research*, 25 (1), 97-114.

Voss, Glenn B., A. Parasuraman, and Dhruv Grewal (1998), "The Roles of Price, Performance, and Expectations in Determining Satisfaction in Service Exchanges," *Journal of Marketing*, 62 (4), 46-61.

Voss, Kevin E., Eric R. Spangenberg, and Bianca Grohmann (2003), "Measuring the Hedonic and Utilitarian Dimensions of Consumer Attitude," *Journal of Marketing Research*, 40 (3), 310-20.

Wakefield, Kirk L. and Julie Baker (1998), "Excitement at the Mall: Determinants and Effects on Shopping Response," *Journal of Retailing*, 74 (4), 515-39.

Wakefield, Kirk L. and Jeffrey G. Blodgett (1994), "The Importance of Servicescapes in Leisure Service Settings," *Journal of Services Marketing*, 8 (3), 66-76.

Wakefield, Kirk L. and Jeffrey G. Blodgett (1996), "The effect of servicescapes on customers' behavioural intentions in leisure service settings," *Journal of Services Marketing*, 10 (6), 45-61.

Wakefield, Kirk L. and Jeffrey G. Blodgett (1999), "Customer Response to Intangible and Tangible Service Factors," *Psychology and Marketing*, 16 (1), 51-68.

Walker, James L. (1995), "Service encounter satisfaction: conceptualised," *Journal of Services Marketing*, 9 (1), 5-14.

Wall, Eileen A. and Leonard L. Berry (2007), "The Combined Effects of the Physical Environment and Employee Behaviour on Customer Perception of Restaurant Service Quality," *Cornell Hotel and Restaurant Administration Quarterly*, 48 (1), 59-69.

Wang, Jeff and Melanie Wallendorf (2006), "Materialism, Status Signalling, and Product Satisfaction," *Journal of the Academy of Marketing Science*, 34 (4), 494-505.

Ward, Philippa, Barry J. Davies, and Dion Kooijman (2003), "Ambient Smell and the Retail Environment: Relating Olfaction Research to Consumer Behaviour," *Journal of Business and Management*, 9 (3), 289-302.

Watson, David, Lee Anna Clark, and Auke Tellegen (1988), "Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scale," *Journal of Personality and Social Psychology*, 54 (6), 1063-70.

Watson, David and Auke Tellegen (1985), "Toward a Consensual Structure of Mood," *Psychological Bulletin*, 98 (2), 219-35.

Watson, Lisa and Mark T. Spence (2007), "Causes and Consequences of Emotions on Consumer Behaviour: A Review and Integrative Cognitive Appraisal Theory," *European Journal of Marketing*, 41 (5/6), 487-511.

Weiner, Bernard (1985), "An attributional theory of achievement motivation and emotion," *Psychological Review*, 92 (4), 548-73.

Weiner, Bernard (1986), *An Attributional Theory of Motivation and Emotion*. Boulder: Westview.

Westbrook, Robert A. (1980), "A Rating Scale For Measuring Product/Service Satisfaction," *Journal of Marketing*, 44 (4), 68-72.

Westbrook, Robert A. (1987), "Product/Consumption-Based Affective Responses and Postpurchase Processes," *Journal of Marketing Research*, 24 (3), 258-70.

Westbrook, Robert A. and Richard L. Oliver (1981), "Developing Better Measures of Consumer Satisfaction: Some Preliminary Results," *Advances in Consumer Research*, 8 (1), 94-99.

Westbrook, Robert A. and Richard L. Oliver (1991), "The Dimensionality of Consumption Emotion Patterns and Consumer Satisfaction," *Journal of Consumer Research*, 18 (1), 84-91.

Westbrook, Robert A. and Michael D. Reilly (1983), "Value-Percept Disparity: An Alternative to the Disconfirmation of Expectations Theory of Consumer Satisfaction," *Advances in Consumer Research*, 10, 256-61.

Wetzels, Martin G.M. (1998), "*Service Quality in Customer-Employee Relationship*," PhD thesis, University of Maastricht.

Wheaton, Blair, Bengt Muthen, Duane Alwin, F., and Gene Summers (1977), "Assessing Reliability and Stability in Panel Models," *Sociological Methodology*, 8 (1), 84-136.

White, Christopher (2006), "Towards an Understanding of the Relationship between Mood, Emotions, Service Quality and Customer Loyalty Intentions," *The Service Industries Journal*, 26 (8), 837-47.

White, Christopher and Yi-Ting Yu (2005), "Satisfaction emotions and consumer behavioural intentions," *Journal of Services Marketing*, 19 (6), 411-20.

Winer, Russell, S. (1999), "Experimentation in the 21st Century: The Importance of External Validity," *Journal of the Academy of Marketing Science*, 27 (3), 349-58.

Wirtz, Jochen and John E.G. Bateson (1999), "Consumer Satisfaction with Services: Integrating the Environment Perspective in Services Marketing into the Traditional Disconfirmation Paradigm," *Journal of Business Research*, 44 (1), 55-66.

Wiseman, Frederick, Marianne Schafer, and Richard Schafer (1983), "An Experimental Test of the Effects of a Monetary Incentive on Cooperation Rates and Data Collection Costs in Central-Location Interviewing," *Journal of Marketing Research*, 20 (4), 439-42.

Wolff, Hans-Georg and Katja Preising (2005), "Exploring item and higher order factor structure with the Schmid-Leiman solution: Syntax codes for SPSS and SAS," *Behavior Research Methods*, 37 (1), 48-58.

Wong, Amy (2004), "The role of emotional satisfaction in service encounters," *Managing Service Quality*, 14 (5), 365-76.

Wong, Amy and Amrik Sohal (2003), "Service quality and customer loyalty perspectives on two levels of retail relationships," *Journal of Services Marketing*, 17 (5), 495-513.

Wood, Stacy L. and Page Moreau (2006), "From Fear to Loathing? How Emotion Influences the Evaluation and Early Use of Innovations," *Journal of Marketing*, 70 (3), 44-57.

Woodruff, Robert B., Ernest R. Cadotte, and Roger L. Jenkins (1983), "Modelling Consumer Satisfaction Processing Using Experience-Based Norms," *Journal of Marketing Research*, 20 (3), 296-304.

Woods, Walter A. (1960), "Psychological Dimensions of Consumer Decision," *Journal of Marketing*, 24 (3), 15-19.

Woodside, Arch G., Lisa L. Frey, and Robert Timothy Daly (1989), "Linking Service Quality, Customer Satisfaction, and Behavioral Intention," *Journal of Health Care Marketing*, 9 (4), 5-17.

Worthington, Roger L. and Tiffany A. Whittaker (2006), "Scale Development Research: A Content Analysis and Recommendations of Best Practices," *The Counselling Psychologist*, 34 (6), 806-38.

Yalch, Richard F. and Eric R. Spangenberg (1990), "Effects of Store Music on Shopping Behaviour," *Journal of Services Marketing*, 4 (1), 31-9.

Yalch, Richard F. and Eric R. Spangenberg (1993), "Using Store Music for Retail Zoning: A Field Experiment," *Advances in Consumer Research*, 20 (1), 632-6.

Yalch, Richard F. and Eric R. Spangenberg (2000), "The Effects of Music in a Retail Setting on Real and Perceived Shopping Times," *Journal of Business Research*, 49 (2), 139-47.

Yani-de-Soriano, M. Mirella and Gordan R. Foxall (2006), "The emotional power of place: The fall and rise of dominance in retail research," *Journal of Retailing and Consumer Services*, 13 (6), 403-16.

Yi, Youjae (1990), "A Critical Review of Consumer Satisfaction," in *Review of Marketing*, Valerie A. Zeithaml, Ed. Chicago, IL: American Marketing Association.

Yoo, Changjo, Jonghee Park, and Deborah J. MacInnis (1998), "Effects of Store Characterisics and In-Store Emotional Experiences on Store Attitude," *Journal of Business Research*, 42 (3), 253-63.

Yu, Julie and Harris Cooper (1983), "A Quantitative Review of Research Design Effects on Response Rates to Questionnaires," *Journal of Marketing Research*, 20 (1), 36-44.

Yu, Yi-Ting and Alison Dean (2001), "The contribution of emotional satisfaction to consumer loyalty," *International Journal of Service Industry Management*, 12 (3), 234-50.

Yuan, Ke Hai (2005), "Fit Indices Versus Test Statistics," *Multivariate Behavioral Research*, 40 (1), 115-48.

Zajonc, Robert B. (1965), "Social Facilitation," *Science*, 149 (3681), 269-74.

Zeithaml, Valerie A., Leonard L. Berry, and A. Parasuraman (1993), "The Nature and Determinants of Customer Expectations of Service," *Journal of the Academy of Marketing Science*, 21 (1), 1-12.

Zeithaml, Valerie A., Leonard L. Berry, and A. Parasuraman (1996), "The Behavioral Consequences of Service Quality," *Journal of Marketing*, 60 (April), 31-46.

Zeithaml, Valerie A. and Mary Jo Bitner (2003), *Services Marketing: Integrating Customer Focus across the Firm* (3rd ed.). New York: McGraw-Hill.

Zeithaml, Valerie A., A. Parasuraman, and Leonard L. Berry (1985), "Problems and Strategies in Service Marketing," *Journal of Marketing*, 49 (2), 33-46.

Zeitlin, David M. and Richard A. Westwood (1986), "Measuring Emotional Response," *Journal of Advertising Research*, 26 (5), 34-44.

Zwick, William, R. and Wayne F. Velicer (1986), "Comparison of Five Rules for Determining the Number of Components to Retain," *Psychological Bulletin*, 99 (3), 432-42.

APPENDIX 1A
Servicescape Dimensions

Dimensions Appearing in the Servicescape Literature (Theoretical and Empirical Research)

Author(s)	No. of Dimensions	Ambience	Design	Signage	Layout	Equip	Music	Hygiene	Social	Other	Context
Lee et al (2008)	7								x	Convenience Information Program Content Facility Souvenirs Food	Festivals
Ezeh and Harris (2007)	3	x	x						x		Theoretical Paper
Ryu and Jang (2007)	6	x[1]	x		x	x			x	Lighting	High-end Restaurant
El Sayed et al (2003)	5		x				x			Crowdedness Location Lighting	Shopping Mall
Hightower (2003)	3	x	x						x		Theoretical Paper
Baker et al (2002)	3		x				x		x		
Turley and Chebat (2002)	5		x[2]		x[a]				x	POP Decorations Exterior General Interior	Theoretical Paper
Wakefield and Blodgett (1999)	3	x	x			x					Hockey games Recreational Center
Wakefield and Baker (1998)	4	x	x		x					Variety	Mall
Wakefield and Blodgett (1996)	5		x		x	x		x		Seating	College Football Baseball Casinos

[1] Included music
[2] These environmental stimuli were grouped together to form one environmental dimension

Author(s)	No. of Dimensions	Ambience	Design	Signage	Layout	Equip	Music	Hygiene	Social	Other	Context
Baker et al (1994)	3	x	x						x		Retail (Experiment)
Bitner (1992)	3	x			x					Signs, Symbols and Artifacts	Theoretical Paper
Baker (1987)	3										

APPENDIX 2A
Servicescape Scale Items

Equipment

Reimer and Kuehn 2005
The equipment is modern-looking

Wakefield and Blodgett 1999
The electronic equipment is of high quality
The electronic equipment is excellent
The equipment is modern-looking

Physical Facilities

Reimer and Kuehn 2005
The physical facilities are visually appealing

Wakefield and Blodgett 1996
This is an attractive faciilty

Grace and O'Cass 2004
Facilities are attractive
Up-to-date facilities

Baker et al 2002
Attractive facilities

Employees
Reimer and Kuehn 2005
The employees are neat-appearing

Lucas 2003
The employees are neat-appearing
Employees are polite and courteous
Employees display warmth in their behaviour
Employees are happy to serve customers
Employees never seem bothered with customers requests
Employees respond quickly in this environment
Prompt service is important to employees here

Wakefield and Blodgett 1999
The employees are neat-appearing

Grace and O'Cass 2004
The employees are neat-appearing

Hightower et al 2002
The employees are neat and well dressed
The employees are helpful*
The employees are friendly*
There are enough employees to service

Brochure and Signs
Reimer and Kuehn
2005
Brochures & other comm materials are visually appealing

Hightower et al 2002
The signs used are helpful to me

Lucas 2003
The signs in this environment provide adequate direction

Temperature
Reimer and Kuehn
2005
Room temperature is pleasant

Lucas 2003
The temperature in this environment is comfortable

Johnson 2002
The temperature in this environment is comfortable

Wakefield and Blodgett 1999
Temperature is comfortable

Hightower et al 2002
The temperature is pleasant*

Music
Reimer and Kuehn
2005
Background music is pleasant

Lucas 2003
The background music makes this envir. more enjoyable

Hightower et al 2002
The background music is appropriate*

Baker et al 2002
Pleasant music
Appropriate music
Bothersome music

Architecture
Reimer and Kuehn
2005
The buildings architecture is visually appealing

Wakefield and Blodgett 1999
Outside appearance is attractive

customers

Baker et al 2002
Well-dressed employees
Friendly employees
Helpful employees

Hightower et al 2002
The architecture is attractive*

Restrooms and Exterior (other than architecture)
The restrooms are appropriately designed*
The parking lot has more than enough space

General

Lucas 2003
Overall the feeling I get from this environment is satisfaction
The overall feeling I get from this environment puts me in a good mood

Wakefield and Blodgett 1994
How accurately do these words describe the quality of this stadium?
How accurately do these words describe your satisfaction with this stadium?

Wakefield and Blodgett 1996

The overall quality of this facility is	terrible-great
	worse than expected-better than expected
	not at all what it should be-what it should be
The overall feeling that I get from this facility is	dissatisfaction-satisfaction
	Puts me in a bad mood-puts me in a good mood

Hightower et al 2002
Generally, I am impressed with the environment when I go to x
In general, the physical environment pleases me
X's physical environment is one of the best in the industry*
I think that X's physical environment is
superior*

Sherman et al 1997
good-bad
negative-positive

Interior Design

Reimer and Kuehn 2005
The interior design is visually appealing

Lucas 2003
The interior décor of this enivoronment is attractive

Wakefield and Blodgett 1999
The interior design is attractive

Hightower et al 2002
The style of the interior accessories are fashionable

Sherman et al 1997
Impressive interior

Odour

Colours

Reimer and Kuehn 2005
The colours of the physical facilties and the interior are pleasant

Lucas 2003
The use of colour adds excitement

Wakefield and Blodgett 1996
The interior wall and colour schemes are attractive

Hightower et al 2002
The colour scheme is attractive

Baker et al 2002
Pleasing colour schemes

Lighting

Reimer and Kuehn 2005
The odour is pleasant

Hightower et al 2002
The environment has a pleasant smell

Sherman et al 1997
Pleasant smelling

Noise levels
Reimer and Kuehn 2005

The noise level is acceptable

Hightower et al 2002

The background noise is acceptable*

Cleanliness
Reimer and Kuehn 2005
The physical facilities are clean

Lucas 2003
This environment is clean

Wakefield and Blodgett 1996
This facility maintains clean restrooms
This facility maintains clean food service areas
This facility maintains clean walkways and exits
Overall, this facility is kept clean

Wakefield and Blodgett 1999
Kept clean

Hightower et al 2002
X is clean

Sherman et al 1997

Reimer and Kuehn 2005
The lighting is comfortable

Lucas 2003
Overall lighting level is appropriate

Hightower et al 2002
The lighting is excellent at X

Space
Lucas 2003
It is easy to walk around this environment (plus 3 more)

Hightower et al 2002
Has more than enough space for me to be comfortable

Sherman et al 1997
Large - small
Roomy - cramped

Layout
Lucas 2003
It is easy to walk around and see what you're looking for

Wakefield and Blodgett 1996
Overall this facilty's layout makes it easy to find where you're going (plus 3)

Wakefield and Blodgett 1999
Layout makes it easy to get around

Hightower et al 2002
The interior layout is pleasing

Sherman et al 1997
Well-organised layout

APPENDIX 2B
Two Factor Solution of Servicescape Scale Items

Pattern Matrix

	Factors and % of Variance	
	35.7%	6.55%
	1	2
The colour schemes were pleasant	**.867**	-.066
The interior layout was pleasing	**.819**	.029
I found the interior design visually appealing	**.782**	-.132
The style of the interior was fashionable	**.748**	-.180
I found the physical facilities comfortable	**.745**	.068
The waiting room was attractive	**.706**	.048
The surgery had a pleasant smell	**.641**	-.014
I found the lighting comfortable	**.536**	.173
The architecture of the surgery was attractive	**.505**	.029
The waiting room was comfortable	**.496**	.227
The background music was appropriate	**.414**	.153
The background music was pleasant	**.363**	.107
The electronic equipment was excellent	.020	**.797**
I found the dental surgery very clean	-.106	**.762**
The electronic equipment was of high quality	-.049	**.743**
The employees were neat in appearance	-.099	**.668**
The bathroom facilities were clean and tidy	.077	**.625**
The flooring was appropriate	.085	**.610**
The dental equipment was modern looking	-.043	**.593**
The materials used were of high quality	.219	**.539**
I found the employees at the dental surgery friendly	-.034	**.422**
I found the employees very helpful	.066	**.387**
The seating in the surgery was comfortable	.308	**.369**
I found my way around the surgery quite easily	.156	**.325**

Extraction Method: Principal Axis Factoring.
Rotation Method: Promax with Kaiser Normalization.
a Rotation converged in 3 iterations.

APPENDIX 2C
Scale Development Questionnaire

1. Consumers' Perceptions of Dental Servicescapes

As part of my PhD research, I am conducting a survey about peoples' perceptions of both the interior and exterior of dental services. This research is being conducted solely for academic purposes, and not on behalf of any particular dental service. I am hoping that you will assist me by setting aside some of your valuable time to complete this questionnaire, it should only take ten minutes at the most to fill out.

The survey consists of three sections. The first section contains questions regarding your opinions of both physical and atmospheric aspects of the service. The second section relates to the particular emotions that you may experience while at the dentist. The final section includes some, very brief, demographic information.

Thank you again for taking the time out to complete this questionnaire!

Daire Hooper
Doctoral Researcher
Faculty of Business
Dublin Institute of Technology

2. SECTION A: The Servicescape

Please visualise the last time you went to the dentist. Recall what it looked like, smelled like and felt like and use your imagination to envision your impression of the experience. As you probably know, a big part of going to the dentist is the impression the physical facilities have on you. Please use your past experience of going to the dentist to fill out this first section of the questionnaire.
*
1.

	Strongly Disagree				**Strongly Agree**
I was generally impressed by the environment at the dental surgery	☐	☐	☐	☐	☐
The physical facilities were visually unappealing	☐	☐	☐	☐	☐
The employees were neat in appearance	☐	☐	☐	☐	☐
I found the dental surgery very clean	☐	☐	☐	☐	☐
The brochures and other communications material were visually appealing	☐	☐	☐	☐	☐
The odour was unpleasant	☐	☐	☐	☐	☐
I found the lighting excellent	☐	☐	☐	☐	☐
The background music was pleasant	☐	☐	☐	☐	☐

	Strongly Disagree				Strongly Agree
I found the interior design visually appealing	☐	☐	☐	☐	☐
There were appropriate reading materials to read in the waiting room	☐	☐	☐	☐	☐
The dental equipment was modern looking	☐	☐	☐	☐	☐
Overall, I would rate the dental surgery's physical environment very highly	☐	☐	☐	☐	☐

Section A: The Servicescape

	Strongly Disagree	2	3	4	Strongly Agree
I found the signs used (e.g. bathroom, enter, exit,) helpful to me	☐	☐	☐	☐	☐
The electronic equipment was of high quality	☐	☐	☐	☐	☐
I found the music at the dental surgery bothersome	☐	☐	☐	☐	☐
Overall, the décor at the dentist's was pleasing	☐	☐	☐	☐	☐
The exterior of the dentist's was appropriate	☐	☐	☐	☐	☐
The surgery was very unhygienic	☐	☐	☐	☐	☐
I found the dental surgery easy to find	☐	☐	☐	☐	☐
The seating in the surgery was comfortable	☐	☐	☐	☐	☐
The flooring was appropriate	☐	☐	☐	☐	☐
The electronic equipment was excellent	☐	☐	☐	☐	☐
The waiting room was attractive	☐	☐	☐	☐	☐
I found my way around the surgery quite easily	☐	☐	☐	☐	☐
I found the employees at the dental surgery friendly	☐	☐	☐	☐	☐

*

	Strongly Disagree	2	3	4	Strongly Agree
I found the physical facilities of this dentist surgery to be superior to others	☐	☐	☐	☐	☐
The dental equipment was of poor quality	☐	☐	☐	☐	☐
The background music was appropriate	☐	☐	☐	☐	☐

The colour schemes were pleasant	☐	☐	☐	☐	☐
I found the lighting comfortable	☐	☐	☐	☐	☐
The interior layout was pleasing	☐	☐	☐	☐	☐
The bathroom facilities were clean and tidy	☐	☐	☐	☐	☐
In general, the physical environment was pleasing	☐	☐	☐	☐	☐
The waiting room was comfortable	☐	☐	☐	☐	☐
I found the employees at the dental surgery unhelpful	☐	☐	☐	☐	☐
The room temperature was pleasant	☐	☐	☐	☐	☐
The surgery had more than enough space to be comfortable	☐	☐	☐	☐	☐

5. Section A: The Servicescape

	Strongly Disagree	**2**	**3**	**4**	**Strongly Agree**
The parking facilities were satisfactory	☐	☐	☐	☐	☐
The surgery had a poorly organised layout	☐	☐	☐	☐	☐
The architecture of the surgery was attractive	☐	☐	☐	☐	☐
I found the physical facilities comfortable	☐	☐	☐	☐	☐
The materials used were of high quality	☐	☐	☐	☐	☐
I found the physical facilities very unclean	☐	☐	☐	☐	☐
The other patients at the surgery were neat and well dressed	☐	☐	☐	☐	☐
The style of the interior was fashionable	☐	☐	☐	☐	☐
The surgery had a pleasant smell	☐	☐	☐	☐	☐
I found the employees very helpful	☐	☐	☐	☐	☐
Overall, the atmosphere at the surgery was pleasant	☐	☐	☐	☐	☐

8. SECTION C: Demographic Information

7. Gender

☐ Male

☐ Female

8. Please state the purpose of your last visit to the dentist:

*

9. Please state when your last visit to the dentist was:

- Less than two weeks ago
- Between two weeks and a month ago
- Between one and two months ago
- Between two and four months ago
- Between four and six months ago
- Between six months to a year ago
- Between one and two years ago
- Between two and three years ago
- Between three and four years ago
- More than four years ago

Just hit 'DONE' below and you are finished!!

Thank you for the time you have spared to fill out this questionnaire. Your input is thoroughly valued!

Kind Regards,
Daire Hooper.

Ms. Daire Hooper
Doctoral Researcher
Faculty of Business
Room 3048.1
Dublin Institute of Technology
Aungier Street
Dublin 2

APPENDIX 2D
Skewness and Kurtosis

	Skewness	Kurtosis
I was generally impressed by the environment at the dental surgery	-.426	-.377
The physical facilities were visually unappealing	-.262	-.949
The employees were neat in appearance	-1.390	1.892
I found the dental surgery very clean	-1.936	3.522
The brochures and other communications material were visually appealing	.142	-.399
The odour was unpleasant	-.277	-.883
I found the lighting excellent	-.472	-.207
The background music was pleasant	.129	-.847
I found the interior design visually appealing	.180	-.741
There were appropriate reading materials to read in the waiting room	-.087	-1.067
The dental equipment was modern looking	-1.310	1.600
Overall, I would rate the dental surgery's physical environment very highly	-.610	-.275
I found the signs used (e.g. bathroom, enter, exit,) helpful to me	-.299	-.239
The electronic equipment was of high quality	-.614	.040
I found the music at the dental surgery bothersome	-.331	-.550
Overall, the decor at the dentist's was pleasing	-.171	-.699
The exterior of the dentist's was appropriate	-.244	-.560
The surgery was very unhygienic	-2.394	5.414
I found the dental surgery easy to find	-.602	-.602
The seating in the surgery was comfortable	-.376	-.351
The flooring was appropriate	-.327	.014
The electronic equipment was excellent	-.565	.114
The waiting room was attractive	.160	-.587
I found my way around the surgery quite easily	-.567	-.266
I found the employees at the dental surgery friendly	-1.124	.999
I found the physical facilities of this dentist surgery to be superior to others	.172	.005
The dental equipment was of poor quality	-1.374	1.737
The background music was appropriate	-.048	-.280
The colour schemes were pleasant	-.172	-.178
I found the lighting comfortable	-.489	.097
The interior layout was pleasing	-.196	-.318
The bathroom facilities were clean and tidy	-.346	-.176
In general, the physical environment was pleasing	-.206	-.479
The waiting room was comfortable	-.262	-.396
I found the employees at the dental surgery unhelpful	-1.546	1.445
The room temperature was pleasant	-.285	-.325
The surgery had more than enough space to be comfortable	-.236	-.556
The parking facilities were satisfactory	.192	-.928
The surgery had a poorly organised layout	-.488	-.128
The architecture of the surgery was attractive	.070	-.458
I found the physical facilities comfortable	-.145	-.181
The materials used were of high quality	-.712	.109

I found the physical facilities very unclean	-1.872	3.073
The other patients at the surgery were neat and well dressed	.411	.082
The style of the interior was fashionable	.062	-.584
The surgery had a pleasant smell	-.092	-.503
I found the employees very helpful	-1.255	1.695
Overall, the atmosphere at the surgery was pleasant	-.612	.100

APPENDIX 2E
Two Factor Solution

Total Variance Explained

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings(a)
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	9.104	37.934	37.934	8.579	35.747	35.747	7.624
2	2.105	8.769	46.703	1.573	6.556	42.303	7.048
3	1.592	6.632	53.334				
4	1.369	5.704	59.038				
5	1.345	5.604	64.642				
6	1.071	4.463	69.105				
7	.787	3.281	72.386				
8	.739	3.078	75.465				
9	.710	2.956	78.421				
10	.657	2.738	81.159				
11	.557	2.320	83.479				
12	.506	2.110	85.589				
13	.470	1.957	87.546				
14	.444	1.850	89.397				
15	.366	1.527	90.923				
16	.332	1.384	92.308				
17	.310	1.293	93.601				
18	.295	1.230	94.832				
19	.259	1.081	95.913				
20	.226	.944	96.856				
21	.224	.933	97.789				
22	.191	.796	98.585				
23	.178	.740	99.325				
24	.162	.675	100.000				

Extraction Method: Principal Axis Factoring.

a When factors are correlated, sums of squared loadings cannot be added to obtain a total variance.

Pattern Matrix

	Factor	
	1	2
The colour schemes were pleasant	.867	-.066
The interior layout was pleasing	.819	.029
I found the interior design visually appealing	.782	-.132
The style of the interior was fashionable	.748	-.180
I found the physical facilities comfortable	.745	.068
The waiting room was attractive	.706	.048
The surgery had a pleasant smell	.641	-.014
I found the lighting comfortable	.536	.173
The architecture of the surgery was attractive	.505	.029
The waiting room was comfortable	.496	.227
The background music was appropriate	.414	.153
The background music was pleasant	.363	.107
The electronic equipment was excellent	.020	.797
I found the dental surgery very clean	-.106	.762
The electronic equipment was of high quality	-.049	.743
The employees were neat in appearance	-.099	.668
The bathroom facilities were clean and tidy	.077	.625
The flooring was appropriate	.085	.610
The dental equipment was modern looking	-.043	.593
The materials used were of high quality	.219	.539
I found the employees at the dental surgery friendly	-.034	.422
I found the employees very helpful	.066	.387
The seating in the surgery was comfortable	.308	.369
I found my way around the surgery quite easily	.156	.325

Extraction Method: Principal Axis Factoring.
Rotation Method: Promax with Kaiser Normalization.
a Rotation converged in 3 iterations.

APPENDIX 2F
Results from Parallel Analysis and Velicer's MAP

Parallel Analysis

Component Number	Actual eigenvalue from Principal Axis Factoring	Criterion Value from Parallel Analysis	Decision
1	9.104	1.79	Accept
2	2.105	1.64	Accept
3	1.592	1.55	Accept
4	1.369	1.46	Reject
5	1.345	1.39	Reject
6	1.071	1.32	Reject

Velicer's MAP

```
Velicer's Minimum Average Partial (MAP) Test:

Eigenvalues
     9.099831
     2.106599
     1.595777
     1.369434
     1.350271
     1.072775
      .787776
      .732945
      .701782
      .651718
      .565533
      .516808
      .470273
      .443359
      .364655
      .332448
      .304159
      .290602
      .263762
      .225872
      .221884
      .189958
      .180724
      .161057

Velicer's Average Squared Correlations
      .000000      .134792
     1.000000      .028568
     2.000000      .023173
     3.000000      .024594
     4.000000      .026762
     5.000000      .026817
     6.000000      .025688
```

```
     7.000000      .028910
     8.000000      .031517
     9.000000      .037349
    10.000000      .040891
    11.000000      .047422
    12.000000      .052701
    13.000000      .063015
    14.000000      .075013
    15.000000      .088001
    16.000000      .103119
    17.000000      .120695
    18.000000      .140660
    19.000000      .185793
    20.000000      .250156
    21.000000      .322060
    22.000000      .488771
    23.000000     1.000000

The smallest average squared correlation is
      .023173

The number of components is
  2

------ END MATRIX -----
```

APPENDIX 2G
Hierarchical Factor Analysis

Correlation Matrix[a]

a. Determinant = .200

Communalities

	Initial	Extraction
I1	.475	.592
I2	.400	.486
I3	.479	.608
I4	.326	.390
I5	.121	.130
I6	.187	.185

Extraction Method: Principal Axis Factoring.

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Factor	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.916	48.597	48.597	2.391	39.850	39.850
2	.930	15.495	64.092			
3	.739	12.313	76.405			
4	.579	9.650	86.056			
5	.466	7.774	93.829			
6	.370	6.171	100.000			

Extraction Method: Principal Axis Factoring.

Factor Matrix(a)

	Factor
	1
I1	.769
I2	.697
I3	.780
I4	.625
I5	.360
I6	.430

Extraction Method: Principal Axis Factoring.
a 1 factors extracted. 6 iterations required.

Parallel Analysis for Hierarchical Solution

```
Run MATRIX procedure:

PARALLEL ANALYSIS:

Principal Components

Specifications for this Run:
Ncases     204
Nvars        6
Ndatsets  1000
Percent     99

Random Data Eigenvalues
         Root        Means     Prcntyle
     1.000000     1.236394     1.373473
     2.000000     1.120714     1.221053
     3.000000     1.034608     1.108248
     4.000000      .952458     1.019343
     5.000000      .875317      .959398
     6.000000      .780510      .877816

------ END MATRIX -----
```

APPENDIX 2H
Schmid-Leiman Solution

Run MATRIX procedure:

F1

	factor1	factor2	factor3	factor4	factor5	factor6
v1	.836	.123	-.301	-.118	-.006	.021
v2	.780	-.086	.098	.112	-.007	-.006
v3	.772	.102	.089	-.123	.113	-.116
v4	.723	-.184	.138	.113	-.081	.168
v5	.707	.029	-.197	.122	-.016	.020
v6	.639	-.001	-.069	-.066	-.026	.290
v7	.623	-.059	.108	.088	-.008	.074
v8	.598	.150	.032	-.140	.034	-.214
v9	.474	.030	.372	.034	-.128	-.130
v10	.414	-.021	.341	.046	.068	-.056
v11	-.061	.739	.244	.039	-.059	.066
v12	-.037	.722	.010	.157	-.076	.060
v13	.193	.643	.112	-.133	.036	.022
v14	.054	.488	-.219	.402	-.089	.054
v15	.135	.479	.049	.082	.177	-.047
v16	.024	.052	.761	-.074	-.055	.056
v17	-.150	.122	.721	.197	-.093	.013
v18	-.087	.021	.634	-.182	.228	.045
v19	-.031	.118	-.035	.886	.035	-.046
v20	.008	.056	.008	.635	.204	-.088
v21	.075	.022	-.071	.052	.844	.048
v22	-.089	-.088	.098	.137	.789	.040
v23	-.054	.136	.008	-.142	.061	.968
v24	.150	-.080	.146	.132	.027	.494

F2

	General1
factor1	.769
factor2	.697
factor3	.780
factor4	.625
factor5	.360
factor6	.430

factor loadings of Schmid-Leiman Solution and h²

Columns 1 - 8

	General1	factor1	factor2	factor3	factor4	factor5	factor6	H² total
v1	.427	.534	.088	-.188	-.092	-.006	.019	.520
v2	.681	.499	-.062	.061	.087	-.007	-.005	.728
v3	.648	.494	.073	.056	-.096	.105	-.105	.703
v4	.649	.462	-.132	.086	.088	-.076	.152	.696
v5	.489	.452	.021	-.123	.095	-.015	.018	.469
v6	.511	.408	-.001	-.043	-.052	-.024	.262	.502
v7	.606	.398	-.042	.068	.069	-.007	.067	.542
v8	.422	.382	.108	.020	-.109	.032	-.193	.387
v9	.595	.303	.022	.233	.027	-.119	-.117	.529
v10	.599	.265	-.015	.213	.036	.063	-.051	.482
v11	.690	-.039	.530	.153	.030	-.055	.060	.789
v12	.579	-.024	.518	.006	.123	-.071	.054	.627
v13	.623	.123	.461	.070	-.104	.034	.020	.633
v14	.453	.035	.350	-.137	.314	-.083	.049	.456
v15	.571	.086	.343	.031	.064	.165	-.042	.485
v16	.606	.015	.037	.476	-.058	-.051	.051	.605

```
v17    .627    -.096     .087     .451     .154    -.087     .012     .645
v18    .430    -.056     .015     .397    -.142     .213     .041     .413
v19    .578    -.020     .085    -.022     .692     .033    -.042     .823
v20    .484     .005     .040     .005     .496     .190    -.079     .524
v21    .375     .048     .016    -.044     .041     .787     .043     .768
v22    .334    -.057    -.063     .061     .107     .736     .036     .677
v23    .409    -.035     .098     .005    -.111     .057     .874     .957
v24    .478     .096    -.057     .091     .103     .025     .446     .460

Columns   9 -  10
       H² G   H² 1st
v1     .182     .338
v2     .464     .264
v3     .420     .283
v4     .421     .275
v5     .239     .230
v6     .261     .241
v7     .367     .174
v8     .178     .208
v9     .354     .175
       .359     .124
       .476     .313
       .335     .292
       .388     .245
       .205     .250
       .326     .160
       .368     .237
       .394     .252
       .185     .228
       .334     .489
       .234     .290
       .140     .628
       .111     .566
       .167     .790
       .229     .231

sum of squared loadings
   General1  factor1  factor2  factor3  factor4  factor5  factor6    total
H²    7.139    1.884    1.086     .822    1.000    1.339    1.151   14.420
%      .495     .131     .075     .057     .069     .093     .080    1.000

percentage of extracted variance explained by general factors (%)
  .495

percentage of extracted variance explained by first order factors (%)
  .505


------ END MATRIX -----
```

APPENDIX 3A
Structural Equation Modelling Notation

Structural Equation Modelling Notation
Latent constructs can be either exogenous or endogenous. Exogenous constructs are indicated with a ξ while endogenous constructs are labelled with an $\acute{\eta}$.

Regression parameters between exogenous and endogenous constructs are denoted with $\gamma_{21.}$ Regression parameters between endogenous constructs are represented by $\beta_{21.}$

It is expected that dependent constructs cannot be perfectly predicted and thus contain an error term referred to as zeta: ζ_1.

Indicators associated with exogenous variables are diagrammed as X_1 while endogenous variables are referred to as Y_1.

The loadings of each indicator on its latent construct are symbolised by lambdas: λ_{32}

One of the major advantages of structural equation modelling is the ability to model measurement error. Measurement error associated with exogenous indicators is labelled with deltas: δ_5. Endogenous variables are illustrated with epsilons: ε_5.

Non-directional relationships are always assumed between exogenous constructs and are designated by φ_{21}.

APPENDIX 3B
Fit Indices and Their Acceptable Thresholds

Fit Index	Acceptable Threshold Levels	Description
Absolute Fit Indices		
Chi-Square χ^2	Low χ^2 relative to degrees of freedom with an insignificant *p* value ($p > 0.05$)	
Relative χ^2 (χ^2/df)	2:1 (Tabachnik and Fidell, 2007) 3:1 (Kline, 2005)	Adjusts for sample size.
Root Mean Square Error of Approximation (RMSEA)	Values less than 0.07 (Steiger, 2007)	Has a known distribution. Favours parsimony. Values less than 0.03 represent excellent fit.
GFI	Values greater than 0.95	Scaled between 0 and 1, with higher values indicating better model fit. This statistic should be used with caution.
AGFI	Values greater than 0.95	Adjusts the GFI based on the number of parameters in the model. Values can fall outside the 0-1.0 range.
RMR	Good models have small RMR (Tabachnik and Fidell, 2007)	Residual based. The average squared differences between the residuals of the sample covariances and the residuals of the estimated covariances. Unstandardised.
SRMR	SRMR less than 0.08 (Hu and Bentler, 1999)	Standardised version of the RMR. Easier to interpret due to its standardised nature.
Incremental Fit Indices		
NFI	Values greater than 0.95	Assesses fit relative to a baseline model which assumes no covariances between the observed variables. Has a tendency to overestimate fit in small samples.
NNFI (TLI)	Values greater than 0.95	Non-normed, values can fall outside the 0-1 range. Favours parsimony. Performs well in simulation studies (Sharma et al, 2005; McDonald and Marsh, 1990)
CFI	Values greater than 0.95	Normed, 0-1 range.

APPENDIX 4A
Service Stations Questionnaire

Service Station Questionnaire

1.	Please circle your level of agreement with each statement:	**Strongly Disagree**						**Strongly Agree**
	The employees were neat and tidy in appearance	1	2	3	4	5	6	7
	The store was very clean	1	2	3	4	5	6	7
	The background music was pleasant	1	2	3	4	5	6	7
	I found the interior design visually appealing	1	2	3	4	5	6	7
	The equipment was modern looking	1	2	3	4	5	6	7
	I found the staff friendly	1	2	3	4	5	6	7
	The flooring was appropriate	1	2	3	4	5	6	7
	The electronic equipment was excellent	1	2	3	4	5	6	7
	The interior design was attractive	1	2	3	4	5	6	7
	I found my way around quite easily	1	2	3	4	5	6	7
	The employees were helpful	1	2	3	4	5	6	7
	The colour schemes were pleasant	1	2	3	4	5	6	7
	The lighting was comfortable	1	2	3	4	5	6	7
	The interior layout was pleasing	1	2	3	4	5	6	7
	The forecourt was clean and tidy	1	2	3	4	5	6	7
	The atmosphere was comfortable	1	2	3	4	5	6	7
	The materials used were of high quality	1	2	3	4	5	6	7
	The store had a pleasant smell	1	2	3	4	5	6	7
	The architecture was attractive	1	2	3	4	5	6	7
	The background music was appropriate	1	2	3	4	5	6	7
	I found the physical facilities comfortable	1	2	3	4	5	6	7
	The equipment was of high quality	1	2	3	4	5	6	7
	The service station appeared to be hygienic	1	2	3	4	5	6	7
2.	Please provide your level of agreement with the following statements:	**Strongly Disagree**						**Strongly Agree**
	I am satisfied with my decision to go to this service station	1	2	3	4	5	6	7
	If I had to do it all over again, I would feel differently about this service station	1	2	3	4	5	6	7
	My choice to go to this service station was a wise one	1	2	3	4	5	6	7
	I think I did the right thing when I decided to go to this service station	1	2	3	4	5	6	7
3.	Please circle your level of agreement with the following statements:							
	The service in store was delivered promptly	1	2	3	4	5	6	7
	The store provided good service	1	2	3	4	5	6	7
	The staff were trustworthy	1	2	3	4	5	6	7
	The service was of a very high quality	1	2	3	4	5	6	7
	The service here suited my needs	1	2	3	4	5	6	7
	The staff were polite	1	2	3	4	5	6	7
	The service was efficient	1	2	3	4	5	6	7
	The staff were helpful	1	2	3	4	5	6	7
	The service here is reliable	1	2	3	4	5	6	7
	The service station provided quality service	1	2	3	4	5	6	7

4. Please state **to what extent** you experienced the following **emotions while in this store**.
Please circle a number beside **each** emotion.

	Not at all	A little	Moder-ately	Quite A bit	A lot		Not at all	A little	Moder-ately	Quite a bit	A lot
Interested	1	2	3	4	5	**Distressed**	1	2	3	4	5
Nervous	1	2	3	4	5	**Guilty**	1	2	3	4	5
Excited	1	2	3	4	5	**Scared**	1	2	3	4	5
Upset	1	2	3	4	5	**Hostile**	1	2	3	4	5
Strong	1	2	3	4	5	**Enthusiastic**	1	2	3	4	5
Irritable	1	2	3	4	5	**Proud**	1	2	3	4	5
Alert	1	2	3	4	5	**Ashamed**	1	2	3	4	5
Inspired	1	2	3	4	5	**Determined**	1	2	3	4	5
Jittery	1	2	3	4	5	**Attentive**	1	2	3	4	5
Active	1	2	3	4	5	**Afraid**	1	2	3	4	5

5. As with the previous question, please provide your level of agreement with the following statements:

	Strongly Disagree						**Strongly Agree**
I would say positive things about this service station	1	2	3	4	5	6	7
I would recommend this service station to someone else	1	2	3	4	5	6	7
I would encourage friends and family to go to this service station	1	2	3	4	5	6	7
I would consider this service station as my first choice in the future	1	2	3	4	5	6	7

6. As in the following example, please place an ✗ closest to the adjective that best describes what you think of this service station:

Example:

Good	**1**	✗	**3**	**4**	**5**	**6**	**7**	**Bad**
Important	1	2	3	4	5	6	7	Unimportant
Boring	1	2	3	4	5	6	7	Interesting
Relevant	1	2	3	4	5	6	7	Irrelevant
Exciting	1	2	3	4	5	6	7	Unexciting
Means nothing	1	2	3	4	5	6	7	Means a lot to me
Appealing	1	2	3	4	5	6	7	Unappealing
Fascinating	1	2	3	4	5	6	7	Mundane
Worthless	1	2	3	4	5	6	7	Valuable
Involving	1	2	3	4	5	6	7	Uninvolving
Not needed	1	2	3	4	5	6	7	Needed

7. **Gender** Male ❑ Female ❑

8. **Age** 16 or under ❑ 17 - 25 ❑ 26 - 35 ❑ 36 - 45 ❑
46 – 55 ❑ 56 - 65 ❑ 66 – 75 ❑ Over 75 ❑

Please sign here to give permission for the above information to be used for academic research purposes:

_____

Thank you for your time and patience!!

APPENDIX 4B
Descriptive Statistics for Age

	Frequency (Counts)	Percentage
16 or under	28	8.0
17-25	60	17.1
26-35	77	21.9
36-45	56	16.0
46-55	69	19.7
56-65	39	11.1
66-75	21	6.0
		.3
	351	100
Missing	4	
Total	355	

APPENDIX 4C
CFA Results for Assessing the Dimensionality of the Servicescape Construct

This table provides a summary of the tables in the following section:

	Design	Hygiene	Space	Equipment	Ambience	Employees
Design	Table 1					
Hygiene	Table 2	Table 7				
Space	Table 3	Table 8	Table 12			
Equipment	Table 4	Table 9	Table 13	Table 16		
Ambience	Table 5	Table 10	Table 14	Table 17	Table 19	
Employees	Table 6	Table 11	Table 15	Table 18	Table 20	Table 21

Table 1
Design

	Estimates	Standardised Loading	t-value	R^2
I found the interior design visually appealing	1.10	0.81	17.05*	0.66
The interior design was attractive	1.10	0.81	17.00*	0.65
The colour schemes were pleasant	0.77	0.58	11.17*	0.34
The materials used were of high quality	0.74	0.62	11.98*	0.38
The architecture was attractive	0.88	0.57	10.87*	0.33
I found the physical facilities comfortable	0.54	0.47	8.58*	0.22
Model fit: Chi-Square=23.98, df=9, P-value=0.00433, RMSEA=0.069				
* = significant at the p = 0.01 level				

Table 2
Design and Hygiene

	Estimates	Standardised Loading	t-value	R^2
I found the interior design visually appealing	1.10	0.81	17.05*	0.66
The interior design was attractive	1.10	0.81	17.00*	0.65
The colour schemes were pleasant	0.77	0.58	11.17*	0.34
The materials used were of high quality	0.74	0.62	11.98*	0.38
The architecture was attractive	0.88	0.57	10.87*	0.33
I found the physical facilities comfortable	0.54	0.47	8.58*	0.25
The service station appeared to be hygienic	0.74	0.50	12.85*	0.56
The store was very clean	0.58	0.75	10.98*	0.40
The forecourt was clean and tidy	0.55	0.63	6.42*	0.15
Chi-Square=73.36, df=26, P-value=0.00000, RMSEA=0.072				
* = significant at the p = 0.01 level				
Phi (φ) = 0.70(.05)				
Lambda-x modification indices indicated high crossloadings between the **I found the physical facilities comfortable** item and the Hygiene dimension (22.61)				

Table 3
Design and Space

	Estimates	Standardised Loading	t-value	R^2
I found the interior design visually appealing	10.08	0.79	16.89*	0.63
The interior design was attractive	1.10	0.86	17.14*	0.64
The colour schemes were pleasant	0.77	0.59	11.39*	0.35
The materials used were of high quality	0.77	0.64	12.68*	0.41
The architecture was attractive	0.88	0.57	10.92*	0.32
I found the physical facilities comfortable	0.56	0.48	9.07*	0.23
The flooring was appropriate	0.72	0.60	10.77*	0.36
I found my way around quite easily	0.48	0.51	9.01*	0.26
The interior layout was pleasing	0.83	0.67	12.29*	0.45

Chi-Square=73.36, df=26, P-value=0.00000, RMSEA=0.072
* = significant at the $p = 0.01$ level
Phi (φ) = 0.88(0.04)
Lambda-x modification indices indicated high crossloadings between the **The interior layout was pleasing** item and the Design dimension (15.65)

Table 4
Design and Equipment

	Estimates	Standardised Loading	t-value	R^2
I found the interior design visually appealing	0.83	0.80	16.97*	.63
The interior design was attractive	0.64	0.77	16.27*	.59
The colour schemes were pleasant	0.93	0.57	11.03*	.33
The materials used were of high quality	1.08	0.66	13.21*	.44
The architecture was attractive	1.05	0.57	11.01*	.33
I found the physical facilities comfortable	0.75	0.52	9.82*	.27
The equipment was modern looking	0.79	0.70	13.41*	.49
The electronic equipment was excellent	0.89	0.52	9.37*	.27
The equipment was of high quality	0.60	0.73	14.00*	.53

Chi-Square=142.48, df=26, P-value=0.00000, RMSEA=0.112
* = significant at the $p = 0.01$ level
Phi (φ) = 0.85(0.04)
Lambda-x modification indices indicated high crossloadings between the **I found the physical facilities comfortable** item and the Equipment dimension (24.99).

Table 5
Design and Ambience

	Estimates	Standardised Loading	t-value	R^2
I found the interior design visually appealing	1.09	.80	16.89*	.64
The interior design was attractive	1.09	.80	16.90*	.64
The colour schemes were pleasant	.78	.59	11.35*	.35
The materials used were of high quality	.75	.63	12.23*	.39
The architecture was attractive	.90	.58	11.16*	.34
I found the physical facilities comfortable	.57	.49	9.18*	.24
The background music was pleasant	.77	.61	11.16*	.38
The lighting was comfortable	.54	.67	9.36*	.22
The atmosphere was comfortable	.34	.41	7.16*	.17
The store had a pleasant smell	.52	.44	7.67*	.19
The background music was appropriate	.97	.77	14.56*	.60

Chi-Square=241.48, df=43, P-value=0.00000, RMSEA=0.114
* = significant at the $p = 0.01$ level. Phi (φ) = 0.70(0.04)
Lambda-x modification indices indicated high crossloadings between the **I found the physical facilities comfortable** item and the Ambience dimension (13.67). **The lighting was comfortable, The store had a pleasant smell** and **The background music was appropriate** all crossloaded onto Design (22.65), (15.50) and (17.07) respectively.

Table 6
Design and Employee

	Estimates	Standardised Loading	t-value	R^2
I found the interior design visually appealing	1.10	.81	10.95*	.65
The interior design was attractive	1.09	.80	16.77*	.64
The colour schemes were pleasant	.77	.59	11.25*	.34
The materials used were of high quality	.75	.63	12.33*	.40
The architecture was attractive	.89	.57	10.89*	.33
I found the physical facilities comfortable	.55	..48	8.81*	.23
The employees were neat and tidy in appearance	.49	.59	9.98*	.35
I found the staff friendly	.52	.62	9.37*	.39
The employees were helpful	53	.56	8.64*	.32

Chi-Square=71.23, df=26, P-value=0.00000, RMSEA=0.070
* = significant at the $p = 0.01$ level
Phi (φ) = 0.44(0.06)
Lambda-x modification indices indicated a crossloading of 12.59 between the **The materials used were of high quality** and the Hygiene dimension.

Table 7
Hygiene

	Estimates	Standardised Loading	t-value	R^2
The store was very clean	.54	.59	7.41*	.35
The forecourt was clean and tidy	.53	.37	5.67*	.13
The service station appeared to by hygienic	.79	.81	8.31*	.65
Model Fit: Not identified Chi-Square=0.00, df=0, P-value=1.00000, RMSEA=0.000 * = significant at the p = 0.01 level				

Table 8
Hygiene and Space

	Estimates	Standardised Loading	t-value	R^2
The store was very clean	.56	.61	10.85*	.37
The forecourt was clean and tidy	.63	.44	7.54*	.19
The service station appeared to by hygienic	.73	.74	13.30*	.54
The flooring was appropriate	.80	.66	11.94*	.44
I found my way around quite easily	.62	.66	11.93*	.44
The interior layout was pleasing	.59	.48	8.26*	.23
Chi-Square=13.66, df=8, P-value=0.09121, RMSEA=0.045 * = significant at the p = 0.01 level Phi (φ) = 0.92(0.05) Lambda-x modification indices indicated a crossloading of 6.29 between the **The forecourt was clean and tidy** and the Hygiene dimension.				

Table 9
Hygiene and Equipment

	Estimates	Standardised Loading	t-value	R^2
The store was very clean	.57	.62	16.59*	.39
The forecourt was clean and	.58	.46	6.64*	.16

tidy				
The service station appeared to by hygienic	.74	.75	12.54*	.56
The equipment was modern looking	.72	.61	10.83*	.37
The electronic equipment was excellent	.68	.55	9.77*	.31
The equipment was of high quality	1.04	.81	14.49*	.66
Chi-Square=8.69, df=8, P-value=0.36916, RMSEA=0.016 * = significant at the p = 0.01 level Phi (φ) = 0.70(0.05) There were no significant modification indices for these two constructs.				

Table 10
Hygiene and Ambience

	Estimates	Standardised Loading	t-value	R^2
The store was very clean	.61	.67	11.04*	.45
The forecourt was clean and tidy	.54	.38	6.15*	.14
The service station appeared to by hygienic	.70	.72	11.70*	.51
The background music was pleasant	.77	.61	10.99*	.38
The lighting was comfortable	.51	.44	7.67*	.20
The atmosphere was comfortable	.37	.45	7.74*	.20
The store had a pleasant smell	.50	.43	7.38*	.18
The background music was appropriate	.97	.77	14.14*	.60

Chi-Square=174.13, df=19, P-value=0.00000, RMSEA=0.152
* = significant at the p = 0.01 level
Phi (φ) = 0.66(0.06)
Lambda-x modification indices indicated high crossloadings between **The background music was pleasant**, **The lighting was comfortable, The atmosphere was comfortable**, **The store had a pleasant smell** and **The background music was appropriate** and the Hygiene dimension (21.08), (10.99), (42.84), (13.89) and (22.55) respectively.

Table 11
Hygiene and Employees

	Estimates	Standardised Loading	t-value	R^2
The store was very clean	.61	.67	11.26*	.44
The forecourt was clean and tidy	.61	.42	7.02*	.18
The service station appeared to by hygienic	.67	.68	11.50*	.46
The employees were neat and tidy in appearance	.51	.61	10.09*	.37
I found the staff friendly	.49	.59	9.72*	.34
The employees were helpful	.54	.57	9.48*	.33

Chi-Square=29.59, df=8, P-value=0.00025, RMSEA=0.087
* = significant at the p = 0.01 level
Phi (φ) = 0.79(0.06)
There were no significant modification indices for these two constructs.

Table 12
Space

	Estimates	Standardised Loading	t-value	R^2
The flooring was appropriate	.91	.75	9.61*	.56
I found my way around quite easily	.53	.56	8.22*	.32
The interior layout was pleasing	.62	.51	7.71*	.26
Model Fit: Chi-Square=0.00, df=0, P-value=1.00000, RMSEA=0.000 * = significant at the $p = 0.01$ level				

Table 13
Space and Equipment

	Estimates	Standardised Loading	t-value	R^2
The flooring was appropriate	.79	.66	11.34*	.43
I found my way around quite easily	.52	.55	9.43*	.30
The interior layout was pleasing	.74	.60	10.41*	.36
The equipment was modern looking	.78	.66	12.02*	.44
The electronic equipment was excellent	.72	.58	10.32*	.34
The equipment was of high quality	.93	.72	13.19*	.52
Chi-Square=20.75, df=8, P-value=0.00785, RMSEA=0.067 * = significant at the $p = 0.01$ level Phi (φ) = 0.82(0.05) Lambda-x modification indices indicated a crossloading of 6.29 between the **The interior layout was pleasing** and the Equipment dimension.				

Table 14
Space and Ambience

	Estimates	Standardised Loading	t-value	R^2
The flooring was appropriate	.85	.70	12.03*	.49
I found my way around quite easily	.59	.62	10.71*	.39
The interior layout was pleasing	.60	.49	8.19*	.24
The background music was pleasant	.59	.47	7.93*	.22
The lighting was comfortable	.62	.55	9.44*	.30
The atmosphere was comfortable	.43	.52	8.98*	.27
The store had a pleasant smell	.59	.50	8.57*	.25
The background music was appropriate	.78	.63	10.99*	.39

Chi-Square=159.89, df=19, P-value=0.00000, RMSEA=0.145
* = significant at the p = 0.01 level
Phi (φ) = 0.79(0.05)
Lambda-x modification indices indicated high crossloadings between **The background music was pleasant**, **The lighting was comfortable, The atmosphere was comfortable**, and **The background music was appropriate** item and the Space dimension (28.72), (22.63), (13.29), and (28.03) respectively.

Table 15
Space and Employees

	Estimates	Standardised Loading	t-value	R^2
The flooring was appropriate	.82	.67	10.95*	.46
I found my way around quite easily	.62	.66	10.75*	.44
The interior layout was pleasing	.57	.46	7.56*	.21
The employees were neat and tidy in appearance	.46	.55	8.87*	.30
I found the staff friendly	.53	.64	10.27*	.41
The employees were helpful	.55	.58	9.41*	.34

Chi-Square=22.12, df=8, P-value=0.00470, RMSEA=0.071
* = significant at the p = 0.01 level
Phi (φ) = 0.68(0.06)
Lambda-x modification indices indicated a crossloading of 10.31between the **I found my way around quite easily** item and the Employees dimension.

Table 16
Equipment

	Estimates	Standardised Loading	t-value	R^2
The equipment was modern looking	.69	.59	9.53*	.34
The electronic equipment was excellent	.70	.57	.9.31*	.32
The equipment was of high quality	1.05	.82	11.92*	.67
Model fit: Not identified Chi-Square=0.00, df=0, P-value=1.00000, RMSEA=0.000 * = significant at the *p* = 0.01 level				

Table 17
Equipment and Ambience

	Estimates	Standardised Loading	t-value	R^2
The equipment was modern looking	.69	.59	10.26*	.35
The electronic equipment was excellent	.70	.57	9.91*	.33
The equipment was of high quality	1.04	.81	13.82*	.66
The background music was pleasant	.82	. 65	11.80*	.43
The lighting was comfortable	.48	.42	7.20*	.17
The atmosphere was comfortable	.32	.39	6.65*	.15
The store had a pleasant smell	.49	.42	7.26*	.18
The background music was appropriate	1.01	.81	14.78*	.66
Chi-Square=136.41, df=19, P-value=0.00000, RMSEA=0.132 * = significant at the *p* = 0.01 level Phi (φ) = 0.55(0.06) Lambda-x modification indices indicated high crossloadings between **The background music was pleasant**, **The lighting was comfortable, The atmosphere was comfortable**, **The store had a pleasant smell** and **The background music was appropriate** item and the Equipment dimension (9.89), (9.19), (12.28), (44.48) and (21.70) respectively.				

Table 18
Equipment and Employees

	Estimates	Standardised Loading	t-value	R^2
The equipment was modern looking	.71	.60	10.11*	.36
The electronic equipment was excellent	.72	.58	9.82*	.34
The equipment was of high quality	1.02	.80	12.81*	.63
The employees were neat and tidy in appearance	.47	.57	8.65*	.32
I found the staff friendly	.53	.64	9.54*	.41
The employees were helpful	.53	.56	8.60*	.32
Chi-Square=9.07, df=8, P-value=0.33626, RMSEA=0.019 * = significant at the p = 0.01 level Phi (φ) = 0.43(0.07) There were no significant modification indices for these two constructs.				

Table 19
Ambience

	Estimates	Standardised Loading	t-value	R^2
The background music was pleasant	. 65	.82	11.80*	.43
The lighting was comfortable	.42	.48	7.20*	.17
The atmosphere was comfortable	.39	.32	6.65*	.15
The store had a pleasant smell	.42	.49	7.26*	.18
The background music was appropriate	.81	1.01	14.78*	.66
Chi-Square=43.79, df=5, P-value=0.00000, RMSEA=0.148 * = significant at the p = 0.01 level				

Table 20
Ambience and Employees

	Estimates	Standardised Loading	t-value	R^2
The background music was pleasant	.83	.66	11.97*	.44
The lighting was comfortable	.44	.39	6.71*	.15
The atmosphere was comfortable	.31	.38	6.60*	.14
The store had a pleasant smell	.44	.37	6.45*	.14
The background music was appropriate	1.06	.85	15.22*	.72
The employees were neat and tidy in appearance	.51	.61	9.37*	.38
I found the staff friendly	.48	.58	8.95*	.34
The employees were helpful	.54	.57	8.84*	.33
Chi-Square=147.25, df=19, P-value=0.00000, RMSEA=0.138 * = significant at the p = 0.01 level Phi (φ) = 0.50(0.07) Lambda-x modification indices indicated high crossloadings between **The employees were neat and tidy in appearance** item and the Ambience dimension (11.31). High crossloadings were also found between **The atmosphere was comfortable, The store had a pleasant smell** and **The background music was appropriate** items and the Employees dimension of (50.56), (14.04) and (12.89) respectively.				

Table 21
Employees

	Estimates	Standardised Loading	t-value	R^2
The employees were neat and tidy in appearance	.43	.52	7.73*	.27
I found the staff friendly	.56	.67	8.93*	.45
The employees were helpful	.55	.58	8.25*	.34
Model fit: Not identified Chi-Square=0.00, df=0, P-value=1.00000, RMSEA=0.000 * = significant at the p = 0.01 level				

APPENDIX 4D
Assessing Discriminant Validity between Constructs

To assess discriminant validity Bagozzi's (1980) formula was used. This formula proposes that if the φ coefficient is significantly less than 1, plus or minus two standard errors then there is support for discriminant validity.

	Design	**Hygiene**	**Space**	**Equipment**	**Ambience**
Design					
Hygiene	Φ = .70 (.05) c.i. = 1.96 (.05) c.i. = 0.098 = .70 + 0.098 **= .798 < 1.0**				
Space	Φ = 0.88 (.04) c.i. = 1.96 (.04) c.i. = 0.0784 = 0.88 + 0.0784 **= 0.9584 < 1.0**	Φ = .92 (.05) c.i. = 1.96 (.05) c.i. = 0.098 = .92 + 0.098 **= 1.018 > 1.0**			
Equipment	Φ = 0.85 (.04) c.i. = 1.96 (.04) c.i. = 0.0784 = 0.85 + 0.0784 **= 0.9284 < 1.0**	Φ = .70 (.05) c.i. = 1.96 (.05) c.i. = 0.098 = .70 + 0.098 **= .798 < 1.0**	Φ = .82 (.05) c.i. = 1.96 (.05) c.i. = 0.098 = .82 + 0.098 **= .91 < 1.0**		
Ambience	Φ = .70 (.04) c.i. = 1.96 (.04) c.i. = 0.0784 = .70 + 0.0784 **= .778 < 1.0**	Φ = .66 (.06) c.i. = 1.96 (.06) c.i. = 0.1176 = .66 + 0.1176 **= .777 < 1.0**	Φ = .79 (.05) c.i. = 1.96 (.05) c.i. = 0.098 = .79 + 0.098 **= .88 < 1.0**	Φ = .55 (.06) c.i. = 1.96 (.06) c.i. = 0.1176 = .55 + 0.1176 **= .66 < 1.0**	
Employees	Φ = .44 (.06) c.i. = 1.96 (.06) c.i. = 0.1176 = .44 + 0.1176 **= .5576 < 1.0**	Φ = .79 (.06) c.i. = 1.96 (.06) c.i. = 0.1176 = .79 + 0.1176 **= .90 < 1.0**	Φ = .68 (.06) c.i. = 1.96 (.06) c.i. = 0.1176 = .68 + 0.1176 **= .79 < 1.0**	Φ = .43 (.07) c.i. = 1.96 (.07) c.i. = 0.1372 = .43 + 0.1372 **= .56 < 1.0**	Φ = .50 (.07) c.i. = 1.96 (.07) c.i. = 0.1372 = .50 + 0.1372 **= .63 < 1.0**

APPENDIX 4E

Covariance Matrix for the Servicescape as a Second Order Construct

	FLOORI	EXCELL	INTERI	WAYARO	HELPFU	COLOUR	LIGHTI	SMELL	ARCHIT	MUSICA	QUALEQ	HYGIEN
FLOORI	1.46											
EXCELL	0.45	1.52										
INTERI	0.63	0.50	1.87									
WAYARO	0.48	0.28	0.39	0.89								
HELPFU	0.25	0.13	0.17	0.31	0.90							
COLOUR	0.39	0.25	0.96	0.30	0.30	1.73						
LIGHTI	0.55	0.25	0.53	0.36	0.25	0.70	1.30					
SMELL	0.47	0.45	0.46	0.33	0.19	0.30	0.30	1.38				
ARCHIT	0.53	0.29	0.94	0.33	0.16	0.60	0.50	0.61	2.40			
MUSICA	0.39	0.28	0.63	0.32	0.23	0.44	0.45	0.41	0.65	1.56		
QUALEQ	0.52	0.73	0.68	0.39	0.20	0.53	0.41	0.63	0.73	0.49	1.64	
HYGIEN	0.51	0.27	0.49	0.43	0.27	0.35	0.35	0.36	0.47	0.37	0.54	0.97

APPENDIX 4F
Inter-correlations for the Second Order Servicescape Constructs

Correlation Matrix of ETA and KSI

	equip	space	design	ambience	hygiene	employee
equip	1.00					
space	0.75	1.00				
design	0.62	0.71	1.00			
ambience	0.54	0.62	0.51	1.00		
hygiene	0.74	0.84	0.70	0.61	1.00	
employee	0.56	0.63	0.53	0.46	0.63	1.00

APPENDIX 4G
Fit Indices for the Final Servicescape Construct

Degrees of Freedom = 99
Minimum Fit Function Chi-Square = 209.58 (P = 0.00)
Normal Theory Weighted Least Squares Chi-Square = 205.15 (P = 0.00)
Estimated Non-centrality Parameter (NCP) = 106.15
90 Percent Confidence Interval for NCP = (68.93 ; 151.13)
Minimum Fit Function Value = 0.59
Population Discrepancy Function Value (F0) = 0.30
90 Percent Confidence Interval for F0 = (0.19 ; 0.43)
Root Mean Square Error of Approximation (RMSEA) = 0.055
90 Percent Confidence Interval for RMSEA = (0.044 ; 0.066)
P-Value for Test of Close Fit (RMSEA < 0.05) = 0.21
Expected Cross-Validation Index (ECVI) = 0.79
90 Percent Confidence Interval for ECVI = (0.68 ; 0.92)
ECVI for Saturated Model = 0.77
ECVI for Independence Model = 5.32
Chi-Square for Independence Model with 120 Degrees of Freedom = 1852.67
Independence AIC = 1884.67
Model AIC = 279.15
Saturated AIC = 272.00
Independence CAIC = 1962.62
Model CAIC = 459.41
Saturated CAIC = 934.61
Normed Fit Index (NFI) = 0.89
Non-Normed Fit Index (NNFI) = 0.92
Parsimony Normed Fit Index (PNFI) = 0.73
Comparative Fit Index (CFI) = 0.94
Incremental Fit Index (IFI) = 0.94
Relative Fit Index (RFI) = 0.86
Critical N (CN) = 228.43
Root Mean Square Residual (RMR) = 0.072
Standardized RMR = 0.054
Goodness of Fit Index (GFI) = 0.93
Adjusted Goodness of Fit Index (AGFI) = 0.91
Parsimony Goodness of Fit Index (PGFI) = 0.68

APPENDIX 4H
Composite Reliability and Average Variance Extracted Calculations

Ambience	Loadings	Standard-ised Error	Loadings squared	Standard-ised Error squared
The background music was pleasant	0.60	0.60	0.36	0.36
The background music was appropriate	0.97	0.06	0.9409	0.0036
Sum	1.57	0.66	2.4649	0.3636
Sum Squared	2.46			
Composite Reliability	**0.79**			
Average Variance Extracted	**2.41**			

Space & Hygiene	Loadings	Standard-ised Error	Loadings squared	Standard-ised Error squared
The flooring was appropriate	0.63	0.60	0.3969	0.36
I found my way around quite easily	0.63	0.61	0.3969	0.3721
The store was very clean	0.63	0.6	0.3969	0.36
The service station appeared to be hygienic	0.72	0.49	0.5184	0.2401
Sum	1.26	1.21	0.7938	0.7321
Sum Squared	1.59			
Composite Reliability	**0.57**			
Average Variance Extracted	**0.41**			

Design	Loadings	Standard-ised Error	Loadings squared	Standard-ised Error squared
I found the interior design visually appealing	0.81	0.34	0.6561	0.1156
The interior design was attractive	0.81	0.34	0.6561	0.1156
The colour schemes were pleasant	0.59	0.65	0.3481	0.4225
The architecture was attractive	0.58	0.67	0.3364	0.4489
Sum	2.79	2.00	1.9967	1.1026
Sum Squared	7.78			
Composite Reliability	**0.80**			
Average Variance Extracted	**0.64**			

Equipment	Loadings	Standardised Error	Loadings squared	Standardised Error squared
The equipment was modern looking	0.67	0.56	0.4489	0.3136
The electronic equipment was excellent	0.55	0.70	0.3025	0.49
The equipment was of high quality	0.75	0.44	0.5625	0.1936
Sum	1.97	1.70	1.3139	0.9972
Sum Squared	3.88			
Composite Reliability	**0.70**			
Average Variance Extracted	**0.49**			

Employees	Loadings	Standardised Error	Loadings squared	Standardised Error squared
The employees were neat and tidy in appearance	0.62	0.62	0.3844	0.3844
I found the staff friendly	0.59	0.65	0.3481	0.4225
The employees were helpful	0.55	0.70	0.3025	0.49
Sum	1.76	1.97	1.035	1.2969
Sum Squared	3.10			
Composite Reliability	**0.61**			
Average Variance Extracted	**0.32**			

APPENDIX 4I
Testing the Discriminant Validity of the Servicescape Construct II

Incorrelations of factors (Eta and Ksi)

	Equipment	Space	Design	Ambience	Employee
Space	0.75				
Design	0.66	0.71			
Ambience	0.44	0.47	0.41		
Employee	0.55	0.59	0.53	0.35	

Intercorrelations of Factors Squared

	Equipment	Space	Design	Ambience	Employee
Space	0.5625				
Design	0.4356	0.5041			
Ambience	0.1936	0.2209	0.1681		
Employee	0.3025	0.3481	0.2809	0.1225	

Average Variance Extracted (from Table 5.12)

	Equipment	Space	Design	Ambience	Employee
Space	0.41				
Design	0.49	0.41			
Ambience	0.49	0.41	0.64		
Employee	0.49	0.41	0.32	0.32	

Whether Discriminant Validity Was Met

	Equipment	**Space**	**Design**	**Ambience**	**Employee**
Space	FALSE				
Design	0	FALSE			
Ambience	0	0	0		
Employee	0	0	0	0	

APPENDIX 5A
SERVQUAL Instrument

Expectations Measures

Tangibles
Excellent XYZ companies will have modern-looking equipment.
The physical facilities at excellent XYZ companies will be visually appealing.
Employees of excellent XYZ companies will be neat-appearing.
Materials associated with the service (such as pamphlets or statements) will be visually appealing in an excellent telephone company.

Reliability
When excellent XYZ companies promise to do something by a certain time, they will do so.
When customers have a problem, excellent XYZ companies will show sincere interest in solving it.
Excellent XYZcompanies will perform the service right the first time.
Excellent XYZ companies will provide their services at the time they promise to do so.
Excellent XYZ companies will insist on error-free records.

Responsiveness
Employees of excellent XYZ companies will tell customers exactly when services will be performed.
Employees of excellent XYZ companies will give prompt service to customers.
Employees of excellent XYZ companies will always be willing to help customers.
Employees of excellent XYZ companies will never be too busy to respond to customer requests.

Assurance
The behaviour of employees of excellent XYZ companies will instill confidence in customers.
Customers of excellent XYZ companies will feel safe in their transactions.
Employees of excellent XYZ companies will be consistently courteous with customers.
Employees of excellent XYZ companies will have the knowledge to answer customer questions.

Empathy
Excellent XYZ companies will give customers individual attention.
Excellent XYZ companies will have operating hours convenient to all their customers.
Excellent XYZ companies will have employees who give customers personal attention.
Excellent XYZ companies will have their customers' bets interests at heart.
The employees of excellent XYZ companies will understand the specific needs of their customers.

Perceptions Measures

Tangibles

1. XYZ has modern-looking equipment.
2. XYZ's physical facilities are visually appealing.
3. XYZ's employees are neat-appearing.
4. Materials associated with the service are visually appealing.

Reliability

5. When XYZ promises to do something by a certain time, it does so.
6. When you have a problem, XYZ shows sincere interest in solving it.
7. XYZ performs the service right first time.
8. XYZ provides its services at the time it promises to do so.
9. XYZ insists on error free records.

Responsiveness

10. Employees of XYZ tell you exactly when services will be performed.
11. Employees of XYZ give you prompt service.
12. Employees of XYZ are always willing to help you.
13. Employees of XYZ are never too busy to respond to your requests.

Assurance

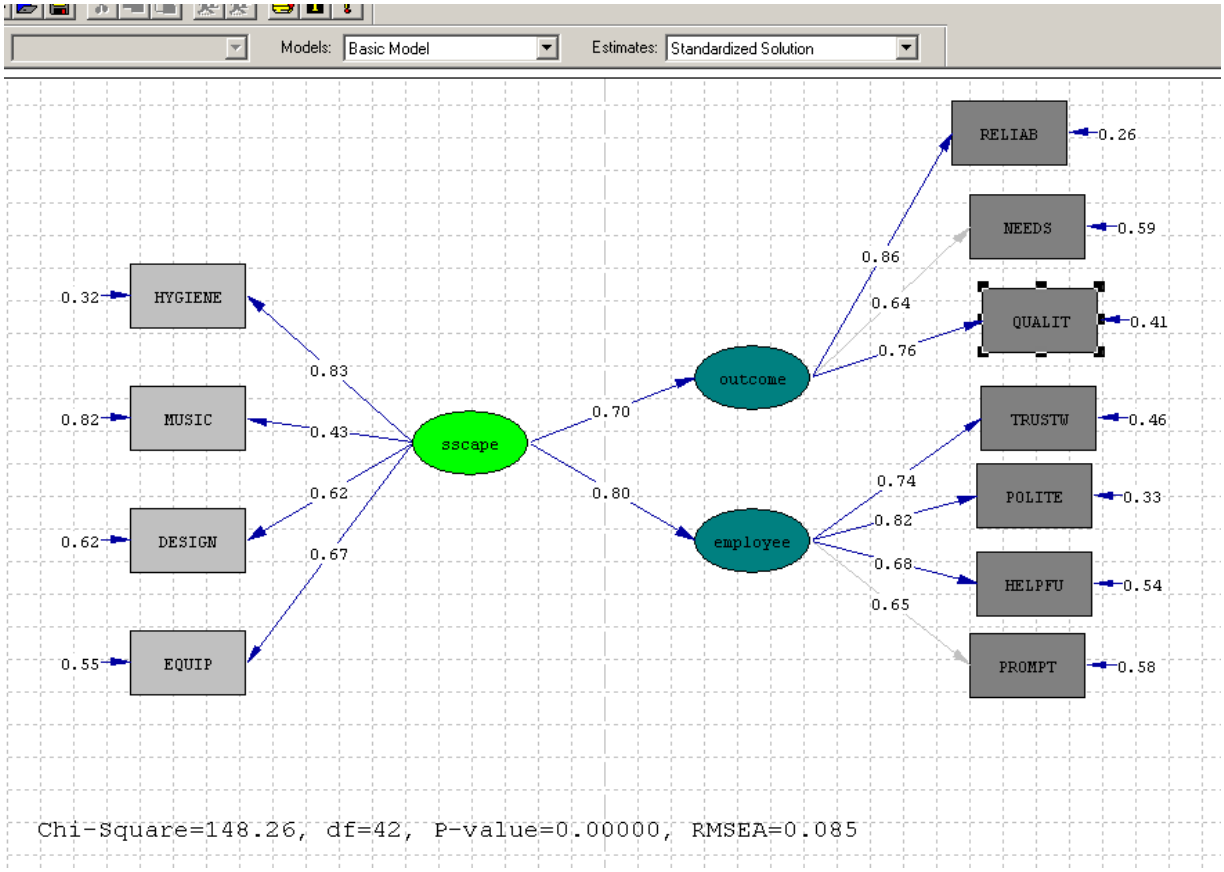
14. The behaviour of employees of XYZ instills confidence in customers.
15. You feel safe in your transactions with XYZ.
16. Employees of XYZ are consistently courteous with you.
17. Employees of XYZ have the knowledge to answer your questions.

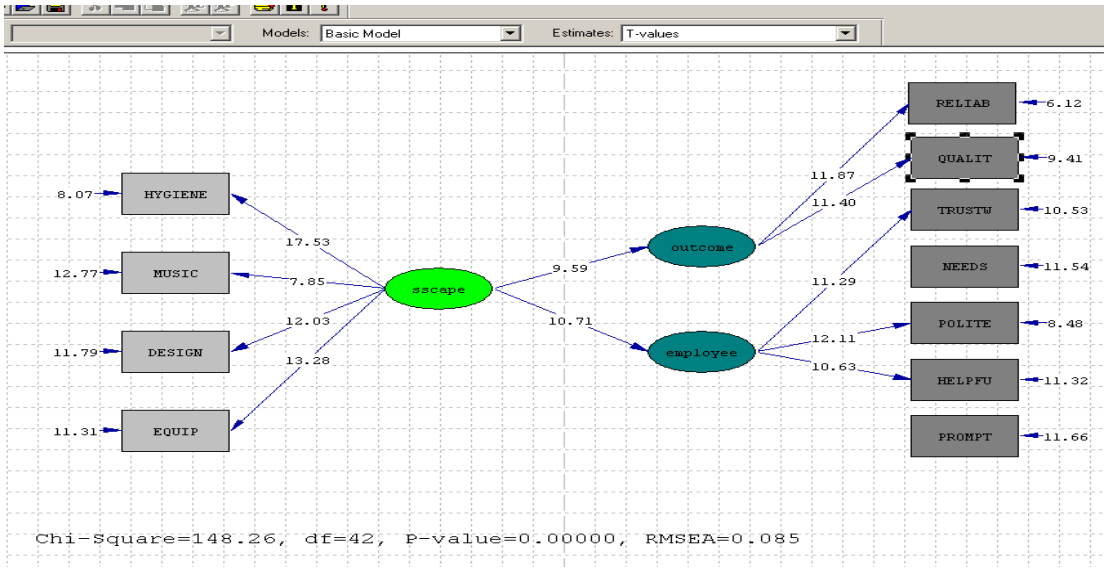
Empathy

18. XYZ gives you individual attention.
19. XYZ has operating hours convenient to all its customers.
20. XYZ has employees who give you personal attention.
21. XYZ has your best interests at heart.
22. Employees of XYZ understand your specific needs.

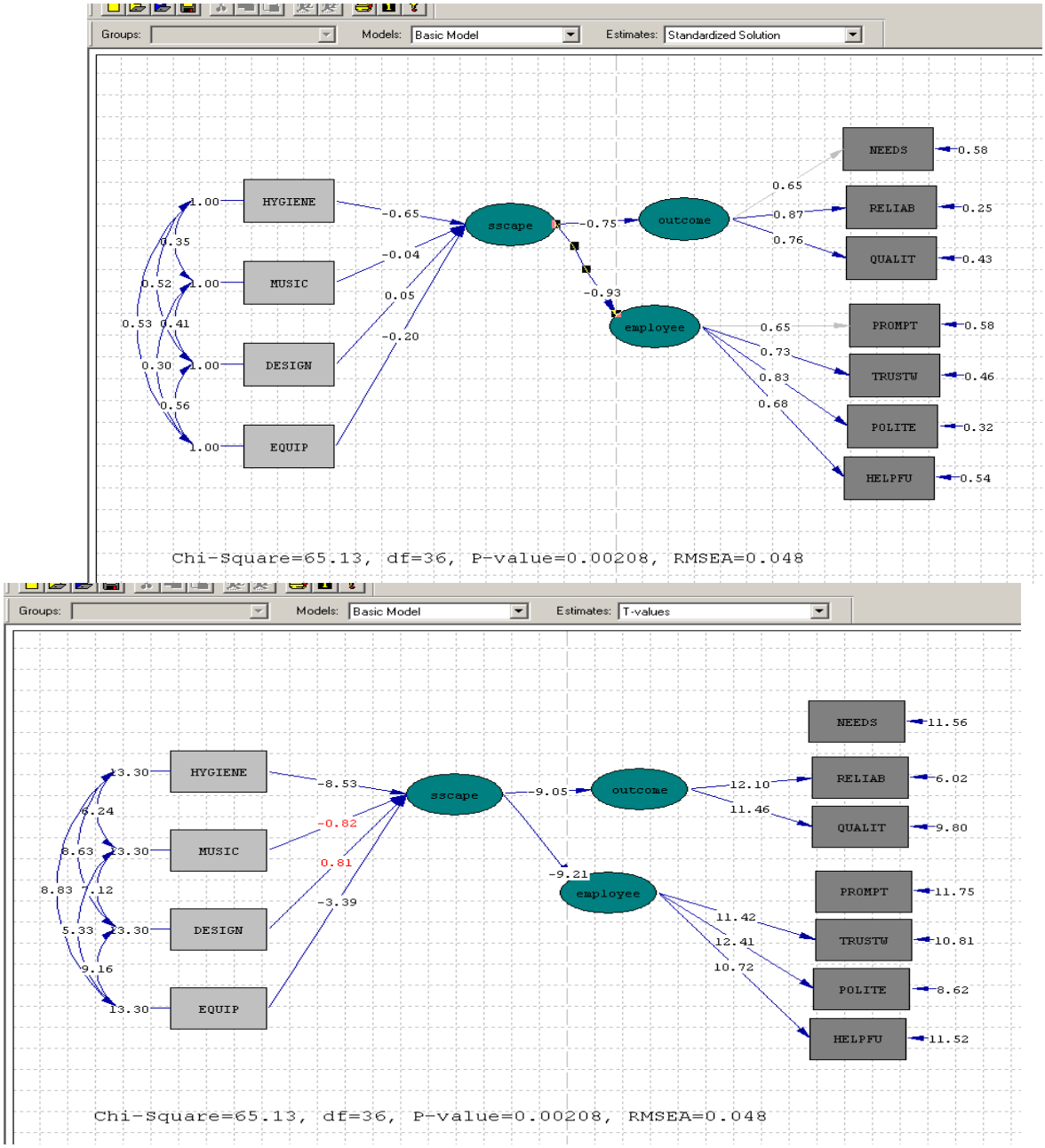
APPENDIX 5B
Formative versus Reflective Models

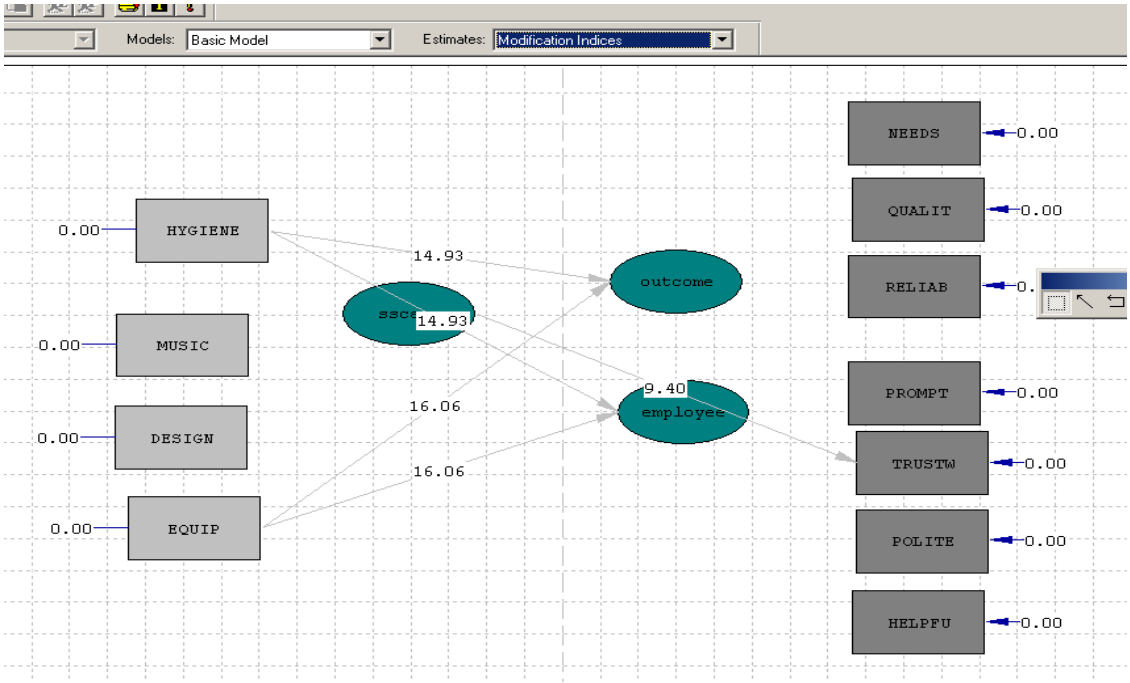
Reflective Model





Formative Model




Models: Basic Model
Estimates: Modification Indices
NEEDS
0.00
QUALIT
0.00
RELIAB
0.
0.00
HYGIENE
14.93
outcome
0.00
MUSIC
14.93
9.40
PROMPT
0.00
0.00
DESIGN
16.06
employee
TRUSTW
0.00
16.06
0.00
EQUIP
POLITE
0.00
HELPFU
0.00

APPENDIX 7A
Phi-matrix for Full Measurement Model

PHI

	employee	service	hygiene	design	equip	satis
employee	0.43					
	(0.06)					
	6.79					
service	0.35	0.57				
	(0.04)	(0.06)				
	8.06	9.08				
hygiene	0.29	0.26	0.31			
	(0.04)	(0.04)	(0.05)			
	7.50	7.10	5.90			
design	0.30	0.34	0.40	1.21		
	(0.05)	(0.06)	(0.06)	(0.14)		
	5.72	5.78	7.21	8.55		
equip	0.24	0.31	0.27	0.56	0.46	
	(0.04)	(0.05)	(0.04)	(0.08)	(0.09)	
	5.77	6.52	6.38	7.11	5.04	
satis	0.31	0.44	0.32	0.37	0.31	1.27
	(0.05)	(0.06)	(0.05)	(0.07)	(0.06)	(0.22)
	6.36	7.93	6.91	4.98	5.56	5.84
behav in	0.29	0.53	0.28	0.82	0.48	0.62
	(0.06)	(0.08)	(0.06)	(0.12)	(0.08)	(0.10)
	4.44	6.77	4.71	7.10	5.79	6.25

APPENDIX 7B
Exploratory Factor Analysis Testing
Common Methods Variance

Communalities

	Initial	Extraction
The employees were neat and tidy in appearance	.414	.401
The store was very clean	.394	.458
The background music was pleasant	.450	.509
I found the interior design visually appealing	.579	.612
The equipment was modern looking	.440	.457
I found the staff friendly	.343	.336
The flooring was appropriate	.381	.377
The electronic equipment was excellent	.324	.246
The interior design was attractive	.604	.625
The employees were helpful	.457	.528
The colour schemes were pleasant	.452	.445
The interior layout was pleasing	.429	.493
The architecture was attractive	.362	.323
The background music was appropriate	.537	.737
The equipment was of high quality	.485	.451
The service station appeared to be hygienic	.483	.494
I am satisfied with my decision to go to this service station	.487	.617
If I had to do it all over again, I would feel differently about this service station	.302	.423
The service in store was delivered promptly	.546	.494
The staff were trustworthy	.510	.475
The service at the centre was of a very high quality	.581	.586
The service here suited my needs	.440	.429
The staff were polite	.674	.678
The service was efficient	.689	.729
The staff were helpful	.584	.663
The service here is reliable	.631	.662
The service station provided quality service	.555	.542
I would say positive things about this service station	.622	.640
I would recommend this service station to someone else	.868	.946
I would encourage friends and family to go to this service station	.862	.888

Extraction Method: Principal Axis Factoring.

Total Variance Explained

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings(a)
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	9.703	32.343	32.343	9.267	30.889	30.889	6.440
2	2.670	8.898	41.241	2.271	7.569	38.458	5.175
3	1.955	6.517	47.759	1.603	5.343	43.801	7.286
4	1.411	4.704	52.462	1.008	3.361	47.162	4.414
5	1.304	4.345	56.807	.913	3.042	50.204	2.582
6	1.130	3.766	60.573	.656	2.185	52.390	6.154
7	1.036	3.452	64.026	.550	1.832	54.222	1.698
8	.899	2.995	67.021				
9	.809	2.697	69.718				
10	.775	2.582	72.300				
11	.716	2.385	74.685				
12	.663	2.211	76.897				
13	.656	2.186	79.083				
14	.611	2.038	81.121				
15	.542	1.806	82.926				
16	.536	1.785	84.712				
17	.500	1.666	86.378				
18	.483	1.610	87.988				
19	.458	1.526	89.514				
20	.424	1.414	90.928				
21	.376	1.255	92.183				
22	.366	1.220	93.402				
23	.333	1.110	94.512				
24	.314	1.048	95.560				
25	.300	.998	96.558				
26	.282	.940	97.498				
27	.263	.876	98.374				
28	.241	.802	99.177				
29	.175	.585	99.761				
30	.072	.239	100.000				

Extraction Method: Principal Axis Factoring.

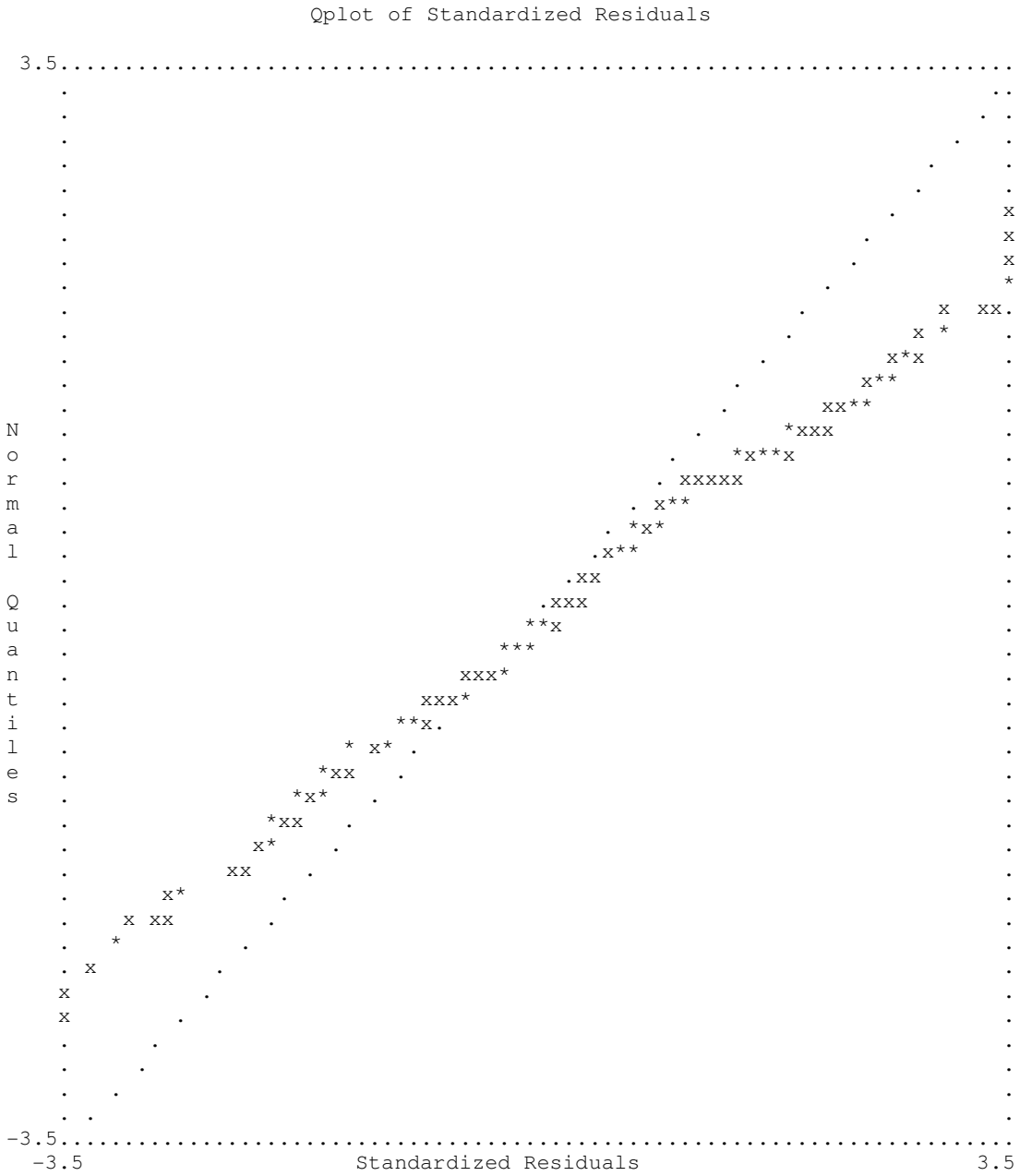
a When factors are correlated, sums of squared loadings cannot be added to obtain a total variance.

Pattern Matrix(a)

	Factor						
	1	2	3	4	5	6	7
The interior layout was pleasing	.775						
The interior design was attractive	.749						
I found the interior design visually appealing	.745						
The colour schemes were pleasant	.675						
The equipment was modern looking	.625						
The architecture was attractive	.512						
The equipment was of high quality							
The flooring was appropriate							
The employees were helpful		.807					
The staff were helpful		.794					
I found the staff friendly		.513					
The staff were polite		.492	.459				
The service in store was delivered promptly		.432					
The staff were trustworthy							
The service here is reliable			.897				
The service station provided quality service			.729				
The service was efficient			.715				
The service here suited my needs			.680				
The service at the centre was of a very high quality			.414				
The electronic equipment was excellent							
I would recommend this service station to someone else				.977			
I would encourage friends and family to go to this service station				.943			
I would say positive things about this service station				.524			
The background music was appropriate					.804		
The background music was pleasant					.709		
The store was very clean						.671	
The service station appeared to be hygienic						.507	
The employees were neat and tidy in appearance						.426	
If I had to do it all over again, I would feel differently about this service station							.635
I am satisfied with my decision to go to this service station							.520

Extraction Method: Principal Axis Factoring.
Rotation Method: Promax with Kaiser Normalization.
a Rotation converged in 7 iterations.

APPENDIX 7C
Q-plot of Standardized Residuals


Qplot of Standardized Residuals
3.5
-3.5
Normal Quantiles
-3.5
Standardized Residuals
3.5

APPENDIX 7D
Covariance Matrix for Full Research Model

Covariance Matrix

	SATIS1	SATIS2	PROMPT	TRUSTW	SUITNEED	POLITE
SATIS1	1.21					
SATIS2	0.81	2.56				
PROMPT	0.27	0.13	0.68			
TRUSTW	0.38	0.25	0.36	1.13		
SUITNEED	0.46	0.38	0.28	0.39	1.13	
POLITE	0.23	0.17	0.30	0.40	0.30	0.40
HELPFUL	0.32	0.30	0.36	0.51	0.34	0.34
RELIABLE	0.41	0.38	0.30	0.41	0.53	0.28
QUALITY	0.50	0.39	0.25	0.48	0.50	0.26
RECOMM	0.61	0.46	0.17	0.36	0.50	0.16
FAMILY	0.64	0.39	0.25	0.47	0.54	0.24
FIRSTCHO	0.67	0.52	0.18	0.54	0.40	0.26
NEAT	0.21	0.02	0.26	0.38	0.15	0.23
CLEAN	0.32	0.04	0.19	0.35	0.20	0.19
DESIGN	0.36	0.27	0.25	0.49	0.27	0.25
MODERN	0.33	0.14	0.15	0.36	0.33	0.19
FLOORI	0.44	0.27	0.28	0.50	0.38	0.27
EXCELL	0.34	0.29	0.24	0.40	0.29	0.18
INTERI	0.41	0.23	0.26	0.46	0.33	0.20
WAYARO	0.27	0.09	0.32	0.46	0.29	0.27
COLOUR	0.24	0.16	0.13	0.37	0.23	0.11
ARCHIT	0.27	0.18	0.12	0.29	0.24	0.17
QUALEQ	0.46	0.37	0.24	0.48	0.50	0.28
HYGIEN	0.49	0.25	0.30	0.43	0.36	0.26

Covariance Matrix

	HELPFUL	RELIABLE	QUALITY	RECOMM	FAMILY	FIRSTCHO
HELPFUL	0.93					
RELIABLE	0.32	0.80				
QUALITY	0.33	0.62	1.09			
RECOMM	0.33	0.49	0.62	2.63		
FAMILY	0.43	0.48	0.62	2.50	2.83	

FIRSTCHO	0.36	0.40	0.57	1.69	1.75	3.57
NEAT	0.31	0.24	0.17	0.18	0.21	0.29
CLEAN	0.27	0.24	0.27	0.33	0.36	0.47
DESIGN	0.27	0.33	0.37	0.72	0.75	0.97
MODERN	0.23	0.23	0.31	0.53	0.53	0.60
FLOORI	0.36	0.30	0.41	0.36	0.41	0.70
EXCELL	0.21	0.37	0.38	0.46	0.48	0.50
INTERI	0.21	0.34	0.42	0.85	0.87	1.05
WAYARO	0.35	0.27	0.32	0.11	0.21	0.19
COLOUR	0.35	0.19	0.26	0.64	0.72	0.53
ARCHIT	0.27	0.26	0.29	0.66	0.74	1.02
QUALEQ	0.31	0.49	0.51	0.72	0.71	0.71
HYGIEN	0.31	0.32	0.41	0.42	0.45	0.53

Covariance Matrix

	NEAT	CLEAN	DESIGN	MODERN	FLOORI	EXCELL
NEAT	0.70					
CLEAN	0.33	0.85				
DESIGN	0.29	0.45	1.85			
MODERN	0.24	0.33	0.85	1.39		
FLOORI	0.24	0.41	0.60	0.53	1.46	
EXCELL	0.21	0.28	0.48	0.48	0.45	1.52
INTERI	0.24	0.40	1.22	0.72	0.63	0.50
WAYARO	0.26	0.32	0.40	0.31	0.48	0.28
COLOUR	0.22	0.28	0.79	0.61	0.39	0.25
ARCHIT	0.26	0.29	1.03	0.66	0.53	0.29
QUALEQ	0.25	0.39	0.74	0.72	0.52	0.73
HYGIEN	0.24	0.43	0.53	0.40	0.51	0.27

Covariance Matrix

	INTERI	WAYARO	COLOUR	ARCHIT	QUALEQ	HYGIEN
INTERI	1.87					

WAYARO	0.39	0.89			
COLOUR	0.96	0.30	1.73		
ARCHIT	0.94	0.33	0.60	2.40	
QUALEQ	0.68	0.39	0.53	0.73	1.64
HYGIEN	0.49	0.43	0.3		