A Comparison of Two Approaches to Measuring Brand Equity in the Hotel Industry

by

Seung Hyun Lee

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2014 by the
Graduate Supervisory Committee:

Timothy Tyrrell, Chair
Muzaffer Uysal
Kathleen Andereck
Gyan Nyaupane

ARIZONA STATE UNIVERSITY

May 2014

ABSTRACT

Branding and brand management have been top management priorities in the hotel industry. Some researchers have concluded that strong branding would be an efficient way for hotels and hotel chains to differentiate themselves from each other. Recent studies have focused on the establishment of a brand equity model and the relevant causal relationships of the model. Most of these studies have used types of desirability scales examining the importance of individual factors in measuring brand equity. However, they ignore the trade-offs that affect and characterize choice. Particularly, the personal decision process implied by the hierarchical brand equity model is absent. This study proposed two alternative measures of brand equity, analytic hierarchy process (AHP) and conjoint analysis (CA), to address these limitations. The AHP and the CA were compared using several validity measures to aid in selecting efficient methods. This study examined the validity of AHP and CA under two data collection methods applied to hotel branding: paper-based survey and online survey. Result showed that the AHP data collection methods were easier, as well as with respect to saving time and costs. Results also indicated that the AHP is equivalent to the CA with respect to predictive accuracy. Practical differences for hotel branding in attribute preferences were clearly observed between the AHP and the CA. The AHP results were consistent with previous studies by awarding high importance to perceived quality and brand loyalty and lower importance to brand awareness and brand image. Managerial implications were provided for results. In terms of practicality in data collection, the study results revealed that the data gathered online leads to a slightly lower internal and predictive validity. A limitation of this study was that the two methods were not perfectly

comparable. Nevertheless, the validity of both AHP and CA seems satisfactory for both methods. The study results also offer useful perspectives to consider when choosing between the two methods, as well as between AHP and CA.

DEDICATION

This is for my parents, for always encouraging and supporting me.

ACKNOWLEDGMENTS

First and foremost, I am immensely grateful to Dr. Timothy Tyrrell, my mentor and dissertation committee chair. This work would not have been possible without his valuable help and patient support. I consider myself very lucky to have Dr. Timothy Tyrrell serve as the major professor for my Ph.D. program. What a great example and admirable scholar! This dissertation would not have been possible without the help, support and patience of my other committee members, Dr. Kathleen Andereck, Dr. Muzaffer Uysal, and Dr. Gyan Nyaupane. I would like to thank the Arizona Board of Regents (ABOR) for financial support for this project. Finally, I am so thankful to my wife and parents for their encouragement and strong support. Especially, my wife who has been with me all the way and has been my momentum to finish this “long voyage.”

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

CHAPTER Page

APPENDIX Page

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Brand equity is valued as a very important concept in business practice as well as in academic research because marketers can gain competitive advantages through strong brands (Aaker, 1998; Keller, 1993, 2000). Many companies develop marketing strategies in order to improve their sales and to make their brands stand out among competitive ones. For most firms, the ultimate goal of marketing success is to generate a brand, which can differentiate their companies (Jung and Sung, 2008). A brand has also been defined as a distinguishing name and symbol (such as a logo, trademark, or package design) intended to identify the goods or services of either one seller or a group of sellers, and to differentiate those goods or services from those of competitors (Aaker, 1991, p. 7).

Various definitions of brand equity can be found in the literature depending on the purpose of study, but there seems to be a basic agreement on the concept of brand equity. The consensus is that brand equity is the value added to the product by the name of brand (Farquhar, 1989). Farquhar (1989) stated that brand is valuable only if it has a meaning to consumers. Brand equity research in marketing has largely concentrated on customer perception. A more specific definition of brand equity is given by Aaker (1991) who defines it as “a set of brand assets and liabilities linked to a brand, its name and symbol that add to or subtract from the value provided by product or service to a firm and/or that firms customers” (p.15). Another definition by Keller (1993) focuses on marketing; he describes brand equity as “the differential effect of brand knowledge on consumer response to the marketing of the brand” (p. 1).

Building brand equity, or strong brands, is considered to be one of the key drivers of a business's success (Prasad and Dev, 2000). High brand equity levels are known to lead to higher consumer preferences and purchase intentions (Cobb-Walgren et al., 1995) as well as higher stock returns (Aaker and Jacobson, 1994). Besides, high brand equity brings an opportunity for successful extensions, resilience against competitors' promotional pressures, and creation of barriers to competitive entry (Farquhar, 1989). Building and properly managing brand equity has become essential for any business organizations and hospitality organizations are no exception (Ahmad and Hashim, 2010).

Simoes and Dibb (2001) argue that branding plays a special role in service companies because strong brands increase customers' trust of the invisible, enabling them to better visualise and understand the intangible and reduce customers' perceived financial, social or safety risk in buying services, which are difficult to evaluate before purchase. de Chernatony and Segal-Horn (2003) maintain that building a service brand is different than building a product brand and that managing a service brand should be conceptually different than managing those of product. In service marketing, the company brand is the primary brand, whereas in packaged goods marketing the product brand is referred to as the primary brand (Low and Lamb Jr., 2000). Researchers have interpreted service brand as a promise to the customer (e.g., Ambler and Styles, 1996; Berry, 2000; de Chernatony and Segal-Horn, 2001; Mistry, 1998). Berry (2000) wrote that the role of brand equity in the service industry is very important because strong brands increase customers' trust of the invisible purchase. Because the service business, including the hospitality industry, is labor-intensive, the customer experience involving interactions with employees plays a critical role in building the value of brand (Kim et al., 2008). A well-established brand can bring some promise to customers in that the name may reduce the risks due to service characteristics.

The hotel industry shares the same characteristics applicable to the services (Kayaman and Arasli, 2007). In the hotel industry, customers often base their purchase decisions on their perceptions of a company's brand (e.g., Marriott, Hilton, and Hyatt). That is, hotel customers tend to lean toward the strongly built outstanding brands for easy selection, meanwhile feeling they are thereby reducing their purchase risk (Zhou and Jiang, 2011). As Prasad and Dev (2000) stated, the stronger the hotel brand equity, the more customers will prefer that hotel brand. More specifically, studies have proposed that strong hotel brand equity can contribute to improved financial performance because it can positively influence consumers to book with a particular brand (Prasad and Dev, 2000). Evidence suggests that independent hotels have lost ground in market share to branded hotels. A study by Forgacs (2003) showed that branded hotels in the United States accounted for more than 70% of the total room supply in 2000, as compared to approximately 61% in 1990. Freitag (2008) revealed that chain-affiliated hotels consistently grew occupancy and Revenue per Available Room (RevPar) faster than non-branded hotels. The significant increase is attributed to the benefits associated with branding. Brand equity had been widely recognized as the most valuable asset to hospitality and tourism companies and has become a top management priority since strong brands provide a series of benefits to service firms, such as greater customer loyalty and higher resiliency to endure crisis situations, higher profit margins, higher market value (O’Neil and Xiao, 2006), more favorable customer response to price change, and licensing and brand extension opportunities (Keller, 2001).

Brand equity can be measured through either a financial or customer-based perspective (Keller and Lehmann, 2006). However, since the financial based approach has limitations in terms of providing unbiased estimates of a brand's intrinsic value by merely accounting purposes (Aaker, 1991, 1996; Keller, 1998), the customer-based brand equity (CBBE) approach is the dominant perspective and the one preferred by a majority of academics and practitioners in marketing research because if a brand has no meaning or value to the consumer it is ultimately meaningless to investors, manufacturers, or retailers (Cobb-Walgren et al., 1995). Although a financial approach may provide a more precise insight into the valuation of brand, it may not be useful for brand managers to establish marketing strategies because financial approach is only limited to a brand's value estimation (Keller, 1993). The customer-based brand equity approach is more practical in a sense that the information offers a strategic vision of customer behavior and managers can develop brand strategies accordingly (Lassar et al., 1995; Prasad and Dev, 2000; Yoo and Donthu, 2001; Kim et al., 2008). Consequently, this study focuses on the customer perspective of hospitality firms.

From a customer's point of view, the major components of brand equity are brand loyalty, perceived quality, brand awareness, brand association and other assets, which include assets such as patents and trademarks (Aaker, 1991). Among these five brand equity dimensions, the first four represent customers' evaluations and reactions to the brand that can be readily understood by customers (Barwise, 1993). They have been widely used to measure customer-based brand equity in previous studies. To measure customer-based hotel brand equity, this study utilized four core categories including brand loyalty, perceived quality, brand awareness, and brand image proposed by Aaker

(1991, 1996) as the basis, and construct the customer-based hotel brand equity attributes, which would help in empirically understanding and evaluating customer preferences related to brand equity.

In consumer preference measurement research, trade objects (e.g. hotel product or service) are defined by, and specified on, a limited number of relevant characteristics (Vriens, 1995). These characteristics are called attributes. These attributes are defined on a number of discrete levels, which can be used to define hotel products. Measuring the extent in which the various characteristics of a trade object are related to its overall attractiveness is referred to as preference structure measurement (Green and Srinivasan, 1990). Several models and methods have been proposed for the measurement of preference structures. Preference structure measurement models (PSM models) are concerned with the way in which the attribute levels are related to the overall attractiveness (Vriens, 1995). Preference structure measurement methods (PSM methods) are concerned with the way in which the relation between the attribute levels and the overall attractiveness is quantified or measured (Vriens, 1995). Preference analysis consists of two alternative approaches: the compositional and decompositional (Green and Srinivasan, 1990). Compositional methods (e.g. traditional self-explicated methods) use direct questions on each attribute and attribute level to estimate consumer preference, while decompositional methods (e.g. conjoint analysis) ask for general judgments on multiattribute product alternatives (Helm, Scholl, Manthey, and Steiner, 2004a).

It is well understood that consumer brand preferences and choices are characterized by variety and complexity in terms of hotel involvement, nature of hotel product (tangible or intangible), frequency of hotel stay, price risk involved, and information requirements. The task of understanding and evaluating customer preferences related to brand equity is necessarily complicated. In other words, since brand equity is reflected in the customer's brand preference (Chang and Liu, 2009), different methods of measuring brand equity are required for different situations. Hence, this study seeks to compare results obtained from the above-mentioned alternatives approaches to measuring hotel brand equity and contribute to the evidence for and against alternative consumer preference measurement.

Statement of Problem

According to Keller (2003), brand equity is a complex multidimensional concept that requires many different types of measurement techniques. However, most empirical studies on brand equity have focused on the construction and definition of the measurements (Prasad and Dev, 2000; Yoo and Donthu, 2001; Washburn and Plank, 2002; Lee and Back, 2008) as well as the causal relationship between brand equity and other customer behavior conceptual variables (Sloot et al., 2005; Hyun, 2009). The dimensions of brand equity feature diverse connotations and structures in different researches (Burmann et al., 2009; de Chernatony and McDonald 2003). In particular, the determination of the importance ratings in the hierarchical brand equity model is absent (Hsu et al., 2012). Moreover, in measuring the importance of hotel product attributes related to brand equity, traditionally researchers ask consumers to rate the importance of attributes one at a time on some type of desirability scale ignoring the trade-offs that affect and characterize choice (Goldberg et al., 1984).

Many trade-offs of attributes can occur; however, hospitality customers typically choose a product or service, which consists of multiple attributes (Lewis et al., 1991).

According to consumer choice theory, the choice of a single product or service will be influenced by a multiplicity of factors that are not always possible to determine by using traditional rating scale surveys (Huber, 1987). Consumers tend to evaluate a product using various attributes of a given product. In reality, when hotel customers make a choice, they do not consider each attribute separately; instead, they consider the product attributes jointly. This means that in a real purchase situation, hospitality customers examine and evaluate alternatives that simultaneously vary across several attributes in making their final purchase selection (Kim et al., 2004). In this perspective, consumers' reaction to multiattribute product alternatives is difficult to measure on interval or ratio scales because customer choices usually involve the evaluation of several attributes. Since a majority of the brand equity factors, such as perceived quality, brand name and image, may be in non-metric form, it is difficult to measure using an interval or ratio scale (Baek et al., 2006).

A common problem of the traditional approach is that respondents are subject to a ceiling effect (Lawson et al., 2006; Oh, 2001). In other words, consumers might consider all brand equity attributes to be very important (e.g. all factors rated '5' on a 5-pointscale, with '5' being very important). Since there is no logical reason to evaluate the different brand equity categories while giving each category equal importance, it is necessary to develop a less arbitrary method of incorporating relative weights into the evaluation criteria. To overcome this problem, one of the advanced approaches used in many disciplines is conjoint analysis (Lawson et al., 2006).

The rationale for conjoint methodology lies in the idea that an individual choice process is never straightforward. In conjoint analysis, a product contains a bundle of different attributes and levels (Dubas and Strong, 1993). Thus, when individuals are faced with the question of selecting a product, they will not only examine the specific attributes but they will also examine the product as a whole. A conjoint analysis is a multivariate technique, which determines the relative importance of a product's multi-dimensional attributes and measures consumers' degree of preferences for each level of each attribute (Green and Wind, 1975; Tull and Hawkins, 1993) and enables an attribute hierarchy to be established (Peral et al., 2012). A conjoint analysis has had strong predictive power of consumer choices among multiattribute product alternatives and was proven as the appropriate method for hospitality and tourism research in predicting consumers' choice among multiattribute alternatives (e.g., Wong and Lam, 2001). In other words, the application of the conjoint analysis is so great in identifying and understanding the combined effects of product attributes on preferences for a product/service (Hobbs, 1996) that the analysis has been utilized to design the most preferred product by hotel customers (Goldberg et al., 1984; Hu and Hiemstra, 1996; Lewis et al., 1991; Wind et al., 1989; Wong and Lam, 2001), travel packages (Mulhbacher and Botschen, 1988) and meeting planning products (Renaghan and Kay, 1987; Hu and Hiemstra, 1996). Despite its popularity in the consumer preference research, little attention has been paid to the measurement of consumer preferences regarding hotel branding using a conjoint analysis.

While there are several types of conjoint analysis (e.g., see Huber, 1997), full-profile conjoint analysis is one of the most common preference measurement methods (Gil and Sa´nchez, 1997), which has previously proved successful (Mulye 1998; Helm et al., 2003; Helm, Scholl, Manthey, and Steiner, 2004a; Helm, Steiner, Scholl, and

Manthey, 2004b; Scholl et al., 2005; Scholz et al., 2006; van Til et al., 2008; Benaïm et al., 2010; Klein et al., 2010), and was employed in this study. However, a well-known problem of conjoint analysis (hereinafter CA) is that of dealing with large numbers of attributes (Meißner et al., 2008). Using a large number of attributes may cause problems of validity due to an information overload of the respondents (Green and Srinvasan, 1990).

In order to solve the multiple-dimension problem (or multiattribute design problem) inherent in this estimation, several empirical studies have shown the general potential of analytic hierarchy process (hereinafter AHP), particularly in complex product evaluation tasks consisting of many attributes (Meißner and Decker 2009). The AHP is an alternative methodology and is well-established in muti-attributive utility measurement for solving multiple criteria decision making (MCDM) problems. Nonetheless, AHP compared to CA is still rather unpopular in marketing research. This issue is of major practical relevance. If, at least in certain situations, CA is not clearly superior in validity to AHP, it becomes highly questionable whether future applications for measuring customers' preferences should be done by CA, as AHP has considerable practical advantages over CA. There are advantages in terms of ease of data collection, data analysis and research design as well as with respect to savings of time and costs in data collection and data analysis (Krapp and Sattler, 2001; Helm et al., 2003, 2004a; 2004b; Scholl et al., 2005; Meißner and Decker, 2009). When considered in empirical studies, CA's superiority frequently has not been found. In some comparative studies AHP shows high predictive accuracy (Mulye, 1998; Helm, Scholl, Manthey, and Steiner, 2003, 2004a; Scholl et al., 2005; Meißner, Scholz, and Decker, 2008; Meißner and

Decker, 2009), while other studies (Tscheulin 1991, 1992; Helm et al., 2004b; Koo and Koo, 2010) favour CA. Because of the inconclusive results of past research concerning the validity of CA compared to AHP further research is needed and the choice of method was thought to be dependent on the decision context (van Til et al., 2008). To address this issue the present study designed an empirical study which compares the validity of CA against AHP as instruments for measuring preferences regarding hotel branding.

Regarding the choice of a data collection method, traditionally, data collection in both AHP and CA has involved the use of paper-based personal interviews. However, recent years have seen a trend toward the use of online surveys (Orme and King, 1998). The internet has become an important and effective tool for administering consumer's surveys (Bonilla, 2010) since collecting the data on the Internet is a more cost effective method compared to a paper based personal interview (Klein et al., 2010), particularly if a probability sample is needed from a target population that is scattered over a wide geographic area. Especially, conjoint analyses are very time consuming and cost-intensive (Klein et al., 2010). This is why conjoint analyses often base on a small sample size or a convenience sample (university students) (Klein et al., 2010). By changing data collection process to the Internet, it is easier to obtain large sample in a shorter amount of time. For this reason, using the Internet especially for conjoint analysis receives growing interest in marketing research (Saltzman and MacElroy, 1999). According to Sethuraman et al. (2005), Internet-based conjoint analysis accounts for forty to fifty percent of all conjoint analysis applications.

Despite its increased use of the Internet for data collection, little is known about problems arising from the application of CA over the Internet and the quality of this data

(Melles, Laumann, and Holling, 2000; Klein, Nihalani, and Krishnan, 2010). Only the articles by Melles et al. (2000) and Klein et al. (2010) pertain to the topic and demonstrated the reliability and validity of CA when administrated over the Internet. For instance, Melles, Laumann and Holling (2000) presented evidence that useful conjoint analysis data can be collected over the internet, although its reliability may be lower than for other data collection methods. Furthermore, Klein et al. (2010) showed that the results of an online-conjoint analysis are of a higher validity than the results of an interviewer-based conjoint analysis. Their analysis indicated that the presence of an interviewer influences the data negatively since the respondents of the interviewer-based survey are feeling observed and controlled by the interviewer and are in more stressful situation than the respondents of the online survey.

The complexity of the problem (number of attributes and attribute levels) might have an influence on the results. Melles et al. (2000) cautioned that the suitability of conjoint analysis over the Internet depends on the number of attributes in the design since a larger number of stimuli increase the complexity and the difficulty of the evaluation task and thus the validity may decrease (Lines and Denstadli, 2004). The literature argues for example that the validity decreases dramatically with the number of 20 stimuli (Büschken, 1994). Klein et al. (2010) assume that particularly when the complexity of the task is high, the interviewer could possibly affect the results in a positive way and suggest the number of attributes or levels should be increased and included in future online studies, which would affect the orthogonal main-effect design and the number of stimuli. Trying to prove or disprove this assumption, the present study considers the same type of

CA as Klein et al. (2010), but use a more complex decision problem which consists of 16 stimuli (cf. the former study use nine stimuli only).

To ensure a high comparability to the results of former studies the present study followed a similar procedure as used in the study of Klein et al. (2010). In an offline survey, for example, the questionnaire is done by paper and pencil and with help of an interviewer, so that the respondent can obtain assistance easily in case of any questions. In the case of an online (Internet-based) survey the questions are displayed in the same order as the paper and pencil questionnaires but on a computer screen without an interviewer present to help the respondent. Since the questions of both surveys were exactly same, the only difference between offline and online data collection modes was the presence of the interviewer in one setting. This permits conclusions about the relevance of the interviewer and thus the influence of the interviewer on the validity (Klein et al., 2010). To ensure a correct behavior of the interviewer without leading the respondent, the interviewer only assists if a respondent asks for help as also advised by Klein et al. (2010). The study by Meißner, Scholz, and Decker (2008) in particular indicates that the AHP approach can also be used in online research settings. However, little is known about the quality of data generated by an online version of AHP.

Most predictive validity studies have focused primarily on comparisons across model types (e.g., several types of AHP and CA). To data, no research has focused on the predictive performance of AHP and CA across alternative data collection methods. To fill this gap in the published literature, the present study was undertaken. AHP’s potential for market share predictions in consumer research settings is still an open research issue (Meißner and Decker, 2009). Nonetheless, there is little academic research on this issue.

Only two studies (e.g. Meißner and Decker, 2009; Scholz et al., 2010) are dealing just with the market simulation. Because of the lack of other studies in that, this study is one of the few in which both hit rates (individual predictions) and choice share accuracy (aggregate predictions) have been used as validity criteria to differentiate models. Hit rate (HR) is a measure of how well a utility function can predict consumer choice. Choice share prediction accuracy (MASE: mean absolute share error) is an aggregate measure of how well the observed market shares (or shares of choice) line up with the predicted market shares.

Research Objectives

The objective of the research is to examine the internal and predictive validity of AHP and CA across offline and online data collection methods applied to hotel branding. Additionally, this study examines the feasibility and convergent validity of the models.

Research Questions

Both AHP and CA mainly differ with respect to their basic conception. In conjoint analysis, the respondent is confronted with a trade-off task; whereas the AHP develops the trade-offs in the course of structuring and analyzing a series of pair-wise comparison matrices (Baglione, 1994). The AHP is a compositional model while CA is a decompositional model.

Each method has its strengths and weaknesses, and there is no prior reason to assume one method will outperform the other. An advantage of conjoint analysis over the AHP is that conjoint can accommodate interactions. Interaction effects, however, have not been found to contribute significant amounts of explained variation in most conjoint studies (Akaah and Korgaonkar, 1983). In addition, conjoint models are more representative of the market place, for example, a product's attributes are viewed together rather than two at a time, however, conjoint analysis is more cognitively demanding than compositional approaches for that same reason, resulting in greater respondent fatigue (Baglione, 1994), and this advantage of a higher degree of realism disappears rapidly as the number of attributes included in the study increases (Mulye, 1998). In contrast, AHP has some potential advantages due to its simplicity and flexibility in dealing with complex problems (e.g. Mulye, 1998). Considering these pros and cons, the feasibility of the methods might depend on the respondents' cognitive ability to perform the task and the difficulty of the task. Besides, respondents' motivation is important in conducting time consuming interviews (Krapp and Sattler, 2001). Since valid measurements are only possible if respondents are able and willing to apply the method in a motivated manner (Helm et al., 2003; Scholl et al., 2005), it is important to explore the differences between the techniques and how user friendly each is. The following research question examines the issue:

RQ1: Are there any differences in the respondents' subjective evaluations of the methods in terms of (a) enjoyment, (b) difficulty and clarity, and (c) realism?

Brand equity remains a widely discussed topic in the marketing literature. Extant literature offers two broad perspectives of brand equity measurement: the financial perspective and the consumer perspective. Unfortunately, a notable limitation of financial measures that are often used for accounting purposes, such as mergers and acquisitions, is that they provide little or no guidance to managers in implementing and evaluating strategies that help build brand equity (Sinha, Ashill, and Gazley, 2008). They ignore the consumer role in the generation of brand value and the various aspects of brand management (such as brand awareness, brand loyalty, brand image and perceived quality). Contrary to this, customer-based measures of brand equity (Aaker, 1991) help managers evaluate their marketing strategies, including product, price, promotion, and distribution strategies. Focusing on brand equity from a customer's perspective enables marketing managers to determine how their marketing efforts, such as positioning and promotional strategies, contribute to the value of their brands in the mind of the customer. Since customers are the ultimate role players of brand equity, as they are the source of cash flow and resulting profit (Prasad and Dev, 2000), the present study focused on consumer-based brand equity. In other words, how customers perceive a product or service with respect to the brand (Kim, Kim, and An, 2003; Capon, 2008) and its importance on the consumer's perception of brand in the hotel industry.

Since consumer-based brand equity involves consumers' perception and attitude towards a brand which has an effect on the purchase intention of the consumer (Keller, 2003), a hotel will have a strong brand equity when customers have a positive perception of, and attitude towards, that hotel's brand (Prasad and Dev, 2000). In turn, brand equity has a great contribution to overall brand preference (Mishara and Datta, 2011). Preference of a brand leads to the intention of purchasing the brand over others (Wang et al., 2008). Customer-based brand equity has been thought of as a prerequisite for brand preference, which in turn affects consumers' intention to purchase (Tolba and Hassan, 2009). As brand equity is reflected in the customer's brand preference, it could be said that brand preference would be reflected in purchase or usage intention (Chang and Liu, 2009). Other empirical studies in the literature supported the positive relationship between customer-based brand equity constructs (brand awareness, brand loyalty, brand image and perceived quality) and brand preference and purchase intentions, and ultimately brand choice and concluded that high brand equity generates greater brand preference, and higher purchase intentions (Cobb-Walgren et al., 1995; Vakratsas and Ambler, 1999; Myers, 2003; Prasad and Dav, 2000; de Chernaony, 2004; Mishra and Datta, 2011; Moradi and Zarei, 2011; Tolba and Hassan, 2009; Chen and Chang, 2008). It is said that strong brand equity results in customers showing a preference for one product over another, although the products could be basically identical (Kotler, 2003).

For many hotel companies, their single biggest asset is their hotel brand (Simon, 2011). Thus, it is relevant to understand what influences consumer brand preferences and how these preferences translate into purchase intentions. Consequently, strong brand equity has become a very important factor that influences consumer brand preferences. Success in brand management arises from understanding and managing brand equity correctly to produce strong attributes that will influence consumers when making their choices (Ukpebor and Ipogah, 2008). To maximize the branding efforts of hotel firms, managers need to understand first-hand how brand equity, as an indicator of brand strategy success, can be measured (So and King, 2010). More importantly, how it can be built based on customer preferences. To enable hotel marketing managers to consider how their marketing programs improve the value of their brands in the minds of consumers, a hotel firm has to be able to measure and evaluate the hotel firm's brand equity according to their customer preferences. Measuring the subjective preferences or the choices of customers is an important task in several scientific disciplines like marketing, psychology, consumer behavioral research, and economics, as well as tourism. In particular, the measurement of consumer preferences or choices becomes a main issue for some research areas, such as marketing and decision analysis (Helm et al., 2004a). Some methods have been developed to measure customer's preferences. Again, two methods that are commonly used are the Analytic Hierarchy Process and conjoint analysis (Yudhistira, 2002).

Table 1

Preference Measurement in Decision Analysis and Marketing

	Decision Analysis	Marketing
Problem	Selection of alternatives	Design of products/services
Objective	Maximum subjective utility	Maximum consumer preferences
Core problem	Modeling and measuring preferences	Modeling and measuring preferences
Selection methods	Scoring methods	Self explanatory methods
	Multiattribute utility theory	Multidimensional scaling
	Analytic Hierarchy Process (AHP)	Conjoint Analysis (CA)

Source: Helm et al., (2004a).

According to Table 1, the AHP and the CA are suggested methods for measuring preferences with the former approach identified as a multiple criteria decision analysis (MCDA) technique and the latter associated with marketing research and practice. The AHP is generally used when decision makers are assumed to maximize subjective utility in decision analysis, while the CA is generally used to measure customer preferences in marketing. If different preference measurement methods come to similar results, high convergent validity can be presumed (Scholl et al., 2005; Meißner and Decker, 2009). Since the AHP and the CA are used for modeling and measuring consumer preferences regarding hotel branding, the following research question results:

RQ2: Do the AHP hotel branding results accord generally with the CA (convergence)? If so, to what extent does the AHP have convergent validity with the CA with respect to (a) importance ratings, (b) part-worth estimations, and (c) estimated overall utilities?

In terms of practicality in data collection, online surveys have several important advantages over traditional paper-and-pencil (offline) surveys that make them particularly attractive to researchers, because data can be collected faster, at low cost, and from a geographically dispersed population using probability sampling techniques (Mulye, 1998). Notwithstanding the increasing popularity of and reliance on the Internet it appears to be less suitable for the collection of high quality data since the advent of the Web has also produced a decrease in respondents' patience for long questionnaires (Netzer and Srinivasan, 2011). This might have a negative impact on the quality of the gathered data. Since both AHP and CA are rather complex methods compared to a regular questionnaire (Klein et al., 2010), it is questionable whether the quality of data gathered online without an interviewer is comparable to that of data collected using traditional paper-based (offline) methods and with the help of an interviewer. The following research questions examine this concern:

RQ3: Do both the offline and online data collection modes lead to comparable results regarding the internal validity of AHP?

RQ4: Do both the offline and online data collection modes lead to comparable results regarding the internal validity of CA?

Even though internet-based surveying gains increasing importance (Fricker et al., 2005), there are certain disadvantages in administering a complex task in a non-personal interview setting, and the reliability of such approaches needs to be weighed against their potential benefits (Mulye, 1998). Unlike traditional paper-based (offline) surveys, the use of online surveys does not permit the asking of questions in case task instructions are unclear. Especially since respondents participating in either a CA study or an AHP study often require detail explanations of the task (Helm et al., 2004b; Klein et al., 2010), it becomes highly questionable whether the validity of the data is negatively affected by internet-based surveying (Duffy, Smith, Terhanian, and Bremer, 2005; Grant, Teller and Teller, 2005; Schillewaert and Meulemeester, 2005; Klein et al., 2010). The following research question addresses this concern:

RQ5: Do both the offline and online data collection modes lead to comparable results regarding the predictive validity of AHP and CA?

Although data for both AHP and conjoint analysis are gathered through traditional offline (paper-and pencil) methods, the recent trend involves the move of the market research industry to Web-based (online) data collection since it is easier to reach a large number of people at relatively low cost and relatively quickly (Klein et al., 2010; Netzer and Srinivasan, 2011). However, since these two data collection modes (offline vs. online) differ in environment, it is questionable whether AHP and CA reflect differences in predictive performance across the two data collection modes. As both methods have the main objective of providing information to predict the preferences (choices) of consumers, the following research question is:

RQ6: Are both AHP and CA fairly comparable in predictive performance across offline and online data collection modes?

Up to now, there is no evidence whether the validity of data might be moderated or not by the data collection method (Melles et al., 2000). This could be a further limitation to a broad application of specific data collection methods like the Internet. The following research question examines the issue:

RQ7: Do both the offline and online data collection modes moderate the differences in predictive accuracy among the AHP and CA?

Structure of the Study

The remainder of the paper is structured as follows:

Chapter 2 first discusses the modeling of consumer choice behavior in general and examines consumer choice theory from different perspective including information-processing models (e.g. multiattribute attitude theory) and rational choice models (e.g. multiattribute value theory and random utility theory). Next, basic descriptions of AHP and CA are given and compared on a theoretical basis in terms of their similarities and differences. In addition, literature review provides details of the three parent disciplines: 1) consumer-based brand equity, which provides the foundations on which to build a brand equity theory in the hotel context; 2) hotel brand equity, which proves the applicability of consumer-based brand equity theory to the hotel context; and 3) the key attributes of hotel brand equity and their levels.

Chapter 3 depicts the design of the empirical study, including defining the target population, designing the sampling plan, specifying the data collection instrument and methods, and explains the methodology used to compare the two classes of models. This chapter discusses the potential confounding variables that will be controlled for in this study.

Chapter 4 first provides the results of summary statistics on demographic as well as traveling characteristics. Next, the two samples (offline and online samples) are compared to understand the nature of potential differences between the two groups.

Finally, the research results are presented in sequence and relative to each research question. A brief summary of the study results is also provided.

Chapter 5 first compares and contrasts the research findings with other relevant research. Next, the research results are discussed and the conclusions are presented. Included in the discussion are the limitations of this research and opportunities for future research.

CHAPTER 2

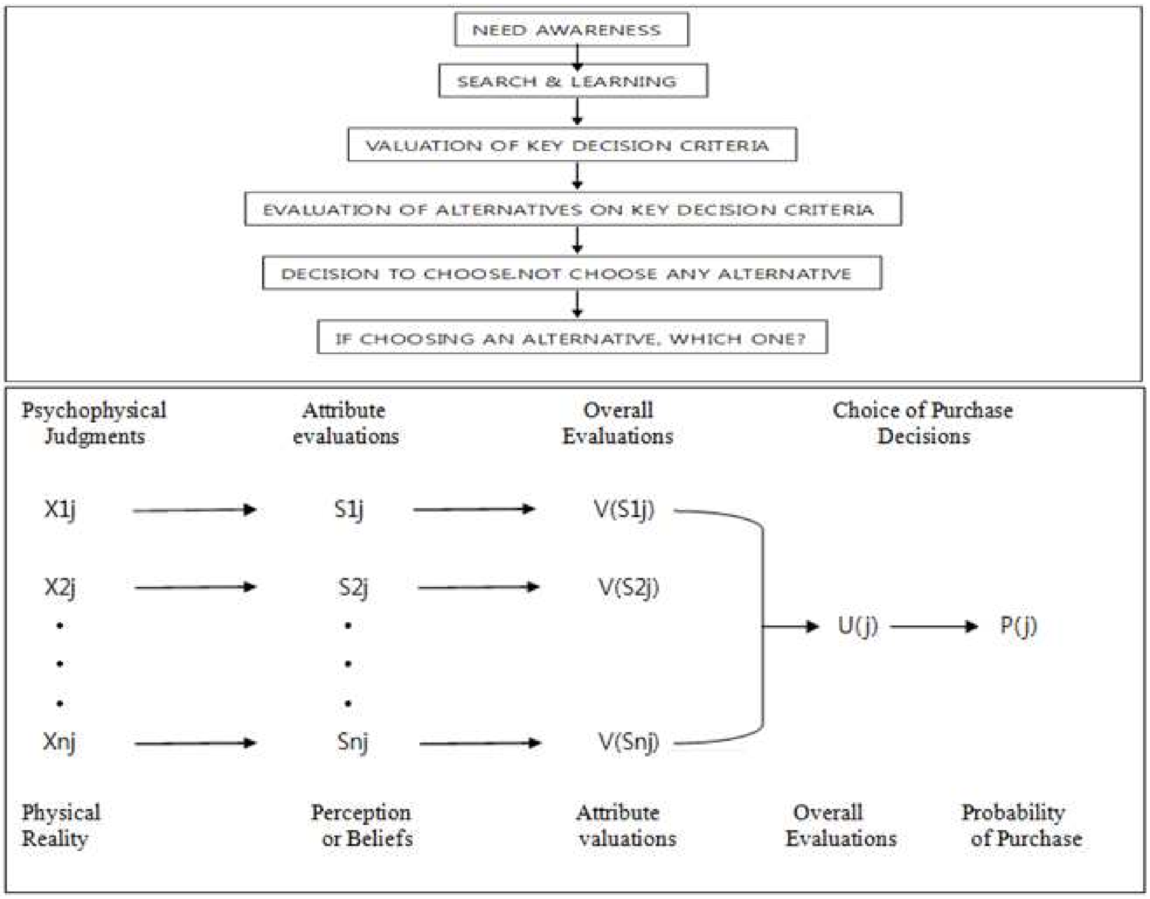
LITERATURE REVIEW

To better understand the importance of consumer-based brand equity on consumer perceptions of a brand, it is necessary to have an overview of consumer behavior (Ukpebor and Ipogah, 2008). An understanding of customers' needs and preferences is the foundation of any successful marketing strategy. A marketing manager, however, who has to decide on the allocation of his or her marketing budget, needs to know more. To predict consumer behavior it is essential for hospitality marketers to understand the consumer decision making process. Understanding the behavior of hospitality and tourism consumers is among the most important challenges facing management. According to Reid and Bojanic (2009) the study of consumer behavior is based on two fundamental ideas: that consumer behavior is rational and predictable and that marketers can influence this behavior. The chapter first discusses how consumers make choices in general and examines consumer choice theory from different perspectives including information-processing models and rational choice models. Next the two multiattribute choice models (AHP and CA) are discussed in more detail and compared on a theoretical basis.

Understanding Consumer Decision Making

When consumers make decisions concerning the purchase of goods and services, a very complex decision-making process takes place (Reid and Bojanic, 2009). Since consumer decision making is extremely complex, hospitality marketing managers constantly strive to learn more about the way consumers reach decisions (Reid and Bojanic, 2009), for this will allow managers to better serve the needs of consumers. Most applied work in consumer choice analysis is based on the rational choice paradigm. It is based on a simple explanation of decision making theory: consumers are hypothesized to approach choice situation with a predefined utility function which defines how the observed attributes of products will be integrated to form overall evaluations of desirability or utility (Kara, 1993). Once the alternatives are evaluated, consumers are hypothesized to choose the option with the highest overall utility or value (McFadden, 1980). Consequently, the development and use of multiattribute utility models have received considerable research attention in the literature.

An important aspect of understanding consumer behaviour is the study of the consumer's decision-making process. This has been a major activity in the marketing and decision analysis area where several approaches to modeling the consumer's choice process have been proposed, each with their unique assumptions, structures, and applications. The process by which consumers compare brands on a set of determinant attributes and make choices is very complicated. This process of complex decision making conceptualized by Louviere (1988a) is illustrated in Figure 1. This figure suggests that consumers form psychophysical (perceptual) as well as value judgments about brands. Psychophysical judgments (e.g., Gescheider, 1976) involve subjective perceptions of physical reality in which individuals form impressions about the position of each determinant attribute based on physical brand characteristics. After consumers form impressions of the positions of various alternatives on the attributes, they make value judgments about how good it is for alternatives to be positioned on each attribute. This evaluation process can be inferred from an analysis of the way in which consumers integrate information about different determinant attributes to form overall impressions of brands (Louviere, 1988a). It is this integration or combining of attribute information that one studies with the conjoint analysis technique in general and information integration theory (Anderson, 1981) in particular.



Source: Louviere (1988a).

Figure 1. Complex decision making.

According to Louviere (1988a), consumers' overall impressions or judgments of the attributes of brands are relative to the set of brands that they consider. Hence, these judgments may change if [1] additional brands are added to or deleted from those already evaluated, [2] new information is acquired that changes the set of determinant attributes by adding or deleting one or more, or [3] consumers' beliefs about the values of attributes are changed by new information prior to choice. Following the comparison, evaluation and impression-formation stage, consumers form a final choice and decide which brand to choose. Commonly, this involves deciding which brand is better, taking into account all available information.

Consequently, a central task in consumer research is to understand how choices are made and what influences choice. Studies that focus solely on asking consumers the degree to which they like or dislike a determined product fail to address this question (Won and Bravo, 2009). Basically, a response to a single question fails to capture how behavior towards a product is influenced by the presence of other attributes (Won and Bravo, 2009). Thus, single evaluations of product-specific features do not tell us much about what determines choice and under what circumstances choice for a particular product will occur. From this perspective multiattribute choice models play a central role in describing and prediction of consumer choice.

Consumer Choice Theory

All contemporary theories and models of consumer behavior argue that consumer choice behavior is a dynamic and a complex process and prediction of this behavior often involves use of some type of multiattribute choice model (Kara, 1993). Theories of consumer choice have been drawn from different perspectives. Among these, two of the most common are the information-processing models and the rational choice models (Bettman et al., 1998; Wilkie, 1994). Information-processing models have integrated various concepts from the behavioral sciences, for example social psychology (Ratchford, 1975; Fishbein, 1967; Rosenberg, 1956; Bonoma and Johnston, 1979; Hauser and Urban, 1979). These models use concepts such as belief, attitude, and intention.

The general assumption underlying the application of social psychology based models to consumer choice situations is that a consumer has a certain level of a particular predictor of behavior for each of the available alternatives, and would select that alternative for which the consumer has the highest level of that predictor (Ajzen and Fishbein, 1980). Information-processing models propose that choice is limited by the notion of bounded rationality (Simon, 1955). In other words, individual capacity for analysis is limited and that decisions are mainly influenced by perceptions and the attitude formation towards the product (Won and Bravo, 2009). In this view, choice occurs as a behavior in response to a decision-making problem. Problems are often influenced by a complex cognitive process that includes perceptions, attitudes, preferences and behavioral intentions toward the product (McFadden, 1986).

Social psychologists interested in the study of attitude formation and, in particular, the way in which attitudes influence behavior toward the object of the attitudes developed what is known as the multiattribute model. The concept of the multiattribute model has its origins in social psychology with the theory that people make rational decisions before they act in a certain manner. The research of attitudes towards objects or behavior is mostly based on work made by Fishbein and/or Rosenberg, (Ajzen and Fishbein, 1980; Bettman et al., 1975; Fishbein, 1963; Rosenberg, 1956). The models try to predict the behavior based on the attitudes that an individual holds versus an object, e.g. a product or brand. The field of multiattribute models has been studied extensively by researchers in fields like economics, psychology and behavioral decision theory (Huber, 1974).

Under the umbrella of information-processing models is the multiattribute attitude theory developed by Fishbein and Ajzen (1975), also known as Fishbein's multiattribute model. The key proposition in Fishbein's theory is that that the preference for a determined product is influenced by the multiplicity of attitude consumers have towards the product and the strength of the belief towards the attributes of the product. The mathematical expression of the basic Fishbein's multiattribute model is formulated as follows.

$$A_j = \sum_{i=1}^{n} B_{ij}\ I_i$$

In this model: "i" is the attribute of the product, "j" is the brand of the product, "A" represents the consumer's attitude towards the brand "j", "I" is the importance given to an attribute "i"; and "B" refers to the strength of the consumer's belief towards attribute "i" of the product "j" (Wilkie, 1994, p. 288).

The multiattribute attitude model suggests that consumers can develop more than a single attitude towards a product. Thus, Fishbein's model proposes that, in order to understand how consumers make choices, researchers must focus on the behaviour of the consumer towards the product (Foxall, 1983). According to Silk (2006), the multiattribute attitude theory has a fairly high level of predictive validity. As a result, it is of interest for practitioners, providing critical information on how marketers could change their consumers' attitudes. By understanding consumers' behaviours towards a particular brand it would be possible to influence consumers' preference towards a certain products by changing the ascribed features of the product (Won and Bravo, 2009).

There are a number of models and several generalized techniques that have been specifically developed to predict behavior in choice situations. These models assume that each alternative in a choice set has a utility, or subjective value that depends only on the alternative and the consumers independently evaluate each available alternative and then choose the one with the highest overall utility (McFadden, 1980). This approach to studying consumer decisions, often attributed to economists and called rational choice theory, has contributed greatly to the prediction of consumer decisions (Bettman, Luce, and Payne, 1998). Rational choice models are typically associated with decision problems in economics in which choice is characterized as the "maximation of value" (Shafir et al., 1993, p.12). These models suggest that choice is influenced by a set of specific product attributes that can be valued with a utility score (Bravo, Won and Ferreira, 2009). The measurement models differ in their approach to the problem. This class of models takes either a compositional or decompositional approach.

With the compositional approach, a respondent's overall evaluation is "built-up" on the basis of the self-explicated weights. In other words, the compositional or self-explicated approach starts with the individual attribute of a product or service and combines them to build an overall preference. It constitutes a class of additive models in which the overall utility for a multiattribute alternative is computed as a weighed sum of that alternative's perceived attribute levels and associated value ratings. Examples of this class of compositional models are the multiattribute utility models. In these models, the consumer provides both the desired level and the relative importance of the attributes that underlie the multiattribute alternatives. These self-explicated measures are then combined to determine the overall preferences for alternatives with similar attributes (Huber, 1974). The most recent development in compositional multiattribute utility models is the Analytic Hierarchy Process (AHP) which through pair-wise comparisons "build-up" on the basis ratio level overall evaluations (Saaty, 1977, 1980).

The second general type of modeling approach for measuring rational choice is the decompositional method. In the decompositional approach respondents are presented with alternatives, defined in terms of a set of attributes, and asked for an overall evaluation (Baglione, 1994). With the decompositional approach the respondent's overall evaluation is broken-down into its constitutive components. In other words, decompositional models start with the consumer's overall evaluation of mutiattribute alternatives and decompose them into "part-worth" or values for individual attribute levels (Kara, 1993). The traditional conjoint models, in which the consumers' overall preferences are used to derive a set of part-worths, fall into this category (Green and Srinivasan, 1978).

The AHP and conjoint models being compared in this study are one of a large group of rational choice models. Theoretically, the AHP and the CA should be similar since both models are based on the additive linear utility function. The AHP is a linear, additive and compensatory compositional model of multiattribute decision making (Jensen, 1983). Thus it is in this sense, similar to conjoint analysis since both models usually assume a compensatory (additive) rule, where negative aspects of a product may be compensated for by other desirable qualities (Orme, 2006). Conjoint analysis begins with the assumptions of the random utility model (RUM). That is, the true utility associated with an alternative is viewed as a random variable. The premise underlying this method is that, when confronted with a purchase decision, consumers assign utilities to each alternative and then select the one with the highest derived utility among available alternatives (Baglione, 1994). This was illustrated in Figure 1.

Utility is decomposed into a deterministic component, the relative influence or each attribute on the overall utility, and the random component, the remainder which is assumed independent and normally distributed. The RUM is linear and additive and is expressed as:

$U_i = D_i + R_i$

Where

$D_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_j X_{iJ}$

i= Index of stimuli (or alternatives)

J= Number of attributes (j=1… J)

U_i= Utility of stimulus i

D_i= Deterministic component of the utility of stimulus i

R_i= Random component of stimulus i

$ß_j$ = Individual's weight for each of the J attributes

X_{ij}= the level of attribute j in stimulus i

It is clear from this representation that a compensatory rule is explicitly assumed. That is, for any given alternative, a weakness in one attribute can be offset, or compensated for, by a perceived strength in another attribute (Louviere, 1988a). Alternative 1, for example, may be preferred overall to alternative 2, even though the contribution of attribute1 to alternative 2 is greater than the contribution of attribute1 to alternative1. Additive models may have some difficulty fitting the data well if many respondents consistently use non-compensatory choice processes, where perceived weakness in one attribute cannot be offset or compensated for by a perceived strength in another attribute (Kara, 1993; Orme, 2006). Yet, even when respondents use non-compensatory processes, the results have been robust (Johnson and Meyer, 1984; Olshavsky and Acito, 1980; and Dawes and Corrigan, 1974).

Multiattribute Choice Models

The rational choice models are all based on the assumption that the buyer makes purchasing decisions based on an evaluation of the attributes of a product or service and a comparison of those evaluations (Kara, 1993). Multiattribute choice models have been developed along two main lines: compositional modeling and decompositional modeling. The simplest compositional multiattribute model is the weighted point model, which has also been used to describe consumer attitudes (Fishbein and Azjen, 1975). In this view, the AHP is a relatively new compositional approach to modeling multiattribute decisions. Alternatively, decompositional multiattribute models are used in conjoint analysis. In the following section, the two classes of models selected for comparison in this study will be discussed in more detail. They also will be compared to each other on a theoretical basis. This study begins with the conjoint analysis model.

Conjoint Analysis Model

The most popular of the multiattribute decision models is conjoint analysis. Conjoint analysis (CA), a decompostional model, has gained a great deal of academic and industry attention as a major set of techniques for measuring how consumers make tradeoffs in evaluating mutiattribute products and services. CA facilitates a respondent's estimation of attribute weights by requiring only overall evaluations instead of individual attribute weights (Baglione, 1994). It is based on the notion that for many consumers, multiattribute choices may be unmeasruable when examined individually according to each alternative's attributes, but they are measurable when considered jointly in an overall evaluation (Green and Rao, 1971). Over the past several years, conjoint analysis has been one of the most prominent methods for measuring customers' preference structures and has been widely used in marketing research practice (Meißner et al., 2008). Perhaps the most prominent lodging industry application of conjoint analysis was the study by Wind et al. (1989) that helped Marriott Corporation design its Courtyard by Marriott brand.

When comparing the validity of conjoint measurement with the AHP approach, one has to distinguish between different types of conjoint measurement methods which can lead to varying results in terms of validity. The number of product attributes selected must be reconciled with the characteristic of the given conjoint method: The full-profile conjoint analysis approach is ideal in the case of a maximum of six attributes, but if more than six attributes must be included, then the adaptive conjoint analysis is the appropriate method (Majláth, 2009). Though nowadays adaptive conjoint analysis and choice-based conjoint methods are very popular, sometimes it is more convenient to use the full-profile approach. Adaptive conjoint analysis must be computer-administered. The interview adapts to respondents’ previous answers, which cannot be done via the "paper and pencil" method. On the other hand, the choice-based conjoint method can be administered by personal computer or via paper and pencil, but results have traditionally been analyzed at the aggregate, or group, level. Aggregate-level analysis is useful for detecting and modeling subtle interactions that may not always be revealed with individual-level models. While these advantages seem to favor aggregate analysis from choice data, academics and practitioners have argued that consumers have unique preferences, and that aggregate-level models which assume homogeneity cannot be as accurate as individual-level models (Orme, 1996). Thus, the full-profile approach proved the better choice in this study, because it calculates a set of utilities for each individual. In the next section the CA is further described.

Assumption of CA. Two basic methodological assumptions are needed in conjoint analysis (Gil and Sa´nchez, 1997). First, a product or service can be described as a combination of levels of a set of attributes. Second, these attribute levels determines consumers' overall evaluation of the product or service. It is based on the assumption that all products are composed of attributes which may have two or more levels.

Generally the steps involved in a conjoint analysis can be summarized as follows (Green and Srinivasan, 1978):

1) Identification of attributes and attribute levels
2) Selection of a preference model
3) Selection of a data collection method,
4) Stimulus set construction,
5) Stimulus presentation,
6) Measurement scale for the dependent variable, and
7) Selection of an estimation method

Identification of attributes and attribute levels. One of the key assumptions underlying the methodology is that an individual's preference for an object can be decomposed into preference scores for components of the object. The identification or generation of the relevant attributes and attribute levels is essential in any rational model of consumer preference (Cattin and Wittink, 1982). Conceptually, the attribute choices for conjoint analysis follow from the customer's product concept choices within the market context. The researcher must identify the so-called the determinant attributes.

These attributes relate to preference and choice and distinguish the choice alternatives in meaningful ways. A very important issue in attribute selection is the trade-off between a realistic description of alternatives, which requires a large number of attributes, and the ease of the task for the respondent, which requires a small number of attributes. There are various ways to reduce the task complexity for a given number of attributes and attribute levels, such as fractional factorial designs (Green, Carroll, and Carmone, 1978; Green, 1974). It is essential that most researchers try to reduce the number of attributes to an essential set of relatively uncorrelated attributes that describe the product concepts in terms of customer choice criteria.

Selection of preference model. Green, Krieger, and Wind (2001) considered three utility (preference) models: (1) vector model, (2) the ideal-point model, and (3) part-worth function model. The vector model estimates the fewest parameters by assuming the linear functional form. The model can be represented as follows.

$$U_i = \sum_{j=1}^{J} \beta_j X_{ij}$$

Where

U_i= Utility for the *i*th stimulus (i=1…I)

$ß_j$ =Individual's weight for each of the J attributes (j=1…J)

X_{ij}=the level of the *j*th attribute for the *i*th stimulus.

In the ideal point model, U_i is negatively related to the squared distance d_i^2 of the location (X_{ij}) of the *i*th stimulus from the individual's ideal point (X_j) where is d_i^2 defined as:

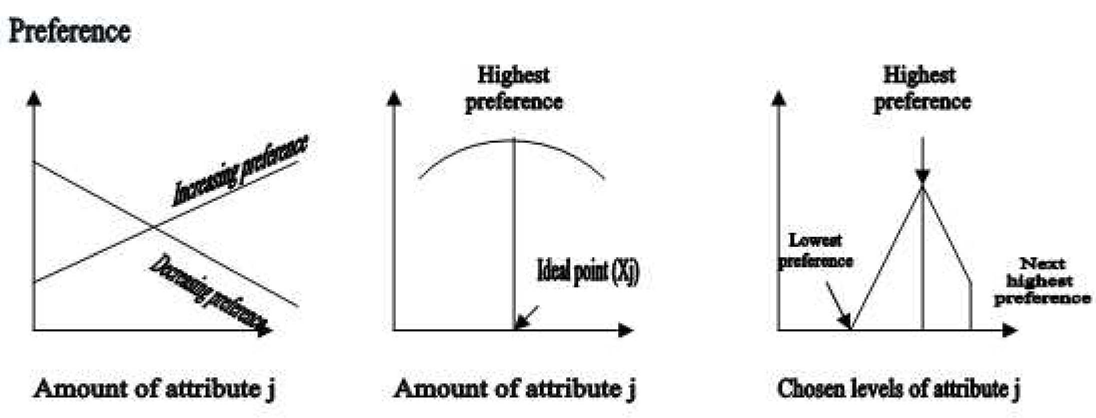
$$d_i^2 = \sum_{j=1}^{J} \beta_j (X_{ij} - X_j)^2$$

Thus, the stimuli closer to the ideal point will be more preferred.

Cattin and Wittink (1982) have found that the part-worth model is the most commonly used model in commercial applications. The part-worth model can be represented as follows:

$$U_i = \sum_{j=1}^{J} f_j (X_{ij})$$

Where f_j is the function denoting the part-worth of different levels of X_{ij} for the *j*th attribute. This model has received wide acceptance because of its ready interpretable part-worth function (Green and Srinivasan, 1990). It provides the greatest flexibility in allowing different shapes for the preference function. The linear model is a special case. Figure 2 illustrates what is meant by linear preferences, ideal point preferences, and discrete (part-worth) preferences. The third graph shows three part-worths.



Source: Green, Krieger, and Wind (2001).

Figure 2. Alternative models of preference.

Selection of a data collection method. The data collection methods in conjoint analysis have largely involved variations of two methods, the full-profile approach (Green and Rao, 1971) and the two-factor-at-a-time approach (Johnson, 1974). The full-profile approach utilizes the complete set of attributes to construct stimulus combinations. The major limitation of this method is the problem of information overload when the number of attributes is large. However, this problem can be dealt with by using a fractional factorial design. Its major advantage is that a more realistic description of stimuli given by defining the levels of each of the attributes. In the two-factor-at-a-time approach, the respondent is asked to rank the various combination of each pair of attribute levels from the most preferred to the least preferred. This method is easy to apply, but lack of realism, possibility of patternized responses, and a large number of combinations to evaluate are major disadvantages.

In summary, the trade-off approach is simpler but requires more evaluations. On the other hand, the full-profile approach when using a fractional factorial design requires fewer but more complex judgments by the respondent. Table 2 shows an illustration of the approach, as applied to consumer evaluations of steel-belted radial replacement tires. The table at the left shows the respondent rankings of each of the combination pairs of brand and treads life. On the right is one of the 18 four-factor stimulus cards that must be sorted by the respondent.

Table 2

Alterative Data Collection Methods

I. Two-Factor-at-a-Time Approach[a]				II. Full-Profile Approach (Sample stimulus card)[b]
	Tread life			Brand
Tire Brand	30,000 miles	40,000 miles	50,000 miles	*Sears*
Goodyear	8	4	1	Tread Life
Goodrich	12	9	5	*50,000 miles*
Firestone	11	7	3	Sidewall
Sears	10	6	2	*White*
				Price
				$55

Note. 1 denotes the best-liked combination and 12 denotes the least-liked combination for a hypothetical respondent.
Sources: [a]Green and Srinivasan (1978); [b]Green, Tull, and Albaum (1988).

Stimulus set construction. This step applies only to full-profile methods. The number of stimuli depends on the number of estimated parameters. If the number of parameters is large, fractional factorial designs (Green, 1974; Green, Carroll, and Carmone, 1978) can be used to reduce the number of combinations to a manageable set. When two or more attributes are highly correlated, creation of a superattribute, which represents both attributes, is suggested (Green and Srinivassan, 1978). Steckel, DeSarbo, and Mahajan (1990) provide another approach for maximizing "orthogonalness" subject to meeting various user-supplied constraints on the attribute levels that are allowed to appear together in full-profile descriptions. Krieger and Green (1988) and Wiley (1977) suggest methods for constructing stimulus sets for conjoint analysis that are Pareto optimal (i.e., no option dominates any other option on all attributes). Huber and Hansen (1986) and Green, Helsen, and Shandler (1988) report empirical results on the question of whether Pareto-optimally designed choice sets provide greater predictive validity than standard orthogonal designs in predicting a holdout set of realistic (Pareto-optimal) full profiles. The results are mixed. Whereas Huber and Hansen's study, utilizing paired comparison preference judgments, suggests that Pareto-optimal choice sets predict better, Green, Helsen, and Shandler's study, utilizing full profiles, indicates the opposite. More recent studies by Moore and Holbrook (1990) and Elrod, Louviere, and Davey (1989) support and extend the findings of Green, Helsen, and Shandler's study. So far, the weight of the evidence suggests that orthogonal designs are very robust even when prediction is made on Pareto-optimal choice sets (Green and Srinivasan, 1990).

Stimulus presentation. Presentation of the stimuli involves three approaches: verbal description, paragraph description, and pictorial representations. Verbal and paragraph descriptions of alternatives are the most commonly used methods (Cattin and Wittink, 1982). These procedures are convenient, straightforward, and inexpensive. However, Green and Srinivasan (1990) reported an increasing use of the pictorial presentation format. These kinds of presentations make the task more interesting to the respondent. They also provide easier and potentially less ambiguous ways of conveying information and thus allow a greater number of attributes to be included in study. In addition to the three approaches mentioned for stimulus presentation, there is also some evidence that conjoint methodology is increasingly being used with actual physical products as stimuli to make the choice process more realistic. Hence, the selection of a presentation method has to be done according to the objectives of the conjoint analysis (Scholz, 2008).

Measurement scale for the dependent variable. The measurement scale can be classified as non-metric (rank order, paired comparison) or metric (ratio scales, rating scales). Traditionally, conjoint data have been collected on a non-metric scale. However, Wittink and Cattin (1989) reported a decrease in the relative popularity of rank order response scales. Rating scales now account for almost a half of all commercial applications.

Estimation methods. Estimation methods in conjoint analysis can be broadly classified into three categories (Green and Srinivasan, 1978):

1. Dependent variable is at most ordinally scaled
    a. MONANOVA which is restricted to the part-worth function model (Kruskal, 1965; Kruskal and Carmone, 1969).
    b. PREFMAP which can be used with either the part-worth function model or the vector model (Carroll, 1972; Carroll and Chang, 1967).
    c. Johnson's non-metric trade-off procedure which can be used with either the part-worth function model or vector model (Nehis, Seaman, and Montgomery, 1976).
    d. LINMAP which uses linear programming, rather than classical statistical methods employed by the other approaches (Srinivasan and Shocker, 1973).
2. Dependent variable is intervally scaled
    a. The Ordinary Least Squares (OLS) method which has an important advantage of providing standard errors for the estimated parameters (Johnston, 1972).
    b. Minimizing Sum of Absolute Errors (MSAE) which permits the researcher to impose a priori constraint on the estimated parameters (Srinivasan and Shocker, 1973).
3. Relate paired comparison data to a choice probability model.

a. LOGIT which is preferred when the attribute weights exhibit a lexicographic structure (McFadden, 1976; Green and Carmone, 1977; Punj and Staelin, 1978).

b. PROBIT which is particularly suited to the case where a dichotomous intention-to-buy scale is used (Goldberger, 1964; Rao and Winter, 1977).

The choice between (1) and (2), and (3) should depend on the scale properties of the dependent variable. Wittink and Cattin (1989) reported that the OLS was the most frequently used method in the 1980s. This trend is consistent with the results indicating that metric analysis is robust regardless of the measurement scale for the dependent variable (Carmone, Green, and Jain, 1978). Lund, Malhotra, and Smith (1988) also reported that the OLS regression method predicted better than either MONANOVA or LINMAP.

The above steps in conjoint analysis and the alternative methods of implementing each of the steps are summarized in the Table 3.

Table 3

Steps involved in Conjoint Analysis

Step	Alternative methods
1. Identification of attributes and attribute levels	Focus-groups and expert panel, surveys of customers, in-depth consumer interviews, expert questionnaires, early studies etc.
2. Selection of a model of preference	Vector model, ideal-point model, part-worth function model
3. Data Collection method	Two-Factor-at-a-Time (Trade-off Analysis), Full Profile (Concept Evaluation)
4. Stimulus set construction for the full-profile method	Fractional factorial design, random sampling from multivariate distribution, Pareto-optimal designs
5.Stimulus presentation	Verbal description (multiple cue, stimulus card), paragraph

	description, pictorial or three-dimensional model representation
6. Measurement scale for the dependent variable	Paired comparisons, rank order, rating scales, constant-sum paired comparisons, category assignment (Carroll, 1969)
7. Estimation method	MONANOVA, PREFMAP, LINMAP, Johnson’s nonmetric tradeoff algorithm, OLS, LOGIT, PROBIT

Source: Green and Srinivasan (1978).

The Analytic Hierarchy Process (AHP)

The most recent development in multiattribute decision making is the analytic hierarchy process (AHP). The AHP, a relatively new compositional approach, has been of substantial impact in business research and particularly in managerial decision making for a long time (Meißner et al., 2008). Historically, the AHP has been applied to the problem of multiattribute decision making of an economic and strategic nature and its principal application is in decisions in which subjective criterion play an important role (Schoner and Wedley, 1989). However, the practical nature of the method, suitable for solving complicated and elusive decision problems, has led to applications in highly diverse areas and has created a voluminous body of literature (Zahedi, 1986).

Assumption of AHP. The AHP assumes that decisions are reached in a hierarchical fashion. It is used to determine the relative importance of a set of criteria or attributes using a hierarchical structure (Saaty, 1977). The measure of preference obtained by the AHP in a multiple criteria decision-making problem under certainty satisfies the definition of an additive value function (Kamenetzky, 1982). It involves an importance-ratio assessment procedure and uses a hierarchy to establish preferences and orderings (Dyer, Fishburn, Steuer, Wallenius, and Zionts, 1992).

The approach can be explained in the following four major steps (Johnson, 1980):

1) Construction of the selected decision problem into a hierarchy;
2) Evaluation of the elements in each level of the decision hierarchy by pair-wise comparisons;
3) Estimation of the relative weights and evaluation of the consistency of judgment; and
4) Synthesis of the relative weights.

Construction of the AHP hierarchy. The most important step in AHP is the decomposition of a complex problem into a hierarchy of interrelated levels. Each level consists of a few manageable elements and each element then decomposed into another set of elements (Saaty 1977). The overall decision objective, such as the objective of making the best decision (or selecting the best alternative) lies at the top of the hierarchy. The decomposition process continues down to the most specific courses of action considered, which are represented at the lowest level of the hierarchy. The lower levels of the hierarchy contain attributes that contribute to the quality of the decision. The last level of the hierarchy contains decision alternatives or selection choices. For instance, in the decision problem of selecting a restaurant, choosing a restaurant comprises the objective in Level 1. Level 2 consists of menu price, location, and service quality. Finally, the selection alternatives (restaurants) constitute the last level, Level 3; in this example (see Figure 3).

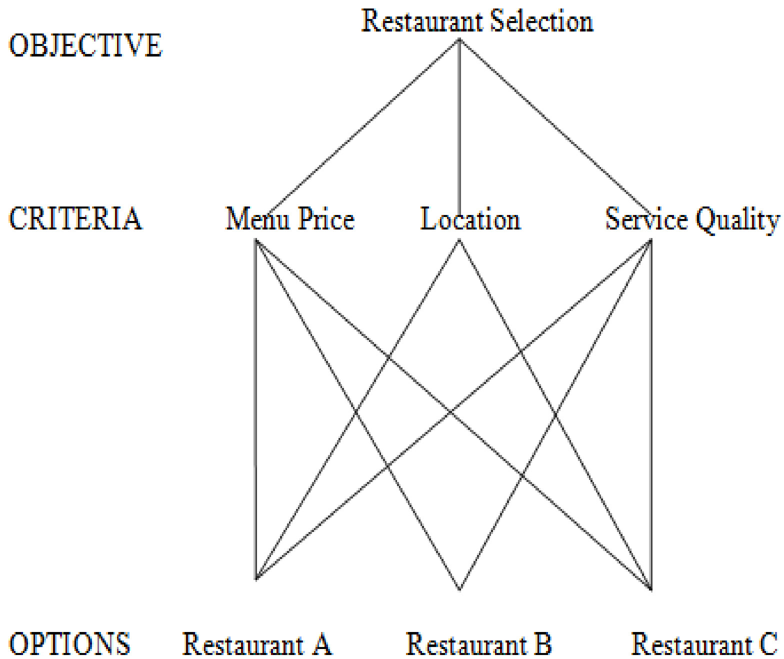


Figure 3. Restaurant decision problem.

Estimation of elements. A measurement methodology is used to establish priorities between the elements in each level of the hierarchy. The pair-wise comparisons are made, using a semantic descriptive 9-point intensity-of-importance scale, with respect to each of the elements at a higher level. The 1-9 point scale provided by Saaty (1996) is given in Table 4. The fundamental scale shows the meaning of numbers from 1 to 9. These numbers indicate the intensity of the relationships between the elements. The research results maintain that the short term memory of the human brain and its ability to internalize can evaluate approximately 7±2 situations (Sönmez and Hacıköylü, 2012). The 1-9 scale reflects the intensity of personal choices that have been developed by Saaty and many AHP users. However, the fundamental scale can be altered to suit an individual's needs, and it can deal with great amounts of information. It is the ratio scale that distinguishes AHP from the traditional decision analysis methods (Wind and Saaty, 1980). A matrix of pair-wise comparisons is constructed by reference to the semantic scale. Comparisons ask the respondent which of two criteria are more important and to what extent they are important. The ratio scale is used in determining the weights of the criteria through pair-wise comparisons (Scholl et al., 2005).

Table 4

Saaty's Fundamental Scale

Intensity of importance or preference	Definition	Explanation
1	Equal Importance or preference	Two activities contribute equally to the objective
3	Moderate importance or preference	Experience and judgment slightly favor one activity over another
5	Strong importance or preference	Experience and judgment strongly favor one activity over another
7	Very strong or demonstrated importance or preference	An activity is strongly favored and its dominance is demonstrated in practice
9	Extreme importance or preference	The evidence favoring one activity over another is the highest possible order of affirmation
2,4,6,8	Intermediate values between the two adjacent judgments for when compromise is needed.	

If alternative *i* (row element) is preferred to alternative *j* (column element) then the (*i*,*j*)'th cell of the matrix measures the strength of preference for the *i*'th over the *j*'th and cell (*j*.*i*) is the reciprocal of that number. In the above restaurant example, the goal is to choose a restaurant based on three attributes: menu price, location and service quality. In step one the relative importance of the three factors is determined by comparing them two at a time (3 comparisons). If location is moderately more important than menu price it will be given a 3 by the respondent. Consequently a comparison of menu price to location is assigned a 1/3. In step two the restaurant alternatives are compared with respect to each of those factors. If Restaurant A is strongly more preferred to Restaurant

B with respect to menu price it will be given a 5 and the comparison of Restaurant B to Restaurant A will be 1/5 or 0.2 with respect to the same attribute. When compared to itself, each element has equal importance. Diagonal elements of the input matrices, always equal one, and the lower triangle elements of the comparison matrix are the reciprocal of upper triangle elements or vice versa. Thus, pair-wise comparisons data are collected for only one half of the matrix elements, excluding the diagonal elements. The complete model for the restaurant choice problem will require an aggregation of importance weights and restaurant preference matrices. The details of estimation and synthesis are presented in Appendix A.

Conceptual Comparison of the AHP and CA

The two classes of choice models, AHP and CA, differ in several ways. Theoretically, the AHP process relies on multiattribute value theory, whereas CA relies on random utility theory, while both models are based on the additive linear utility function. And both models can be said to rely on utility theory and results can be compared. To fulfill the assumptions of the additive utility model that AHP and CA employ, attributes have to be independent from each other and the attribute levels must be mutually exclusive and exhaustive (Orme, 2002b).

While the AHP follows a compositional approach CA follows a decompositional approach. This produces different respondent tasks. In the AHP the respondent must evaluate each alternative against other alternatives in terms of which is more important and by how much. This is a ratio-scaled judgment. In conjoint analysis the task is to rank or rate a complete product or package and is thus an ordinal or interval type of measurement.

The CA model usually uses an experimental design in constructing product profiles. The algorithm can estimate the part-worth of attribute levels for an individual as well as a group of respondents. The AHP does not require an experimental design and utilities are estimated at the individual respondent level. The CA model usually uses the least squares or maximum likelihood method of estimation, whereas AHP usually uses eigenvalue (or the right eigenvector) estimation.

The two models also differ in terms of the algorithms and the choice processes used. CA models frequently use a full-profile approach for data collection that requires alternative processing. It assumes that attributes are evaluated simultaneously. This is considered a more realistic description of stimuli. CA estimates individual utility functions that are used to predict choice behavior. However, AHP uses an attribute processing approach. This two-factor-at-a-time approach is easy to apply but lack of realism is its major disadvantage. It uses a hierarchical evaluation of attributes and estimates the priority weights at each level.

In summary, conjoint analysis models have been very popular in marketing and can be classified in the compensatory models category. There is an increasing trend in the application of AHP in marketing and consumer behavior. AHP can be also classified in the compensatory (additive) utility models but differs from conjoint analysis in several ways. In order to facilitate a theoretical comparison, a summary of similarities and difference between AHP and CA can be seen in Table 5 (cf. Mulye, 1998; Kara, 1993; Helm et al., 2003, 2004b; Kallas et al., 2011). These theoretical differences are the starting point for further comparison of the models.

Table 5

Conceptual Comparison of AHP and CA

	AHP	CA
Assumptions	Preferential independence of the attributes	Preferential independence of the attributes
Utility model	Weighted additive utility model	Additive part-worth model
Theoretical background	Compositional approach Weights estimates rely on the Multiattribute Value Theory, representing the attributes’ rank	Decompositional approach Coefficient estimates rely on the Random Utility Theory, representing the attributes’ rank
Scale used	Ratio scaled	Rank order, rating or constant-sum scale
Algorithm	Attribute processing approach Estimates priority weights	Alternative processing approach Estimates individual utility function
Experimental design	Complete design and some Incomplete Pairwise Comparisons (IPC)	Fractional factorial design
Estimation model	Eigenvalue estimation Approximation method Mean transformation Row/column geometric mean The harmonic mean	Least square estimation LINMAP MONANOVA PREFMAP MSAE
Choice process	Hierarchical evaluation of attributes Hierarchical evaluation of alternatives	Simultaneous evaluation of attributes Simultaneous evaluation of alternatives
Data collection (Interview expense)	Paired comparison of attributes and alternatives	Complex evaluation of complete stimuli (ranking, rating or paired comparisons)
Results	Relative utilities of all attribute levels; attribute weights	Part-worths of all attribute levels
Primary focus in most applications	To aid the decision making process	To measure consumer values, prediction of brand choice
Application range	Selection problems and/or design problems	Design problems

Sources: Mulye, (1998); Kara, (1993); Helm et al., (2003, 2004b); and Kallas et al., (2011).

Determining Attributes in Preference Measurement

The first stage in the design of a conjoint analysis study as well as an AHP study is the selection of the attributes. Regardless of the approach, to conduct an effective study, correctly identifying the relevant attributes is key. The chosen attributes should be relevant for respondents, since the conclusions drawn about consumer choice would change if we ignore the existence of important factors (Lancaster, 1991). The presence of either too many or irrelevant attributes may lead to an overly complex decision for respondents and may result in more inconsistent and random responses (Bennett and Blamey, 2001). Typical CA studies employ up to six attributes (Helm et al., 2004a; Green and Srinivasan, 1990; Orme, 2002a). Using more attributes may overstrain respondents leading to simplification strategies that result in distorted preference structures (Green and Srinivasan, 1990). AHP allows more attributes within a hierarchy, but in order to conduct a fair comparison only four attributes are observed.

Firstly, a set of attributes and attribute levels has to be defined. This step is crucial for any preference elicitation technique because reliable and valid results can only be obtained if the determinant attributes and attribute levels are evaluated (Helm et al., 2004a). So one main focus of any preference measurement study should be the investigation of determinant attributes and attribute levels (Orme, 2002b). According to Helm et al. (2004a), there are several sources of information about structuring sets of attribute, e.g. focus groups, surveys of customers, in-depth consumer interviews, expert questionnaires, and early studies. In this study the key attributes of customer value related to hotel brand equity will be extracted from early studies based on brand equity.

In the following sections, a brief overview of the meaning of brand equity is provided prior to exploring the application of brand equity research in the hotel context. These sections will review the existing literature and research in the field of hotel branding. The key components of brand equity are introduced and their application in the hotel context is considered.

Brand Equity

To date, brand equity has been comprehensively studied and generally accepted as a precious asset for organizations. Ever since Aaker (1991, 1996) identified the explicit dimensions of brand equity and Keller (1998) identified the sources of brand equity, the concepts of brand loyalty, brand awareness, perceived quality, brand associations, and brand image have been well-associated with brand equity and widely tested empirically in related studies (Low and Lamb, 2000; Kim and Kim, 2005; Boo et al., 2009).

As mentioned earlier, there exist so many definitions about brand equity according to different researchers or varied market situations, but basically Farquhar's definition, in which he describes that brand equity is the added value that a brand gives a product, is the most widely accepted (Farquhar, 1989, p.24). In this perspective, hotel brand equity can be seen as the added value to hotel products by such brands as Sheraton, Marriot, Hilton, Hyatt, and Westin (Kim and Kim, 2006).

Service Branding and Hotel Brand Equity

Many studies have suggested that brand equity should be an important research domain because of its strong association with marketing strategy and firms' sustainable competitive advantage (Keller, 2003; Pappu et al., 2005; Tasci et al., 2007). Generally, brand equity has been accepted as the primary source of capital for many industries (Bendixen et al., 2004) and considered to create customer loyalty, enhance consumer trust, and reduce the perceived risk, especially in services (Lee and Back, 2008).

Since brand equity is now an important subject in brand management and inseparably linked to customer value (preferences) (Lawer and Knox, 2006), its role in hospitality firms is deemed very important (de Chernatony et al., 2005; Kay, 2006). Zhou et al. (2008) have declared that service industries are highly customer value-oriented. While research claims that the number of empirical studies connecting brand equity of service firms with their managerial strategies is increasing (Keller, 2003), relatively limited empirical evidence can be found with respect to the consumer-based equity of service brands in the hospitality industry (Konecnik and Gartner, 2007) due to the fact that most conceptual and empirical research has focused on products and not services (Boo et al., 2009).

Few articles regarding brand equity of hospitality can be found, especially focusing on restaurants and hotels. Many early studies are related to qualitative theoretical analysis; researchers began to use the quantitative approach in the empirical studies. For instance, a study by Cobb-Walgren et al. (1995) is the first study which adopts Aaker (1991) model to measure customer-based brand equity, comprising four dimensions: brand awareness, brand association, perceived quality, and brand loyalty. Prasad and Dev (2000) established a hotel brand equity index model, including the two dimensions of brand awareness and brand performance. Bailey and Ball (2006) rethought the meanings of hotel brand equity. They believe it is necessary to integrate the different views to study hotel brand equity, and they proposed a new concept for hotel brand equity based on the point of view of both the property owner and the customer. The research by Kim, Kim and An (2003); Baldauf, Cravens, and Binder (2003); Kim, Dimicelli, and Kang (2004); Kim, Jin-Sun and Kim (2008); Kim and Kim (2004); Kim and Kim (2005); Kim and Kim (2007); Atilgan, Aksoy, and Akinci (2005); Kayaman and Arasli (2007); Nel, North, Myburg, and Hern (2009); Sun and Ghiselli (2010); So and King (2010); and Zhou and Jiang (2011) concentrated on the brand equity of hotels and restaurants (see Table 6). By considering both the literature review and Aaker's (1996) fundamental concept of brand equity, it can be concluded that hotel brand equity is composed of four major components: awareness, association/image, perceived quality, and loyalty. Each component is discussed briefly below.

Table 6

Research on Hotel Brand Equity Dimensions

Brand equity dimensions	Context	Researcher
Brand loyalty, brand awareness, perceived quality, brand image	Impact of consumer-based brand equity on firms' financial performance in the luxury hotel sector	Kim, Kim, and An (2003)
Brand awareness, perceived quality, brand loyalty	Impact of perceived brand equity on brand profitability, brand sales volume, perceived customer value and purchase intention in the value hotel chain.	Baldauf, Cravens, and Binder (2003)
Brand awareness: top of mind brand, brand recall, brand recognition	Measuring brand equity of restaurant chains as the monetary equivalent of the total utility a consumer attaches to a brand.	Kim, Dimicelli, and Kang (2004)
Brand loyalty, perceived quality, brand awareness, brand image	Relationship between brand equity and firms' financial performance in luxury hotels and chain restaurants.	Kim and Kim (2005)
Brand association, perceived quality, brand awareness, brand loyalty	Relationship between hotel's equity dimensions in mid-scale hotels	Kim, Jin-Sun, and Kim (2008)
Brand loyalty, perceived quality, brand awareness, brand image	Relationship between four components of brand equity and restaurant firms' performance	Kim and Kim (2004)
Brand association, perceived quality, brand awareness, brand loyalty	Determinants of mid-scale hotel brand equity	Kim and Kim (2007)
Brand loyalty, perceived quality, brand awareness, brand association	Four determinants of overall brand equity in the beverage industry	Atilgan, Aksoy, and Akinci (2005)
Perceived quality, brand loyalty, brand image, brand awareness	Interrelations of the four brand equity components in five-star hotels	Kayaman and Arasli (2007)
Brand image, brand awareness, brand loyalty, perceived quality	Comparing brand equity across selected hotel brands (low, medium or high-priced)	Nel, North, Myburg, and Hern (2009)
Brand awareness, brand association, brand quality, brand loyalty	Interrelationship among brand equity dimensions in the lodging industry	Sun and Ghiselli (2010)
Brand awareness, brand meaning	Building and measuring hotel brand equity	So and King (2010)
Brand loyalty, perceived quality, brand awareness/associations	Impact of customer-based brand equity on perceived value and revisit intentions in the budget hotel segment	Zhou and Jiang (2011)

Dimensions of Brand Equity

To operationalise the brand awareness, brand loyalty, and brand image dimensions suggestions from Kim and Kim (2006) study were employed.

Brand awareness. Brand awareness is “the ability for a customer to recognize or recall that a brand is a member of a certain product category” (Aaker, 1991, p. 91). Brand awareness can be viewed as the strength of a hotel brand’s presence in the consumer’s mind.

Brand image. Keller (1993, p. 3) defined brand image as “perceptions about a brand as reflected by the brand associations held in consumer memory.” Hotel brand images are the memories related to a hotel brand. It can be a set of meaningful associations with the hotel. Brand image refers to brand perceptions projected by these associations.

Perceived quality. Perceived quality is “the customer’s judgment about a product or service’s overall excellence or superiority” (Zeithaml, 1988, p. 3). Perceived quality of hotel brand can be defined as the consumer’s subjective evaluation of a hotel product or service quality.

Brand loyalty. Brand loyalty is defined as “the attachment that a customer has to a brand” (Aaker, 1991, p. 65). Hotel brand loyalty can be regarded as the preference and continued revisiting of a hotel as a result of satisfaction with the service or products of the hotel.

In summary, strong brand equity means that customers have high brand-name awareness, maintain a favorable brand image, perceive that the brand is of high quality, and are loyal to the brand (Kim and Kim, 2004). It is important for hotel companies to focus their efforts on building up these aspects of brand equity to varying degrees that suit their target customers in order to strengthen their brand positioning in the consumer's mind to encourage consumption such as a hotel stay (Huang, 2010). Table 7, below, shows the attribute levels identified from the literature review and grouped according to the dimensions described above to form the basis of this empirical investigation. After selection of attributes and attribute levels, these attributes and their respective levels were then used to compile hypothetical profiles or scenarios. The research design is covered in the following sections. The type of data and how the data was collected are also explained in that section.

Table 7

The List of Four Attributes as well as Detailed Attribute Levels' Descriptions

Attributes	Levels	Descriptions	Drawn from the literature
Brand awareness	Advertising	Extent to which the intended targeted customers are aware of an advertising message	Cobb-Walgren, Ruble, and Donthu (1995); Netermeyer et al., (2004); Kim, Dimicelli, and Kang (2004)
	Top-of-mind brand	The hotel name that is foremost on the mind of customers	
	Brand popularity	The degree to which consumers feel the brand is popular with and used by others.	
	Brand familiarity	The degree to which consumers are familiar with the brand name	
Brand image	Clean image	A very clean and orderly image	Kim and Kim (2005); Kim, Jin-Sun, and Kim (2008); Kim and Kim (2004); Sun and Ghiselli (2010)
	Elegant atmosphere	A luxurious, stylish, prestigious, and suitable place for high class clientele	
	Feels like home	A comfortable, quiet, and restful image	
	Good value	The room rate and other fees for using the facilities at the hotel are reasonable to pay	
Perceived quality	Error-free service (Assurance)	Staff of the hotel has knowledge and confidence to answer guests.	Kim and Kim (2004); Kim and Kim (2005); Mola and Jusoh (2011) Markovic and Raspor (2010); Malik, Naeem, and Nasir (2011); Parasuraman, Zeithaml, and Berry (1988)
	Prompt service (Responsiveness)	Service without delay (e.g. promptness of check-in and check-out)	
	Courteous service (Empathy & Tangibles)	Friendly, polite, and respectful service with neat, clean and appropriately groomed appearance	
	Reliable service (Reliability)	Handling of complaints and problems sincerely	
Brand loyalty	Friends' recommendation	Positive comments about the hotel brand from other people	Kim and Kim (2005); Kim, Jin-Sun, and Kim (2008); Kim and Kim (2004); Sun and Ghiselli (2010)
	Frequent customer	The hotel brand would always be my first choice compared to other hotels (e.g. frequent guest programs)	
	Previous experience	The hotel brand that never disappoints me, guarantees satisfaction, and always meets my expectations	

CHAPTER 3

SURVEY METHODOLOGY

In Chapter 2, the two choice models used in this study, AHP and CA were fully described and compared on a theoretical basis. In addition, the chapter identified the key attributes of hotel brand equity, namely, brand awareness, brand image, perceived quality and brand loyalty, based on the literature review (e.g., Aaker, 1996; Keller, 2003). In this chapter the distribution and the construction of the surveys, as well as the delimitations of the study will be discussed. Next, this chapter describes the design of the empirical study and discusses the steps involved in the implementation of conjoint analysis (CA) followed by analytic hierarchy process (AHP), including attribute selection, experimental design, survey design, data collection, and model specification.

Sample and Procedure

In order to be able to compare the validity of both AHP and CA across offline and online data collection modes, it had to be assured that the study designs of the surveys were identical. Thus the offline (paper-based) survey and the online survey consisted of the same questions and every attempt was made to keep the instrument as equivalent as possible across the two data collection methods. In both versions at the beginning of the questionnaire one screening question was asked to identify if respondents had stayed at a hotel in the last 12 months. Those who met this criterion were allowed to continue with the survey.

For the offline survey, a convenience sample of college students were recruited and interviewed using a paper-and-pencil questionnaire. The survey was administered to undergraduate students from three different tourism classes on October 24, October 29 and November 15 in 2012 from a large Arizona university. Contacts were made with the instructors for approval of the survey. The interviewer then communicated with the participants concerning the purpose and description of the survey. Then, they were instructed on how to complete the forms correctly and were given extra credit points for participating. Additionally, they were ensured that their participation would be confidential and anonymous.

For the online survey, the population of interest consisted of domestic travelers who stayed in economy, mid-range or upscale hotels. To obtain a representative sample (or as close to it as possible) as well as broad geographic representation, respondents were sourced from a well-reputed professional sampling company's online consumer panel. The online consumer panel contained a sample of respondents balanced according to US census data validated by various demographics criteria. The survey was deployed on their website from March 5, 2013 to March 7, 2013. The database provided by the online research company is designed to be representative of the U.S. population, and respondents have agreed to be contracted for surveys. The participation incentive was offered to potential survey respondents and they were paid $8 each as compensation for their time.

Delimitations

The scope of this study is delimited to those who have stayed in economy, mid-range or upscale hotels in the past 12 months and are over 18 years of age. Thus, in both versions respondents were screened prior to participating in the survey. The primary screening criteria required respondents to be over 18 years of age and have spent at least one night in a hotel in the previous 12 months.

To achieve a comparison of the two data collection methods, the offline and online samples were recruited separately. However, the two samples did not consist of the same test persons. In particular, a student (offline) sample, as a result of limited resources both in time and finances, is not considered to be representative of the target population. Nonetheless, this study chose college students for several reasons. Given the experimental nature of the study, the use of a student sample allowed for control of the administration of the study tasks. Also, according to Helm et al. (2004b) a general knowledge of the decision problem is a pre-condition when selecting respondents for preference measurement studies. From this perspective, in the current study on hotel branding the hotel design problem was considered to be particularly suitable and relevant for those enrolled in tourism-related majors because the tourism students interviewed in this study have enough knowledge to give valid answers and are familiar with the problem. Convenience samples of college students are often utilized for validity testing (Klein et al., 2010). The literature reflects an extensive use of student samples (e.g., Mulye, 1998; Helm et al., 2003, 2004a; 2004b; Scholl et al., 2005; Klein et al., 2010).

Questionnaire Design

In both versions, the questionnaire consists of four parts. In part 1 an introduction of the decision/design problem is given by describing and explaining the attributes and attribute-levels along with a cover letter including the propose of the study, the expected amount of time required for completion of the questionnaire, and the confidentiality of the responses. Part 2 and 3 consist of the AHP and CA questionnaire where the order was varied systematically to compensate for order effects. Finally, the respondents had to answer ten choice tasks including four alternatives in part 4. At the end of the survey some additional questions on socio-demographics as well as traveling characteristics were asked. Further questions concerning hotel involvement and knowledge were added.

In order to rule out the possibility of potential demographic influences, all respondents answered both AHP and CA questionnaires, but in randomly determined order. Because the predictive validity of the (holdout) choices for each method may depend on the order in which the methods were applied, the present study used both orders. Thus, in order to account for and equalize potential order effects about half of the respondents completed the AHP task first, while the other half did the CA task first. This manipulation allows the researcher to determine the extent to which task order effects exist and to examine the results for their presence. Additionally, respondents in the offline (paper based) sample requested to take the time they needed for answering the questions, whereas for the online sample the time taken to complete the survey was recorded by the online research company.

Conjoint Experimental Design

There are important steps in designing a conjoint analysis which must be followed. The first step was to select the most important attributes - those which were most frequently chosen in the previous research (see Table 8). Next step was to choose an appropriate conjoint analysis approach. Based on the review of the previous studies where the conjoint analysis was used (Mulye, 1998; Helm et al., 2003, 2004a, 2004b; Scholl et al., 2005; Klein et al., 2010), full-profile approach, the most common method of data collection in conjoint research (Gil and Sa´nchez, 1997), was chosen to be used. In CA, it is possible to assess what consumers truly value in a product or a service (estimating the utilities). CA assumes that all of the attributes are independent from each other. In general, conjoint is an appropriate approach when the number of attributes is not very large. It is suggested to limit the number of presented stimuli to 20 (Voeth, 2000). Respondents in the present study were requested to rank, rate or score a set of profiles (cards) according to their preference, one at a time. In CA, each profile describes a complete product or service consisting of a different combination of levels of all attributes. The attributes and their respective attribute levels identified from the literature review used in the current study are summarized in table 8.

Table 8

Attributes and Their Levels

Brand equity attributes	Levels	Data type
Brand Awareness	Advertising Top-of-Mind brand Brand popularity Brand familiarity	Discrete
Brand Image	Clean Image Elegant atmosphere Feels like home Good value	Discrete
Perceived Quality	Error-free service Prompt service Courteous service Reliable service	Discrete
Brand Loyalty	Friends’ recommendation Frequent customer Previous experience	Discrete

Once the attribute and attribute levels have been selected, they must be combined into hypothetical products for respondents to rate or rank. The attributes and levels in Table 8 gave rise to 192 possible profiles (4^3 x 3). Based on previous research, Johnson and Orme (1996) and Pignone et al. (2012) suggest that it would be a tedious task for respondents to answer all the questions when the number of profiles is too high. So in order to make the task easier for respondents and to reduce the possible number of profiles to a manageable level, while still allowing the preferences to be inferred for all of the combinations of levels and attributes, CA uses what is termed as fractional factorial design to present a suitable fraction of all possible combinations of profiles. The resulting set is called orthogonal array. Orthogonal array/design considers only the main effect of each attribute level, and not the interaction effects between attributes. Using the orthogonal array experimental design, from all the possible combinations of attributes and levels, 16 combinations were chosen as stimuli to be used in this research, which are shown in Table 9. Four control stimuli (holdout tasks) were added to the given design. These four hold-out stimuli were not used by the conjoint procedure for estimating the utilities but used to assess the validity of the utilities in prediction.

Table 9

The Results of Orthogonal Array

Profile	Brand Awareness	Brand Image	Perceived Quality	Brand Loyalty
1	Brand familiarity	Elegant atmosphere	Courteous service	Previous experience
2	Advertising	Good value	Reliable service	Previous Experience
3	Advertising	Elegant atmosphere	Prompt service	Friends' recommendation
4	Brand popularity	Good value	Prompt service	Frequent customer
5	Top-of-Mind brand	Clean Image	Prompt service	Previous experience
6	Brand popularity	Elegant atmosphere	Reliable service	Friends' recommendation
7	Brand familiarity	Good value	Error-free service	Friends' recommendation
8	Brand familiarity	Feels like home	Prompt service	Friends' recommendation
9	Advertising	Clean Image	Error-free service	Friends' recommendation
10	Brand popularity	Clean Image	Courteous service	Friends' recommendation
11	Top-of-Mind brand	Feels like home	Reliable service	Friends' recommendation
12	Advertising	Feels like home	Courteous service	Frequent customer
13	Top-of-Mind brand	Elegant atmosphere	Error-free service	Frequent customer
14	Top-of-Mind brand	Good value	Courteous service	Friends' recommendation

15	Brand popularity	Feels like home	Error-free service	Previous experience
16	Brand familiarity	Clean Image	Reliable service	Frequent customer
17*	Top-of-Mind brand	Clean Image	Error-free service	Friends' recommendation
18*	Brand popularity	Clean Image	Courteous service	Previous experience
19*	Brand popularity	Elegant atmosphere	Courteous service	Friends' recommendation
20*	Advertising	Feels like home	Error-free service	Friends' recommendation

Note. *Holdout profiles

Sixteen stimulus cards were prepared; each card contained a combination of attributes from the orthogonal array. Within the conjoint task, respondents had to rank the sixteen combinations according to their preferences using 1 and 16 to indicate highest and lowest preference, respectively. Ranking is often used since it provides similar results compared to ratings. Carmone et al. (1978) compared the overall conjoint model goodness of fit under several forms of input data (raw data, rankings, and six-point rating scales), and found conjoint analysis to provide robust results regardless of the type of input data scales, with superior recoveries for rankings in some cases. On this issue as advised by Mulye (1998) the present study followed the conjectures of Green and Srinivasan (1978) that ranking scales may be more reliable than rating scales for several reasons. First, it is easier for a respondent to express which option is preferred over another option than to indicate the magnitude of the preference. Second, the conjoint axiom of strict ranking of alternatives (Louviere, 1988b) is likely to be satisfied better by the ranking task than the rating task. Third, the ranking scale is a comparative scaling procedure requiring consideration of all the profiles in the evaluation task, while in the rating task respondents can give judgments independent of the other profiles in the experiment. As a result, the rating task results in a large number of ties and at times can lead to intransitive judgment if respondents do not follow the instruction of considering all profiles in their task (Mulye, 1998). Ties also occur due to the limitations imposed by the gradation of a rating scale (Mulye, 1998).

The preference model chosen is the part-worth function model, which is suitable for discrete data. Ranking data provided by the respondents were analyzed with the use of ordinary least squares regression because of its robustness most common for metric as well as for ordinary measurement scales (Wittink and Cattin, 1981; Mishra et al., 1989). This is confirmed by the findings of Green and Srinivasan (1978) and Wittink and Cattin (1982) that there is little difference between ordinary least squares and monotonic regression results.

AHP Experimental Design

In AHP, first, a decision problem, e.g. determining the individually most preferred alternative from a given set of products/concepts, is modeled as a hierarchy. In other words, all attributes being relevant for judging alternatives have to be arranged within a hierarchy. At the top level, the main objective of the decision problem has to be specified and decomposed into several second level attributes and corresponding attribute levels on the third level. Further, Helm et al. (2003) explains that there are two types of hierarchy: complete and incomplete hierarchy. In the former case different alternatives (products or concepts or rather “stimuli" in the CA terminology) are considered at the bottom level of the hierarchy, whereas in the latter attribute levels are shown. The use of incomplete hierarchies only covering attribute levels, instead of complete stimuli at the bottom level, is advisable because it allows for the estimation of part-worths directly, thereby making AHP suitable to solving multiattribute design problem, and is thus chosen in this study. For the evaluation of hotel brand equity in this empirical study the decision problem was structured in a 3-level hierarchy and is modeling an incomplete hierarchy. The same attributes and levels as in case of CA were used following the proposed hierarchical structure as depicted in Figure 4.

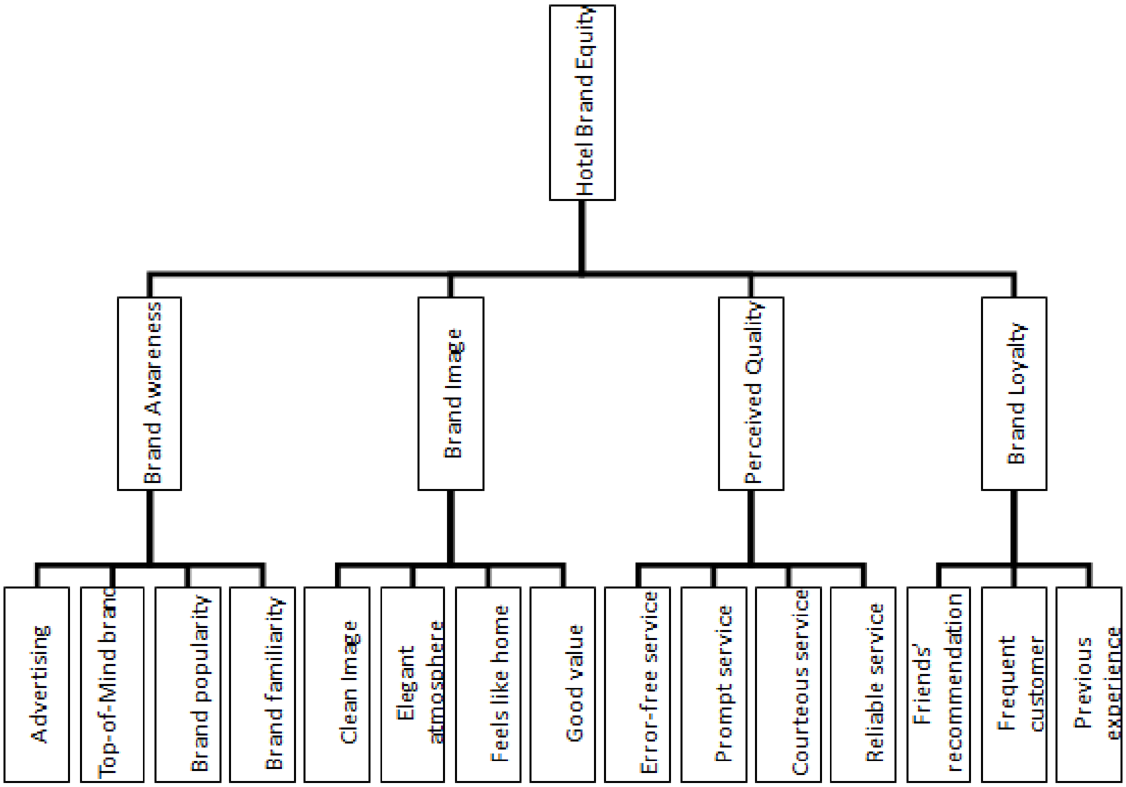


Figure 4. Hierarchical hotel brand equity model.

To perform a fair comparison to CA, the resulting hierarchy is evaluated in a bottom up manner. This means that respondents judge all pairs of attribute levels on the bottom level and then proceed with pair-wise comparisons of attributes on the next higher level of the hierarchy. In this way, the respondents are first introduced to the attributes' range and levels.

After the hierarchical model of the problem is set, in order to implement the AHP, respondents were asked to make two types of pair-wise comparisons: (a) a pair-wise comparison of the levels within each attribute; and (b) a pair-wise comparison of the attributes. In Saaty's AHP, pair-wise comparisons are made on a scale ranging from -9 to 9, according to the original description of AHP. The scale also used verbal statements ranging from "the criteria are equally important" (value of 1) to "extremely more important" (value either -9 or 9). Although several scales have been proposed for this process (Ji and Jiang, 2003; Finan and Hurley, 1999; Salo and Hamalainen, 1997), in this study an adaptation of the Saaty fundamental scale was used as it is easier to understand for respondents who are not skilled in complex mathematics or with the AHP method. However, when applying Saaty's scale to preference measurement a five-point scale instead of a nine-point scale was used in the present study, since the pilot test involving approximately 34 student respondents unskilled in the use of AHP has shown that respondents often did not understand how to use the scale and reported confused when using a more complex scale. In the similar vein, other recent studies (Pecchia et al., 2009a; Pecchia et al., 2009b; Pecchia et al., 2011c) have utilized a reduced scale and found that the results achieved with a five-point scale are equivalent to those achieved using the nine-point fundamental scale.

The data of pairwise comparisons were obtained using a reduced version of Saaty's (1980) scale ranging from "the criteria are equally important" (value of 1) to "extremely more important" (value either -5 or 5) and analyzed by applying the well-known eigenvalue approach, which is at the heart of AHP (Saaty, 2003), to compute the relative importance of attributes and levels.

A Comparison of AHP and CA

Along with the research questions AHP and CA were compared using measures for internal validity, convergent validity, and predictive validity. However, according to

Helm et al. (2004a) validity measures are not sufficient for a comparison concerning the feasibility of methods in marketing practice, since valid results can only be achieved if the respondents are able and willing to accept the assumptions of the methods. Thus to compare the practical feasibility of AHP and CA, the present study collected the respondents' subjective evaluations concerning interest, complexity and clarity, and realism of the task. Using the nine-point semantic scale respondents in this study were asked how enjoyable (interesting), how difficult and confusing (complexity) and how realistic the task was to evaluate. The respective evaluation questions to assess respondents' experience with the preference measurement task are shown in Table 10. A negative evaluation of the task is assumed to be an indicator for the possible use of simplifying strategies.

Table 10

Measures for Judging the Practical Feasibility of the Methods

Feasibility measure	Items (scale 1-9)
Enjoyment	How much did you enjoy the survey as a whole? low enjoyment (1) vs. high enjoyment (9)
Difficulty/clarity	How difficult and confusing was it to respond to the questions asked? difficult evaluation tasks (1) vs. easy evaluation tasks (9)
Realism	How realistic is this form of questioning? low realism (1) vs. high realism (9)

Source: Meißner and Decker, (2009).

For internal validity, both AHP and CA provide goodness-of-fit measures to assess the degree of inconsistency in the responses but these measures cannot be compared directly. The goodness-of-fit of the CA is measured by the coefficient of the determination. There is no similar criterion for the AHP but the inconsistency ratio calculated from the maximum eigenvalue can be used for the purpose of the analysis. To come up with a fair comparison, by not eliminating any respondents from the sample on the basis of the coefficient of determination (in the case of CA) or the consistency measure (in the case of AHP), the present study avoided favoring either method on the basis of goodness-of fit measures that are not directly comparable. Instead, an explorative comparison of AHP and CA with respect to the robustness towards different consistency levels was conducted.

To test for convergent validity Spearman’s coefficient of rank correlation comparing the rankings produced by AHP and CA was employed.

To compare the models' predictive validity, ten holdout choice tasks were included to measure the predictive accuracy of the two models. Each choice task consisted of four alternatives which were described on all attributes included in the study. To make the holdouts more challenging, the holdout choice set was designed to be Pareto-optimal in which none of the alternatives dominates the other alternatives (cf. Green and Srinivasan, 1990; Elrod, Louviere, and Davey, 1992). The predictive accuracy of both methods was checked by comparing the predicted overall utilities of the holdout stimuli with the actual choice in the presented holdout task. Three measures of predictive accuracy were used for comparisons: (1) Hit rate: the proportion of hits, where a hit is a choice correctly predicted by the model, (2) the mean absolute error (MAE) of predictions, (3) the root mean squared error (RMSE) measure. There measures have been frequently used in the literature for similar type of studies (Helm et al., 2003, 2004a, 2004b; Meißner, Scholz, and Decker, 2008; Meißner and Decker, 2009; Klein et al., 2010; Scholz et al., 2010).

The following section describes the potential confounding variables that will be controlled for in this study and the methods of measurement for each construct.

Control Variable

According to Helm et al. (2004a), the validity of results could be affected differently by the intensity of information usage. Several characteristics of the respondents such as respondents' product involvement, prior knowledge, and level of education may influence the extent of processing information (Vriens, 1995). For example, most of the earlier studies have used students as respondents. If reliability increases with the educational level of respondents, as is indicated in Tashchian, Tashchian, and Slam (1982), the use of such student samples may have biased the results. In addition, researchers of consumer behavior have postulated that a consumer's knowledge of products, past experiences and involvement levels affects the choice process (Hansen, 1972; Howard, 1977; Howard and Sheth, 1969). Some studies such as Bettman and Park (1980), Newell and Simon (1972) and Rothschild (1975) have presented empirical findings supporting this postulation. In order to make sure some variables would not influence the validity and reliability of AHP and CA across offline and online survey modes, five factors were considered as control variables since the two data sets came from different populations.

Gender, Age, and Education Level

It is theoretically discussed that demographic characteristics such as gender, age, and education level may influence validity and reliability measures (Tscheulin and Blaimont, 1993; Sattler et al., 2001; Sattler and Nitschke, 2003; Helm et al., 2004a; Klein et al., 2010). Nonetheless, empirical studies on this issue show mixed results. Tscheulin et al. (1982) found a significant influence of demographic factors: age and education are significantly related to task completion, and education is significantly related to validity.

Furthermore, Tscheulin and Blaimont (1993) found a negative correlation between education level and predictive validity. Krapp and Sattler (2001) identified formal education level as an indicator for cognitive ability. On the other hand, other studies from Sattler, Hensel-Borner, and Kruger (2001) revealed that the validity of preference measurement is not influenced by factors such as age, educational background and other demographic characteristics of respondents.

Regardless of the mixed results from former studies, in order to compare the validity of two data sets, where one data set was the result of a paper-based survey and the other data was the result of an online survey, it is important to previously control for the possibility of potential demographic influences since the two samples did not consist of the same test persons. Otherwise, different sample characteristics might lead to biased results.

Hotel Visitation Knowledge and Involvement

Previous research concerning consumer behavior has emphasized the importance of the relationship between product involvement and product knowledge (Lin and Chen, 2006; Park and Moon, 2003; Leian and Widdows, 1999). Adoption theory suggests that the most important factors for the evaluation process are objective/ subjective product knowledge (Brucks, 1986) and involvement (Celsi and Olson, 1988). Both product involvement and knowledge are related concepts which affect behavior and cognition, but their effect on information processing is different. Knowledge represents the ability to process information, and it is argued that experts use this ability to narrow down the amount of information processed and focus on a limited set of attributes, whereas involvement reflects a motivation to process information which increases the propensity to make compensatory judgments (Denstadli et al., 2012).

Scholz (2010) stated that product knowledge may unknowingly influence the validity of conjoint analysis results. Tscheulin et al. (1982) found that the respondents' prior knowledge of the subject matter is significantly related to validity. In the same vein, Kempf and Smith (1998) suggest that consumers with higher levels of product knowledge are more diagnostic and better informed than those who have lower levels of product knowledge. The higher the level of product knowledge a consumer possesses, the less chance there is that he/she will generate evaluation bias (Bian and Moutinho, 2011). Given these findings, the higher knowledge of respondents implies better abilities to evaluate attribute based product concepts with regard to its utility due to higher cognitive resources (Moreau, Lehmann, and Markman, 2001). Since some respondents could have some prior knowledge about the product category whereas others might not, the validity of different preference measurement methods may be dependent on the different levels of knowledge of the respondents. High product knowledge is an important precondition for preference measurement (Helm et al., 2011). There are three distinct but related ways in which consumer knowledge is conceptualised and measured: objective knowledge, subjective knowledge, and experience (Flynn and Goldsmith, 1999). To determine whether the knowledge of respondents would influence the result of this study a four-item scale developed by Smith and Park (1992) was used to measure an individual’s self-assessed ratings of hotel knowledge. These four items were slightly modified to fit in the hotel context from the original format of Smith and Park’s scale and rated on 5-point, Likert-type scale (1 = strongly disagree, 5 = strongly agree). The choice of measuring self-assessed knowledge is supported by Meeds (2004) who found that self-assessed knowledge was a better predictor of respondents' cognitive responses and general attitudinal evaluation in comparison to other kinds of knowledge.

Whereas knowledge indicates a general ability for product evaluation, involvement shows the motivation to actually use cognitive resources for the evaluation process (Helm et al., 2011). The involvement may have an indirect impact on the validity of different preference measurement methods as it influences the level of cognitive control in a decision making (cf. Kroeber-Riel and Weinberg, 1996; Felser, 1997). Krapp and Sattler (2001) argued that both validity and evaluation are influenced by a respondent's involvement with a particular product class and his or her cognitive ability. According to Felser (1997) the effort put into a task depends on the involvement. Since the various preference measurement methods are prone to simplifying effects to varying degrees (Wright, 1975) their predictive validity might depend on the respondents' involvement. Strebinger et al. (2000) showed that a low (task) involvement may have a negative effect on the estimated preferences thus resulting in lower validity results. Furthermore, according to Bian and Moutinho (2011) consumers with a higher level of product involvement are more likely to be able to distinguish between attributes; so that they can match their subjective utilities to an (objective) attribute and express stable preferences. On the other hand, the differences between attributes might not be easily recognised, if the level of product involvement is low, due to consumers' lack of motivation, effort and even capability in relation to processing information. Since only the measurement of stable preferences can lead to reliable and valid results (Darmon and Rouziès, 1994), involvement is a crucial precondition to assess stable preferences. Hotel involvement was measured using three items derived from Zinkhan and Locander's (1988) involvement scale, each being measured on an eight-point semantic scale. These items were contextualized for hotels and consisted of the level of interest in the product (hotel) class relative to others, the frequency of hotel stays, and level of self-reported involvement with the class. The paper (offline) version of the survey is presented in Appendix B.

CHAPTER 4

RESULTS

In this chapter, specific results of the comparisons between two classes of models, AHP and CA, are presented. First, the results of summary statistics are presented. Second, since the responses to the AHP model and the CA model were obtained from different respondents, the two samples (offline group vs. online group) are compared demographically in order to identify differences between the two groups. Third, the validity and reliability of the constructs of interest are also assessed, followed by an analysis of the data, to answer the research questions. Finally, the chapter provides the results which answered the research questions.

Respondents to the Offline Version

A convenience sampling method was used to recruit and collect data from students in different tourism programs at Arizona State University. Data were collected using a paper and pencil survey. Lecturers who agreed to administer the survey distributed the questionnaire to students during the teaching session, and respondents completed the questionnaire on a voluntary basis. 190 questionnaires were received during several lectures of tourism management. Of those, 181 (95%) completed both the AHP and CA tasks. Data from nine students (5%) were excluded from the analyses because they did not complete one or both tasks. The final usable sample size for the offline group was 181.The demographic and travel characteristics of the student respondents are shown in Table 11.

Table 11

Demographic Characteristics of the Paper-based Sample (n=181)

Characteristics	Frequency (%)	Characteristics	Frequency (%)
Gender		Hotel type	
Male	56 (30.9)	Budget/Economy	26(14.4)
Female	125 (69.1)	Mid-priced	59 (32.6)
		Upscale	80 (44.2)
Age, mean years (SD)	22 (3.4)	Luxury	10 (5.5)
		No answer	6 (3.3)
Marital status		Frequency of hotel stays	
Single	171(94.5)	1-3 days	90 (49.7)
Married	6 (3.3)	4-7 days	57 (31.5)
Others	3 (1.7)	8-15 days	22 (12.2)
No answer	1 (.6)	16 or more days	10 (5.5)
		No answer	2 (1.1)
Enrollment status		Purpose of visit*	
Full-time student	171 (94.5)	Vacation	144 (80.4)
Part-time student	7 (3.9)	Business	19 (10.6)
Others	2 (1.1)	Visit friends, relatives	37 (20.7)
No answer	1 (.6)	Others	9 (5.0)
Class Standing			
Freshman	10 (5.6)		
Sophomore	37 (20.4)		
Junior	76 (42.0)		
Senior	56 (30.9)		
Graduate	2 (1.1)		

Note. *The multi-response analysis was applied (total percent of cases: 116.8%)

Of the total sample, more than two-thirds of respondents were female (69.1%) and 30.9% were male. The average age was 22 years old (SD=3.4) in the sample, and almost all were single (94.5%). With respect to class standing, the sample consisted of juniors (42.0%), seniors (30.9%) and sophomores (20.4%). Almost all of the student respondents (94.5%) indicated full-time enrollment status. Using the multi-response analysis, four-fifths (80.4%) of the respondents stated that the reason for their last hotel visit was vacation (68.9% of all responses). Most of them stayed in mid-price or upscale hotels, for between one and three days.

Respondents to the Online Version

Qualtrics, a professional field research firm administered the survey instrument via the Web. Data was collected using it's online consumer panel. A total of 504 respondents took part in the surveys and fully completed both the AHP and CA tasks. Thus, data from a total of 504 were used in the analyses. General sample characteristics and demographics are presented in Table 12.

Table 12

Demographic Characteristics of the Web-based Sample (n=504)

Characteristics	Frequency (%)	Characteristics	Frequency (%)
Gender		Income	
Male	247 (49.0)	$25,000 or less	39 (7.7)
Female	257 (51.0)	$25,001-$50,000	109 (21.6)
		$50,001-$75,000	106 (21.0)
Age, mean years (SD)	43(13.8)	$75,001-$100,000	116 (23.0)
		$100,001-$125,000	44 (8.7)
		$125,001-$150,000	40 (7.9)
Marital status		$150,001 or above	50 (9.9)
Single	166 (32.9)		
Married	309 (61.3)	Region*	
Others	29 (5.8)	Northeast	103 (20.4)
		Midwest	116 (23.0)
Education		South	144 (28.6)
Less than high school	5 (1.0)	West	141 (28.0)
High school	61 (12.1)		
Some College	92 (18.3)	Hotel type	
Associate's degree	60 (11.9)	Budget/Economy	64 (12.7)
Bachelor's degree	175 (34.7)	Mid-priced	290 (57.5)
Master's degree	82 (16.3)	Upscale	131 (26.0)
Doctoral degree	13 (2.6)	Luxury	19 (3.8)
Professional degree	16 (3.2)		
Employment status		Cost of nightly stay	
Employed full-time	289 (57.3)	$60 or less	65 (12.9)
Employed part-time	54 (10.7)	$61-$100	214 (42.5)
Unemployed	24 (4.8)	$101-$150	135 (26.8)
Homemaker	47 (9.3)	$151-$200	51 (10.1)

Retired	59 (11.7)	$201 or more	39 (7.7)
Unavailable for work	11 (2.2)		
Student	14 (2.8)	Frequency of hotel stays	
Others	6 (1.2)	1-3 trips	231 (45.8)
		4-7 trips	143 (28.4)
Ethnicity		8-15 trips	78 (15.5)
American Indian or Alaska Native	2 (.4)	16 or more trips	52 (10.3)
Asian/Asian American	37 (7.3)		
Black or African American	46 (9.1)	Purpose of visit	
Native Hawaiian or other Pacific islander	1 (.2)	Vacation	293 (58.1)
Hispanic/Latino	31 (6.2)	Business	97 (19.2)
White/Caucasian	378 (75.0)	Visit friends, relatives	87 (17.3)
Others	9 (1.8)	Others	27 (5.4)

Note. *The Northeast census region includes Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont.
The Midwest census region includes Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin.
The South census region includes Alabama, Arkansas, Delaware, the District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia.
The West census region includes Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

Gender is distributed nearly evenly with 49% male and 51% female respondents. The average age in the sample was 43 years old (SD=13.8). The majority of the respondents were self-identified as White (75.0%). About three fifths (61.3%) of the respondents were married, and 32.9% were single. In terms of education level, the respondents were highly educated: 34.7% had a bachelor's degree and 18.9% had a graduate degree. More than half of the respondents (57.3%) were employed, 11.7% were retired, and 4.8% were unemployed. An annual income of $50,001 or more was reported by 70.5%. With respect to the regions, of the total respondents, 28.6% lived in the South, 28.9% lived in the West, 23.0% lived in the Midwest, and 20.4% lived in the Northeast. More than half of the respondents (58.1%) indicated that the main purpose of their last hotel visit was vacation. A majority of the respondents stayed at mid-price hotels for between one and three days, and paid nightly room rates between $61 and 100.

Control Variable Analysis

In order to compare the models' predictive performance as well as the validity of two data sets, where one data set was the result of a offline survey and the other data set was the result of an online survey, it is important to compare the two samples in order to understand whether they differed in terms of some sample characteristics, since the respondents who completed both the AHP and CA tasks were not randomly assigned to either the paper-based or the online questionnaire. There were five potential control variables under consideration, as identified earlier in the review of literature (Hansen, 1972; Howard, 1977; Howard and Sheth, 1969; Bettman and Park, 1980; Newell and Simon, 1972; Tscheulin and Blaimont, 1993; Sattler et al., 2001; Sattler and Nitschke, 2003): gender, age, education, hotel knowledge, and hotel involvement. However, in this study the level of education was not considered as one of the covariates because a student sample would have limited the education range.

Gender and Age

Due to different sample characteristics, comparisons are needed to explore the nature and direction of biases associated with non-random allocation or non-experimental approach (Madigan et al., 2000) because of their potential influence on dependent variables such as validity and reliability measures. Thus, to assess accuracy and bias differences between the offline and online surveys, the two samples (or the two data sets) were compared with respect to the key demographic variables of gender and age prior to answering the research questions. The results of the analyses are presented in Table 13.

Table 13

Comparison of the Offline and Online Samples in terms of Gender and Age

Demographic variable	Statistical test	Test statistic	Degree of freedom	p-value	Effect size
Gender	χ^2 homogeneity	$\chi^2(1)=17.626$	df=1	p=.000	phi (φ) =.160
Age*	Welch t	Welch t= -32.757	Rounded df=637	p=.000	est. ω^2=.314

Note. *The t-test using the Welch correction for nonhomogeneity was used; thus, degrees of freedom are reduced due to unequal variances via Aspin-Welch corrections to standard t-tests.
*Age was non-normally distributed; therefore, it was log-transformed.

A chi-square homogeneity test indicated that the samples did differ with respect to gender ($\chi^2(1)=17.626$, $p<.01$, phi (φ) =.160) with a small effect size using Cohen's (1988) criteria of.1=small effect, .3=medium effect, .5=large effect. Obviously, the offline sample was, on average, younger than the online sample. This difference is statistically significant with a rather moderate effect size (Welch t(636.976)=-32.757, p<.000, est. ω^2=.314). Comparison of the key demographic characteristics of the offline and online samples indicated that two samples were not homogeneous and differences in the results might be influenced by systematic deviations in the samples.

Hotel Knowledge and Involvement

Control variables (also referred to as covariates) are assumed to be measured without error, which should be reliable, as well as valid (Polit and Beck, 2008). Some variables, such as gender and age in this case, can be measured directly and reasonably reliably, so that they do not cause any problems and can be used as control variables; others (such as hotel knowledge and involvement in this case) that rely on a subjective scale may not meet this assumption. Therefore, prior to comparing the means of the two sampled groups with respect to their hotel knowledge and involvement, modified scales should be examined psychometrically and proven to be valid and reliable, not assumed to have the same dimensionality, reliability, and validity of original scales (Furr, 2011).

Confirmatory factor analysis. In order to assess the validity and reliability of the scales and items, a confirmatory factor analysis (CFA) was conducted using AMOS 19.0 software. Multivariate normality is an important assumption of confirmatory factor analysis. Multivariate normality was assessed separately for each sample using Mardia (1970)'s coefficient of multivariate kurtosis and using criteria critical ratio of 2.58±.01 significance level (1%). Mardia's coefficient for the offline sample was 2.011, with a critical ratio of 1.205, which is well below the critical value of ±2.58, indicating no multivariate normality violations, whereas multivariate kurtosis for the online sample did not show a normal distribution. Although it is hard to test multivariate normality directly, the achievement of univariate normality among variables is recommended (Hair et al., 1992; Meyers et al., 2006). Therefore, establishing univariate normality among a collection of variables can help gain multivariate normality (e.g., Bollen, 1989; Gold, Malhotra, and Segars, 2001; Meyers et al., 2006). For the online sample, the normality assumption was not violated with an acceptable range of skewness and kurtosis statistics with the value ranging from -.044 to .862; all were within the ±1 range (Hair et al., 2010; Tabachnick and Fidell, 2007). Hence, it was safe to assume that multivariate normality appeared to generally exist because all measurement items for the online sample showed a relative small level of skewness and kurtosis (Meyers et al., 2006). Maximum likelihood estimation was employed in the CFA since it is the most widely used estimation that demonstrates robustness against moderate violation of multivariate normality, if it indeed existed (Hair, Joseph, Anderson, Tatham, and Black, 1995).

Maximum likelihood estimation is considered favorable to other estimation methods when sample size is medium to large (Tabachnick and Fidell, 2007).

Prior to that, there was a maximum of 0.3% missing values (e.g. below the 5% threshold where listwise deletion is not recommended) on single variables in the offline sample; there were missing completely at random (Little's MCAR test, $p=.809$). Therefore, they were replaced with the expectation maximization (EM) algorithm (Enders, 2010).

Goodness-of-fit indices. The fit indices (see Table 14) indicate that the factor models fit the data reasonably for both samples on the basis of the fit criteria established in the literature.

Construct reliability and validity. Convergent validity was examined by the factor loadings, AVE and CR (Hair et al., 2006). The standardized factor loadings (λ) for all items in each latent construct were significant ($t > 1.96$, $p < .01$), providing evidence for convergent validity (Anderson and Gerbing, 1988). As shown in Table 14, all average variance extracted (AVE) values reached or exceeded the .5 cutoff value proposed by Fornell and Larcker (1981) but a benchmark value of 0.4 is also acceptable (Diamantopoulos and Siguaw, 2000), further supporting the convergent validity for each construct.

All composite reliability (CR) coefficients reached or exceeded the cutoff point of .6 suggested by Nunnally and Bernstein (1994), thereby indicating good construct reliability and adequate convergent validity (Anderson and Gerbing, 1988). The Cronbach's α reliability coefficients are all above .70, the recommended minimum threshold (Nunnally and Bernstein, 1994) for both samples, thereby indicating high internal reliability and consistency within latent constructs.

Table 14

Results of Confirmatory Factor Analysis

Construct	No. of Item	Offline sample (n=181)				Online sample (n=504)			
		λ*	AVE	CR	α	λ*	AVE	CR	α
Hotel involvement	3-item scale	.71 to.86	.64	.59	.83	.81 to .83	.70	.68	.78
Hotel knowledge	4-item scale	.55 to .81	.48	.79	.87	.55 to .87	.57	.86	.83
Cut-offs[a]		≥.50	≥.50	≥.60	≥.70	≥.50	≥.50	≥.60	≥.70

Note. *all λ significant at $p<0.01$, all t values>1.96

[a]Sources: Based on Anderson and Gerbing, (1988), Fornell and Larcker, (1981), Nunnally and Bernstein, (1994)

Goodness-of-fit indices

	Absolute						Relative					Parsimonious	
	χ^2/ *df*	P[b]	GFI	AGFI	RMSEA	SRMR	CFI	NFI	IFI	RFI	TLI	PNFI	PCFI
Offline sample	1.13	.327	.978	.952	.027	.042	.997	.971	.997	.953	.994	.602	.608
Online sample	3.97	.000	.971	.938	.077	.038	.979	.973	.979	.956	.967	.602	.606
Cut-offs [c]	<5	>.05	>.90	>.80	<.08	<.08	>.95	>.90	>.90	>.90	>.95	>.50	>.50

Note. χ2=Chi-Square; df=Degree of Freedom; GFI=Goodness of Fit; SRMR=Standardized Root Mean Square Residual; RMSEA=Root Mean Square Error of Approximation; AGFI=Adjusted Goodness of Fit; TLI=Turker_Lewis Index; NNFI=Non Normed Fit Index; NFI=Normed Fit Index; IFI=Increment Fit Index; CFI=Comparative Fit Index; RFI=Relative Fit Index; PNFI=Parsimony Normed Fit Index; PGFI=Parsimonious Goodness of Fit.

[b] the χ^2 test is prone to show a significant lack of model fit in studies with large sample size, so its results cannot be assessed in isolation. Also, the higher the comparative fit index (CFI) and the lower the root mean square residual (SRMR), the better the fit.

[c]Sources: Based on Bagozzi and Yi, (1988), Baumgartner and Homburg, (1996), Cote et al., (2001), Diamantopoulos and Siguaw, (2000), MacCallum et al., (1996), Ping (2004), Marsh and Hocevar, (1985).

The discriminant validity of the measures was assessed in three ways. First, a 95% confidence interval was constructed for the correlation of each pair of latent variables. None of the confidence intervals included 1.0 or -1.0 for both samples, providing support for discriminant validity (Anderson and Gerbing, 1988). Second, Fornell and Larcker's (1981) criterion requires the average variance extracted (AVE) for a pair of constructs to be greater than the squared factor correlation between the two constructs ($\varphi 2$). All AVE estimates exceeded the squared interconstruct correlations (see Table 15). Third, chi-square difference tests were used by comparing a restricted model in which a factor correlation was fixed at one with the original unrestricted model (Anderson and Gerbing, 1988). All the restricted models had substantially poor model fits, resulting in significant chi-square difference statistics. Thus, the discrimination between the constructs was determined to be sufficient although moderately correlated.

Table 15

Results of Discriminant Validity Tests

		Unconstrained Model		Constrained Model		Change
	Interconstruct Correlation ($\varphi 2$)	χ^2	df	χ^2	df	$\Delta\chi^2$/ df
Offline	.68 (.46)	20.650	13	37.440	14	16.79/1*
Online	.71 (.50)	51.660	13	86.923	14	35.263/1*

Note. *Significant at $p<.001$

Tests of invariance between both groups. According to Milfont and Fischer (2010), measurement invariance needs to be tested for cross-group comparisons (especially for mean comparisons). In order to assess measurement invariance, multi-group confirmatory factor analyses (MGCFA) are performed. These involve various hierarchical model testing steps with the two samples simultaneously (Byrne, 2008). One must first test for configural invariance, metric invariance and factor covariance invariance (Steenkamp and Baumgartner, 1998). Testing of error variance invariance is not obligatory, however, tests for error variance are also provided. Several hierarchical models were constructed to test for four types of invariance. The results are presented in Table 16.

Table 16

Results of Invariance Test for Both Groups (Offline and Online)

Model	χ^2	*df*	RMSEA	TLI	CFI	$\Delta\chi^2$	Δdf	P	ΔCFI
Configural invariance	66.335	26	.048	.972	.983	-	-	-	-
Metric invariance	70.176	31	.043	.978	.983	3.841	5	.573	.000
Covariance invariance	70.527	34	.040	.981	.985	.352	3	.950	.002
Full error variance invariance	132.414	41	.057	.960	.961	61.887	7	.000*	.024

Note. *Significant at p<.001

Configural invariance was tested by allowing all the loadings, covariance and error variances to be free across both groups. In the next step, metric invariance was assessed by constraining the matrix of factor loadings to be invariant between both groups. As can be observed in Table 16, the chi-square difference between the constrained and unconstrained models is insignificant, meaning the loadings of indicators are in fact, invariant between both groups. The next step is testing for factor covariance invariance for both groups. The results show that the model is invariant for both groups, since no significant change in the chi-square statistics can be observed. Finally, tests for error variance invariance (added error variance constraints) were conducted and revealed that invariance was not supported for the error variances ($\Delta\chi^2$ (7) = 61.887, p<.001).

Although the Chi-square difference test should be conducted to compare the fit, it suffers from the same well-known problems as the chi-square test for evaluating overall model fit (e.g., very sensitive to sample size and unbalanced group sizes). According to Cheung and Rensvold (2002), ΔCFI is unbiased by model complexity or sample size. A recent study by Meade et al. (2008) proposes ΔCFI values of ≤ 0.02 as an acceptable criterion for judging model invariance. Given that the CFI difference (ΔCFI = 0.02) remains within Meade et al.'s guideline of 0.02, it can be concluded that the measurement model is mainly invariant between both groups, meaning that the understanding of the two concepts (hotel knowledge and hotel involvement) between offline and online groups is basically the same. Thus, the two sample data can be pooled together, so respondents can be compared on their scores on the latent variable across groups. The average scores of the items related to hotel involvement and hotel knowledge were calculated for the use of further analysis.

The means according to sample groups are summarized in Table 17. An independent t-test was used to compare the means of the two sampled groups in terms of their hotel knowledge and involvement, respectively. The results of the analyses are presented in Table 17. The results indicated that the two samples were significantly different in terms of their hotel knowledge, with the online sample possessing significantly more hotel knowledge with a small effect size ($t(683)=-2.367$, $p<.05$, partial $\eta^2= .008$). On the other hand, no significant hotel involvement level difference was found between the two samples ($t(683)=.325$, $p=.745$, partial $\eta^2= .000$).

Table 17

Mean Comparisons of Knowledge and Involvement

Construct	Offline group	Online group	t-test	p	η_p^2
Subjective knowledge	3.5290	3.6820	t= -2.367	.018*	.008
Involvement	4.9890	4.9425	t=.325	.745	.000

Note. *Significant at p<.001

In the subsequent analyses, it is important to consider the differences observed in Tables 13 and 17. For this reason demographics (gender and age), as well as hotel knowledge and hotel involvement were used as control variables (covariates) for further analyses, because they may be significantly related to a specific dependent variable (such as validity and reliability measures). The purpose of this control is to control for possible confounding of online/offline comparisons and thus reduce the influence of extraneous variables so that changes in the dependent variable can be attributed to the independent variable. In the following sections, the results are presented in sequence and relative to each research question.

Practical Feasibility

RQ1: Are there any differences in the respondents' subjective evaluations of the methods in terms of (a) enjoyment, (b) difficulty and clarity, and (c) realism?

Before examining several validity measures, it is important to compare the feasibility of the methods (AHP vs. CA) from respondents' perspective, because valid measurements are only possible if respondents are able and willing to apply the method in a motivated manner (Scholl et al., 2005). For example, difficult or cognitively demanding questions may lead to less reliable responses and to a worse validity of the preferences measured (Helm, et al., 2004a).

Due to the fact that CA requires holistic judgments on several attribute levels, it can be assumed that the cognitive burden should be higher for CA than for AHP. In the latter only two attributes or attribute levels have to be evaluated at a time. AHP questionnaires might be much easier to answer and it can be assumed that respondents evaluate CA to be more difficult. Because of the holistic approach of considering all the relevant attributes simultaneously, CA should be evaluated as being more realistic compared to AHP as often suggested in the literature.

Post-Survey Feedback

To compare the feasibility of AHP and CA, respondents were asked for feedback about their experience with the preference measurement task. Specifically, they were asked to rate on a nine-point scale the difficulty and clarity of the task, the degree of enjoyment derived from completing the task, and the extent to which their responses were realistic. This is only a perceived assessment of the ability of the method to capture preferences because respondents did not see the final estimated preferences. The average respondent ratings of perceptions according to data collection modes are summarized in Table 18. To provide an overall measure of the respondents' subjective evaluations, the data were pooled across the two data collection modes, namely, a mixture of the two modes.

Table 18

Mean Comparisons of Feasibility Measures

	Data collection mode											
	Offline				Online				Mixed			
Feasibility Measure	AHP	CA	t value (p*)	η_p^2	AHP	CA	t value (p*)	η_p^2	AHP	CA	t value (p*)	η_p^2
Enjoyment	3.91	3.68	1.974 (.000)	.021	6.43	6.00	6.375 (.000)	.075	5.77	5.39	6.453 (.000)	.057
Difficulty	5.53	4.46	5.844 (.000)	.159	6.53	5.61	9.767 (.000)	.159	6.27	5.31	11.359 (.000)	.159
Realism	5.52	5.12	2.728 (.007)	.040	6.28	5.81	6.630 (.000)	.080	6.08	5.86	6.947 (.000)	.066

Note. Responses to three questions were obtained on a semantic nine point scale “How much did you enjoy this survey?” (1=boring, 9=interesting), “How difficult was it to respond to the questions asked?” (1=difficult, 9=easy) and “How realistic is this form of questioning in choosing a hotel?” (1=unrealistic, 9=realistic)
*Significant at $p<.001$

According to Table 18 all feasibility measures show a higher mean value for AHP for both data collection modes and their mixture, suggesting it is more interesting, easier and more realistic. Using paired t-tests, the results indicated that highly significant differences ($p<.01$) between AHP and CA were observed for all three measures. In terms of the perceived task difficulty, clarity, and enjoyment, as expected, AHP has a clear advantage. This advantage is certainly due to the lower complexity level of its evaluation tasks. Contrary to expectations, the higher degree of realism for CA by presenting whole stimuli assumed in the literature did not result in significantly higher ratings for CA. One possible explanation could be that the ranking tasks are too complex and cognitively challenging which may result in increased respondent fatigue and boredom, and decreased task motivation. Such situations are sometimes referred to as task overload (Cattin and Weinberger, 1980). It is also possible that respondents really tend to evaluate pairs of attributes rather than complete alternatives when choosing a hotel.

Survey Duration

A further important measure for the feasibility of a method is the time required by the respondents to fill in the questionnaires for each method. The time to finish a questionnaire is a major concern in consumer research because respondents' willingness to participate in a study diminishes with the length of the survey, and the costs of a survey increase with survey length (Meißner and Decker, 2009). Furthermore, longer questionnaires may fatigue the respondents, resulting in unreliable evaluations and undesired cancellations of the survey (Meißner and Decker, 2009). From this perspective, the time required by the respondents to complete each method was recorded. The mean completion times according to data collection modes are summarized in Table19.

In the offline mode, respondents took 4.63 minutes on average to fill out the AHP questionnaire, whereas the CA survey required 5.52 minutes. While the average survey length of AHP was 2.83 minutes in the online mode, the CA survey took 5.39 minutes on average. When pooled, the AHP survey took an average 3.03 minutes, compared with 5.40 minutes for the CA. Using paired t-tests, the results (see Table 19) revealed that AHP puts significantly less burden on the respondent with respect to survey length: offline mode, $t(45)= -3.235$, $p<.001$, $\eta_p^2=.189$; online mode, $t(378)= -14.550$, $p<.001$, $\eta_p^2=.359$; mixed mode, $t(424)= -14.681$, $p<.001$, $\eta_p^2=.337$. The favorable results of AHP suggest that even product evaluation problems with higher numbers of attributes should be possible.

Table 19

Mean Comparisons of Completion Time

	Data collection mode											
	Offline[a]				Online[b]				Mixed			
	AHP	CA	t value (p)	η_p^2	AHP	CA	t value (p)	η_p^2	AHP	CA	t value (p)	η_p^2
Time	4.63	5.52	-3.235 (002)*	.189	2.83	5.39	-14.550 (000)*	.359	3.03	5.40	-14.681 (000)*	.337

Note. [a]Students (n=46) were randomly selected from a class to measure the time needed to accomplish each task
[b]Due to technical difficulties the completion times of 123 respondents (24.4%) in the online sample were not recorded and therefore unavailable, resulting in n=381.
[b]To screen univariate outliers, the well-known boxplot outlier labeling rule was used, which is based on multiplying the Interquartile Range (IQR) by a factor value of 2.2 (Hoaglin, Iglewicz, and Tukey, 1986; Tukey, 1997) and thus were set to missing (n=2), resulting in n=379.
[b]Due to significant skew and kurtosis, logarithmic transformations were used to achieve normality; however, for ease of interpretation, untransformed means are presented in the table.
*Significant at p<.001

Summarizing the above results, significant preference was found in the feasibility of the AHP as a practical method. For all measures of the methods' feasibility, AHP is at least preferred by the respondents. Respondents found the AHP survey to be a better overall experience, to be more realistic, less boring, and more engaging. Furthermore, the AHP questionnaire needed significantly less time than CA questioning. Thus, it can be concluded that the AHP has clear advantages in terms of ease, speed and costs.

The main goal of preference elicitation methods is to get a valid and reliable model of the respondent's preference structure. The following sections describe several measures primarily applied in this study and compare the validity of the results obtained by AHP and CA, respectively, along with the research questions. The presentation follows the classification of validity measures in content, convergent, internal, and predictive validity.

Content Validity

Among other things the content validity measures the plausibility of a study (Klein et al., 2010). Content validity measures are generally used to test whether or not plausible a-priori assumptions are fulfilled (Scholl, et al., 2005). A detailed evaluation of the content validity is to compare the relationships between the estimated part-worths. However, in the present study, there are no general and plausible a-priori assumptions on the preference order for attribute levels (e.g., monotone). Due to the non-unique a prior ordering for levels within attributes, all attributes were modeled as discrete such as categorical variables with no a priori assumptions about which would be preferred. Thus, tests for content validity were unnecessary.

Convergent Validity

RQ2: Do AHP hotel branding results accord generally with CA (convergence)? If so, to what extent does AHP have convergent validity with CA with respect to (a) importance ratings, (b) part-worth estimations, and (c) estimated overall utilities?

Within a test comparing different preference elicitation methods, the convergent validity is a further aspect of judging a method's validity. Convergent validity relates to the amount of agreement among maximally different methods of measuring the same construct (e.g., observed preference, estimated part-worth utility, and estimated overall utility) (Leigh et al., 1981). In the present study, if two different methods (AHP vs. CA) reflect the same construct (e.g., are positively correlated), their convergent validity should be high (Scholl et al., 2005).

Attribute Importance Convergence

The most common way of examining the convergent validity is to correlate the importance of attributes measured by two methods (Stillwell et al., 1983; Van Ittersum et al., 2007). In the following, the Spearman's coefficient was adopted for correlating ranked data in order to contrast the convergent validity of AHP with CA. The absolute correlation coefficient can be used to examine the convergent validity. Correlations below .35 are generally considered low, while those above .45 are considered moderate to high (Van Ittersum et al., 2007). This study used a correlation of .35 as the cut-off level for concluding whether or not two methods show convergent validity.

In order to make the CA importance weights comparable to the ones of AHP, the differences in part-worth utilities for the worst and best level of each attribute were normalized, such that they sum to one as the AHP weights do. Table 20 is a summary of the results. As the table indicates, the importance weights were noticeably different in ranks and sizes across the two models for both survey modes and the mixture.

Table 20

Comparison of Mean Attribute Importance Estimates

Attribute	Offline AHP	Offline CA	Online AHP	Online CA	Mixed AHP	Mixed CA
Brand Awareness	.183(4)	.324(1)	.167(4)	.258(2)	.183(4)	.275(2)
Brand Image	.213(3)	.298(2)	.193(3)	.295(1)	.195(3)	.296(1)
Perceived Quality	.376(1)	.198(3)	.376(1)	.241(3)	.365(1)	.229(3)
Brand Loyalty	.227(2)	.179(4)	.264(2)	.206(4)	.257(2)	.199(4)
Rank Order Correlation Coefficients	Aggregate	-.800[a]		-.600[a]		-.600[a]
	Individual	(-.120)[b]		(-.097)[b]		(-.103)[b]

Note. The ranking of attribute levels is depicted in parentheses
[a]Based on aggregate model (also known as pooled analysis)
[b]Based on disaggregate model
[b]Average Rs are computed via averaging the Fisher z-transformed r, since the raw Rs are not interval-scaled.

Results from the offline mode indicated that while perceived quality (.376) is the most important attribute for AHP, this attribute is only at the third rank for CA. Inversely; brand awareness (.324) is at the first place for CA, but only at the fourth for AHP. For the online mode AHP identifies perceived quality (.376) and brand loyalty (.264) as the top two attributes. Brand image (.193) and brand awareness (.167) are ranked as less important attributes. In the CA results, in contrast, brand image (.295) is the highest ranked attribute which is ranked third in the AHP, and the second most important attribute was brand awareness (.258), while it was the least important attribute in the AHP. For the mixture, while in the AHP results perceived quality (.365) and brand loyalty (.257) ranked as the most important two attributes, in the CA results the former attribute ranked third and the latter comes in the fourth place.

To examine the nature of the differences, the Spearman rank correlation coefficient was computed independently for each survey mode. The aggregate rank correlations between the attribute importance weights of AHP and CA varies between -.80 and -.60. All values are considerably lower than the threshold of .30 proposed by Van Ittersum et al. (2007), demonstrating that a lack of convergent validity exists among the methods for measuring the importance of attributes. In addition, individual rank correlation coefficients between the methods was computed for each respondent and then averaged after applying Fisher's Z-transformation to the correlations. The resulting coefficients (ranging from -.120 to -.097) on the individual level are all below .30, the recommended minimum threshold for all survey modes. Thus, there again is almost no convergent validity (e.g., the methods result in different preference structures).

Because the two methods use different procedures for collecting data and calculating part-worths, it is necessary to test whether the importance weights obtained from the methods are the same (convergent validity). If the importance weights are identical and not significantly different, equal predictive validity can be expected for the two methods (Hensel-Börner and Sattler, 2000), because all further computations are based on the estimated part-worth utilities. Otherwise, there might be differences in the preference structures identified by the methods.

As further analysis, Paired Hotelling's T-square test, a specialized form of multivariate analysis of variance (MANOVA), was applied to assess the equivalence of the relative attribute importance weights. Paired Hotelling's T-square test is an extension of paired t- test and is used to account for the effect of correlations when there is more than one outcome variable (Fukui et al., 2010). Paired Hotelling's T-square test revealed that the average importance weights significantly differed between the two methods in all cases: offline (T^2 = 175.062, F(4, 177)=43.268, p = 0.00, η^2=.494); offline (T^2 = 392.564, F(4, 500)=97.715, p = 0.00, η^2=.439); mixed (T^2 = 550.498, F(4, 681)=137.293, p = 0.00, η^2=.446).

Further exploration (see Table 21), by means of univariate F tests using Bonferroni adjusted alpha levels of .0125 per test, showed the deviations are highly significant for all attributes (all F values significant at p<.001). Effect sizes η_p^2 ranges between .1 and .4, which are considered large in size (Cohen, 1988). So it can be overall concluded that the two preference measurement methods yield different relative attribute importance weights. This confirms similar findings, which have shown considerable discrepancies between them, in particular with respect to the attribute weights (Helm et al., 2003, 2004a; Scholl et al., 2005; van Til, et al., 2008; Kallas et al., 2011; Ijzerman et al., 2012).

Table 21

Results of Paired Hotelling's T-square Tests

	Offline		Online		Mixed	
Attributes	F (p-value)	η_p^2	F (p-value)	η_p^2	F (p-value)	η_p^2
Brand Awareness	60.571 (0.00)	0.252	86.592 (0.00)	0.252	144.617 (0.00)	0.175
Brand Image	40.692 (0.00)	0.184	198.604 (0.00)	0.184	233.062 (0.00)	0.254
Perceived Quality	133.815 (0.00)	0.426	212.907 (0.00)	0.426	339.090 (0.00)	0.331
Brand Loyalty	17.077 (0.00)	0.087	66.995 (0.00)	0.087	83.982 (0.00)	0.109

Different methods of collecting data and estimating attribute importance weights have been found to show low levels of convergence (Jaccard, Brinberg, and Ackerman, 1986). In this regard, differences in the resulting attribute importance weights may be due to the different way of deriving the attribute weights. However, it is not possible to determine which method comes closest to the true preference structure without judging which method predicts better.

Spread of Weights

Another noticeable difference between AHP and CA was the range of the attribute weights from the first to the last ranked attribute. The difference between the first and the last ranked importance weights was calculated to represent the range (R) of the attribute importance weights for each method. In AHP these values (offline: R=.193; online: R=.209; mixed: R=.182) are noticeably lager than those in CA (offline: R=.145; online: R=.089; mixed: R=.097 in all three cases. This suggested that the range in attribute importance weights obtained from AHP is larger than with CA. The latter method produces a lower range in the aggregate importance weights for both modes. However, the aggregate perspective does not reflect much information about the distribution of individual importance weights, because substantial differences might be leveled out when pooling the data (Scholz, et al., 2010).

Following the approach employed by Scholz et al. (2010), Lorenz curves and Gini coefficients were used to compare the individual attribute importance distributions across methods. Although the range is a frequently used measure of variability (dispersion), it is very limited, because it is based solely on the two most extreme values in the distribution and does not fully reflect the pattern of variation within a distribution. In contrast, Lorenz curves and Gini coefficients provide the complete information about the distribution of the attribute importance weights (see, e.g., Scholz, et al., 2010).

To assess the inequality of individual attribute importance weights, the Gini coefficient was computed at the individual level for each method using the following equation and then averaged across individuals, as is the recommended approach in the literature.

$$G = \frac{n+1}{n} - \frac{2\sum_{i=1}^{n}(n+1-i)w_i}{n\sum_{i=1}^{n} w_i}$$

, where the $w_i \ldots\ldots w_n$ are the attribute importance weights indexed in ascending order ($w_i \leq w_{i+1}$)

Using paired t-tests (see Table 22), the average Gini coefficients of AHP were found to be significantly larger than those of CA in all three cases: offline G(AHP)=.288 vs. G(CA)=.268, p<.10 (marginally); online G(AHP)=.261 vs. G(CA)=.243, p<.05; mixed G(AHP)=.268 vs. G(CA)=.249, p<.05). This implies that AHP individual attribute importance weights tend to show more curvature than CA (see Figures 5-7). The findings of the current study are consistent with those of Ijzerman, van Til, and Bridges (2012) who found that AHP's individual attribute importance weights are more distinctive than those of CA. A possible explanation for this dissimilarity is that the varying anchor points of AHP allow respondents better able to distinguish between attributes and lead to greater spread (dispersion) of weights (Pöyhönen and Hämäläinen, 2001).

Table 22

Mean Comparisons of Gini Coefficient of Inequality

Survey mode	AHP	CA	t test (p-value)	η_p^2
Offline	.288	.268	1.886 (.061)*	.019
Online	.261	.243	2.479 (.013)**	.012
Mixed	.268	.249	3.078 (.002)***	.014

Note. The Gini coefficient of inequality is the most commonly used measure of inequality and is defined as twice the area between the 45-degree line and the Lorenz curve. The coefficient varies between 0 and 1. A Gini coefficient equal to 0 indicates that all attributes have the same importance (perfect equality), whereas value equal to 1 indicates complete inequality (Scholz, et al., 2010).

*Significant at p<.10; **Significant at p<.05; ***Significant at p<.01

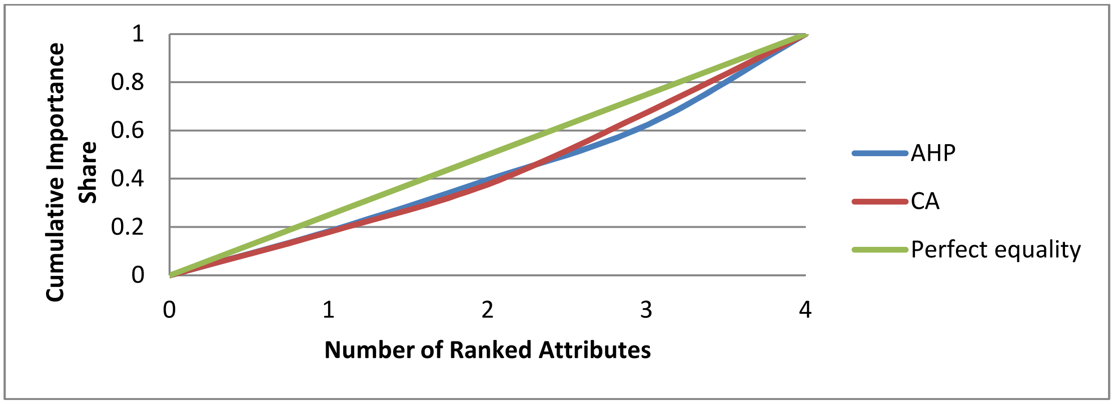


Figure 5. Aggregate Lorenz curves for offline study.

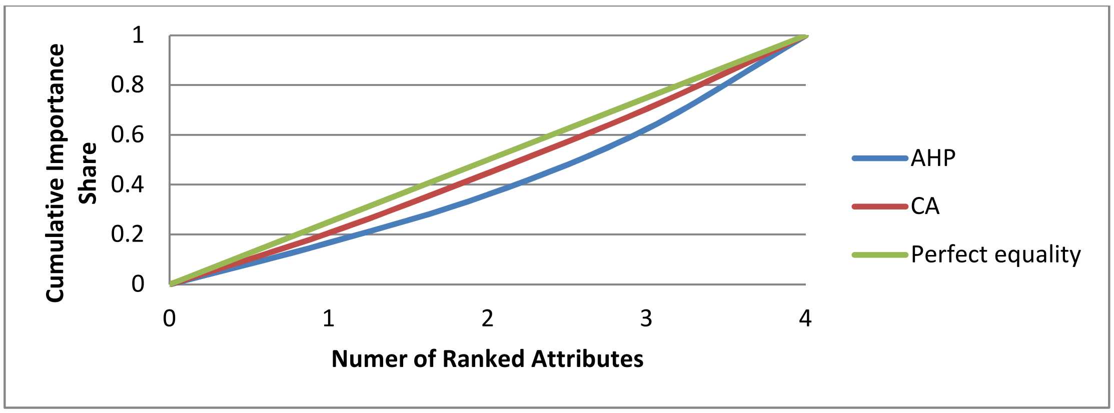


Figure 6. Aggregate Lorenz curves for online study.

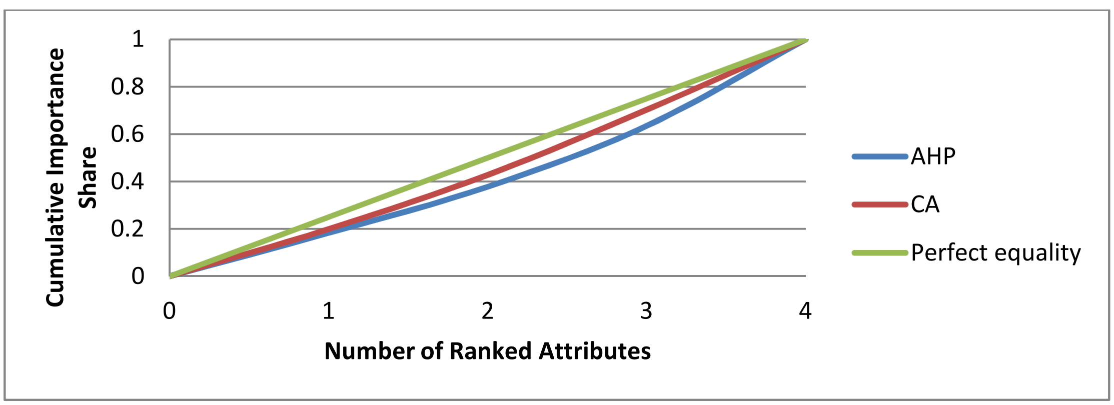


Figure 7. Aggregate Lorenz curves for mixture.

Attribute Level Convergence

A second method to examine convergent validity is to compare the rankings of part-worths derived by AHP and CA and thereby determine the rank correlation between part-worth utilities on the aggregate level (Meißner and Decker, 2009). To allow proper comparison of AHP and CA, the part-worth utilities of AHP are calculated by multiplying the preference weights of the attribute levels with the corresponding importance measures. Conjoint Analysis part-worths often are negative and sum to zero and, in order to be able to make the conjoint analysis part-worth utilities comparable to the ones of AHP it is necessary to normalize them so that they range between 0 and 1 as the AHP weights do. Figures 8, 9 and 10 depict the transformed part-worth utilities of both AHP and CA. Although both methods are conceptually different, the part-worth utilities show high structural similarity (equality) on the aggregate level. The rank correlations between AHP and CA part-worth utilities are very high in all three cases, ranging between .696 and .807; thus, all values exceeded the cutoff point of .3 suggested by van Ittersum et al. (2007). These results indicate that there is a strong degree of convergence between AHP and CA part-worth utilities.

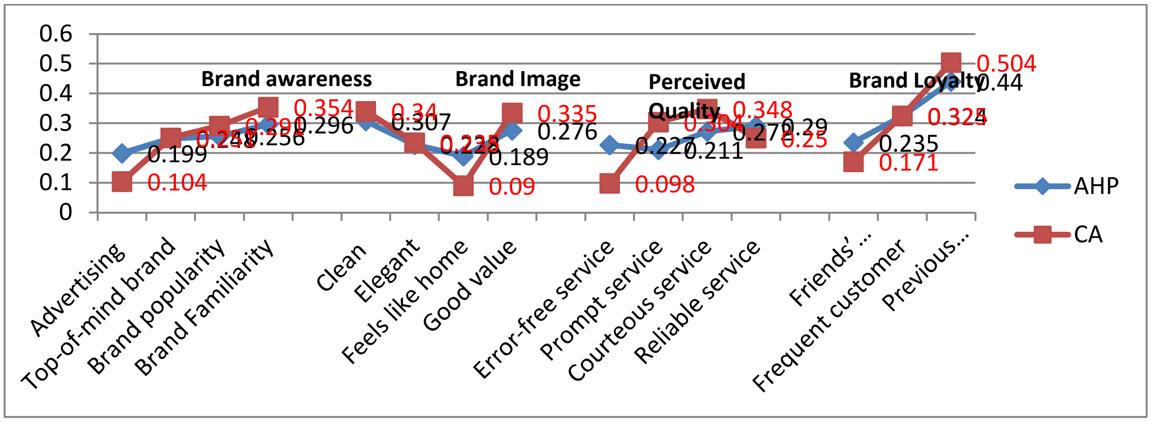


Figure 8. Aggregate part-worth utilities of offline study.

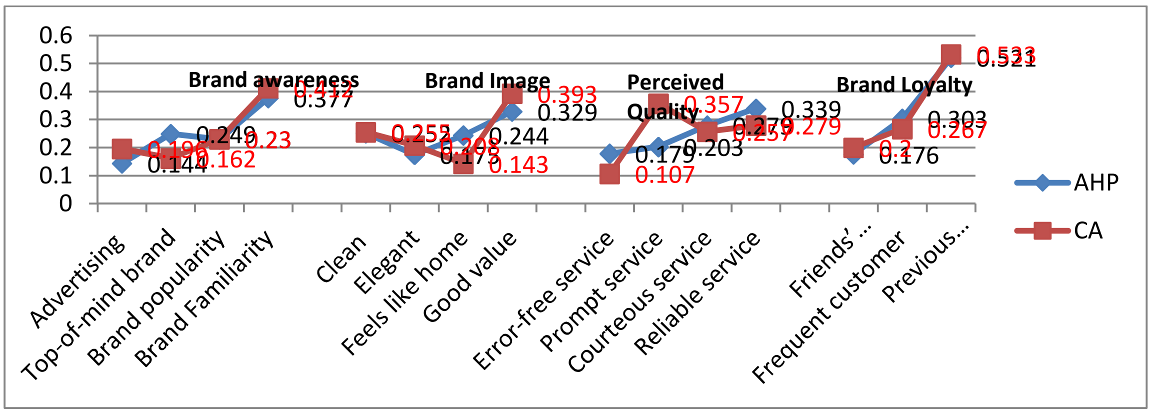


Figure 9. Aggregate part-worth utilities of online study.

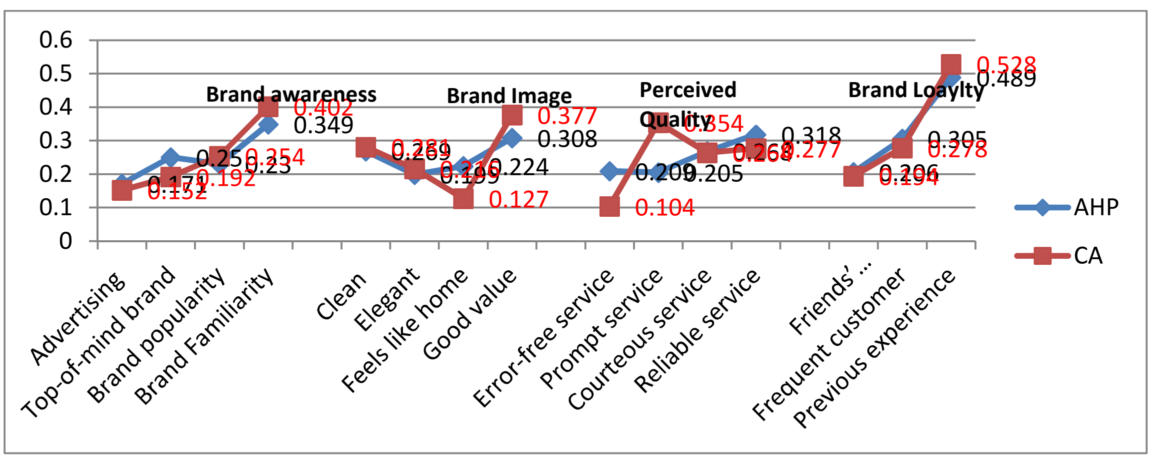


Figure 10. Aggregate part-worth utilities of mixture.

Holdout Convergence

A third method to analyze the convergent validity is to determine the rank correlation between the predicted values of the holdout sample (Mulye, 1998). In the following, the convergent validity was measured through Spearman's coefficient comparing the utilities of the holdouts computed by the part-worths of AHP and CA. The results are reported in Table 23. The average rank correlation coefficient (after Fisher-Z-transforming the correlations) varies between .478 and .643, so all values exceeded the proposed .3 cutoff value. Thus, the results show a high convergence between the methods. The present findings seem to be consistent with other research which showed good holdout convergence between the two methods (Mulye, 1998).

Table 23

Holdout Convergence

Data collection mode	Offline	Online	Mixed
Rank Order Correlation	0.643*	0.478*	0.527*

Note. The individual correlation coefficients were Fisher z-transformed prior to calculating the mean correlation coefficient.
*Significant at $p<.001$

In summary, the convergent validity gave mixed results. Both models are equivalent with regard to convergent validity in both the relative part-worths as well as the overall (holdout) utilities. More specifically, AHP did exhibit a high degree of convergent validity with respect to part-worth estimates and estimated overall utilities when compared with CA. However, there are substantial differences between the attribute importance weights of AHP and CA. Since the importance weights do not have the same structure in these models, it may be concluded that there is almost no convergent validity. With respect to the differences regarding attributes, it cannot be said which method provides the better results. Also, it is necessary to investigate the predictive validity of the two models on the individual level, as well as on the aggregate level and it will be investigated in the further sections.

Internal Validity

RQ3: Do both the offline and online data collection modes lead to comparable results regarding the internal validity of AHP?

The internal validity examines if the measurement of the data is consistent and to what extent the empirical input data matches the estimated data (Klein et al., 2010; Backhaus et al., 2003). Thus, the quality of responses can be assessed by evaluating the internal validity (consistency) of the preference evaluation tasks. To evaluate the degree of consistency for the entire hierarchy, average consistency ratio (ACR) was used for AHP. In case of CA the coefficient of determination R^2, measuring the goodness-of-fit of the preference model, was considered, as will be described in the next section. Due to the different approaches of the two models, theses values cannot be compared in a direct way, but it is possible to analyze the robustness of AHP and CA considering different levels of consistency.

To evaluate the internal validity of AHP, AHP uses a consistency ratio (CR) to test for a good model (Saaty, 1980). A consistency ratio (CR) is computed for each comparison matrix on each level of the hierarchy. This measures the internal consistency of the judgments entered into the matrix according to Saaty (1980). If the consistency ratio is 0.10 or less, the judgments are reasonably consistent and acceptable. A CR of 0.0 means that the judgments are perfectly consistent. Saaty suggests that if that ratio exceeds

0.1 the set of judgments may be too inconsistent to be reliable. In practice, CRs of more than 0.1 sometimes have to be accepted. Some authors have proved that it is possible to increase this threshold to 0.2 when the hierarchy is complex and it is not practical for the respondents to discuss the questionnaire results. Karpetrovie and Rosenbloom (1999) found that it is possible to answer rationally and consistently and obtain a consistency ratio above 0.1. The important issue is that respondents understand what they are doing, so as not to discard rational responses, which could lead to a loss of valuable information (Bodin and Gass, 2003). To this end, some studies have used consistency ratio of 0.2 as cut off (Cook, Angus, Gottberg, Smith, and Lonhurst, 2007). For this purpose, a threshold of $CR \leq 0.2$ is generally considered to be a minimum requirement to get at least a low level of consistency. A sufficiently high level of consistency may be obtained with the condition $CR \leq 0.1$. If $CR > 0.2$, the pair-wise judgments are just about random, and are completely untrustworthy and inconsistent. In order to evaluate the degree of consistency for the entire hierarchy, the arithmetic mean of all consistency ratios (ACR) can be simply used for AHP (Saaty, 1980; Scholl et al., 2005).

Saaty (1980) also proposes another procedure to test the overall consistency of the hierarchy by defining a Consistency Ratio of the Hierarchy (CRH). CRH is computed as the weighted mean of the consistency ratios (CR) of all comparison matrices. Again, overall consistency is acceptable if CRH is less than 0.10 (up to 0.20 is tolerable) (Saaty, 1980).

Table 24 shows the average values of CR for the total sample according to survey modes. The average CR (ACR) ranges from .144 to .164, so all values lie below the threshold value of 0.20 proposed by Cook, Angus, Gottberg, Smith, and Lonhurst (2007), thereby indicating an acceptable internal consistency. Furthermore, respondents were grouped by the degree of consistency of judgment on the basis of CR. According to Table 24 the absolute and the relative frequencies (%) are shown separately for respondents with unacceptable, low/moderate, and high consistencies.

For the offline mode, about 38.1% of the respondents are considered to be highly consistent in their judgments. A further class of about 40% of the respondents is judged to have low or moderate consistency. Approximately one-fourth of respondents are classified as completely inconsistent. For the online mode, about 42.5% of respondent are judged to have a high level of consistency. More than one-third of respondents are considered to be low or moderately consistent in their judgments. Only 20.4% of the respondents are classified as having an unacceptable level of consistency. For the mixture, nearly four-fifths (80%) of the respondents have CR values below 0.20 and are judged to have a high or low/moderate consistency. The remaining (21.2%) respondents have CR values above 0.2, and are classified as completely inconsistent.

Table 24

Interrelation between Task Order and Consistency Classes (AHP)

	Offline			Online			Mixed		
		Task order			Task order			Task order	
Consistency classes	Total (n=181)	AHP first (n=97)	CA first (n=84)	Total (n=504)	AHP first (n=251)	CA first (n=253)	Total (n=685)	AHP first (n=348)	CA first (n=337)
High CR $\leq$ 0.1	69 (38.1%)	38 (39.2%)	31 (36.9%)	214 (42.5%)	111 (44.2%)	103 (40.7%)	283 (41.3%)	149 (42.8%)	134 (39.8%)
Moderate/low CR Є (0.1,0.2]	70 (38.7%)	35 (36.1%)	35 (38.7%)	187 (37.1%)	95 (37.8%)	92 (36.4%)	257 (37.5%)	130 (37.4%)	127 (37.7%)
Inconsistency CR >0.2)	42 (23.2%)	24 (24.7%)	18 (21.4%)	103 (20.4%)	45 (17.9%)	58 (22.9%)	145 (21.2%)	69 (19.8%)	76 (22.6%)
ACR	.155	.164	.144	.146	.141	.152	.149	.147	.150
χ^2 (p value)	.637 (.727)			1.980 (.372)			.992 (.609)		
Cramer's V	.059			.063			.038		

Given that there may be a systematic task order, internal validities measured by CR were also decomposed on the combination of task order (AHP first or CA first) and degree of the respondents' consistency. Chi-squared tests were conducted to determine the extent to which task order effects exist. By this test (see Table 24), no significant order effects were found concerning the groups: offline mode χ^2 (2) =.637, p>.10; online mode χ^2 (2) =1.980, p>.10; mixture χ^2 (2) =.992, p>.10. Concerning the order in which the methods were applied, no clear effect can be found for AHP.

For RQ3 an ANOVA (analysis of variance) was performed to determine the difference between offline and online modes regarding the internal validity of AHP. According to Table 25 the average CR (ACR) was marginally lower for the online mode than for the offline mode (.121 vs. 130, respectively). Note that in the AHP a higher consistency ratio reflects more inconsistency. The difference was not significant, $F(1,683)=1.121$, $p=.290$, $\eta_p^2=.002$.

At the same time, an analysis of covariance (ANCOVA) was conducted to control for possible confounding of online/offline comparisons. Considering different sample characteristics, as stated previously, gender, age, hotel knowledge, and hotel involvement were included as covariates in the analysis. The purpose of ANCOVA is the following: to increase the precision of comparison between groups by reducing within-group error variance; and, to "adjust" comparisons between groups for imbalances by eliminating confounding variables (Polit and Beck, 2008).

The marginal means according to survey modes are also summarized in Table 25. The ANCOVA test again shows similar results. The marginal mean of the CRs for the offline mode is .130 and .120 for the online mode. There again were no significant differences between them regarding internal validity measured by CR, $F(1,679)=.716$, $p=.398$, $\eta_p^2=.001$. The covariates had no effect on the results. It can be overall concluded that the two survey modes are generally equivalent in regards to the internal validity of AHP.

Table 25

ACR as a Measure of Internal Validity of AHP

	Offline	Online	F-value (p-value)	η_p^2
Internal Validity (ANOVA)	.130	.121	1.121 (.290)	.002
Internal Validity (ANCOVA)	.130	.120	.716 (.398)	.001

Note. CR was non-normally distributed due to significant kurtosis; therefore, it was square-root transformed. For ease of interpretation, untransformed means were reported in the table.
Lower CR indicates better internal validity.

RQ4: Do both the offline and online data collection modes lead to comparable results regarding the internal validity of CA?

The internal validity of a conjoint model measures the goodness of the model. It is defined as the correlation between the observed and estimated preferences presented by Pearson's r or Kendall's tau (τ) (Green and Srinivasan, 1978). According to Baier and Säuberlich (1997), the internal validity (consistency) is the correlation between the estimated utility values and the observed utilities in the calibration question, which is measured by the determination coefficient R^2 (percent of total variation in the preferences (utilities) explained by the model). This gives a first indication to the reliability of judgments (Melles, et al., 2000).

Depending on the degree of consistency of respondents' judgments, the fit values of R^2 range from a low of 0 to a high of 1.0. The larger R^2 is, the more consistent are their judgments and the better are they represented by the additive linear utility function and its estimated parameters. Thus, higher values of R^2 indicate a higher internal validity. According to Scholl et al. (2005), R^2 values between 0.9 and 1 indicate that most variances are explained and a high level of consistency is achieved. The values with smaller than 0.7 indicate that the conjoint model could not explain the variances sufficiently and there is an unacceptable level of consistency. As a threshold for acceptable consistency, a value of 0.7 is frequently proposed by Scholl et al. (2005).

Table 26 shows the average values of R^2 for CA for the total sample according to survey modes. In all cases, the average R^2 varies between .854 and .921; thus, all values far exceeded that cutoff point of .7 suggested by Scholl et al. (2005), indicating a satisfactory internal validity for CA. Table 26 also presents internal validities measured by R^2 separately for respondents with unacceptable, low or moderate, and high consistencies.

For the offline mode, almost half of respondents (51.9%) are judged to be highly consistent, and more than one-third of respondents (35.4%) are classified as having a low or moderate consistency. The remaining (12.7%) ones are considered to have an unacceptable level of consistency. For the online mode, about one third respondents (30.8%) are classified as having a high consistency, and almost half of respondents (49.6%) are considered to be low or moderate consistent in their judgments. The remaining (19.6%) of respondents are judged to be completely inconsistent. For the mixture, the majority of respondents (82.2%) are judged to have a high or low and moderate level of consistency based on the condition $R^2 \geq 0.7$, while only 17.8% of respondents are classified as having an unacceptable level of consistency.

Table 26

Interrelation between Task Order and Consistency Classes (CA)

	Offline			Online			Mixed		
		Task order			Task order			Task order	
Consistency classes	Total (n=181)	AHP first (n=97)	CA first (n=84)	Total (n=504)	AHP first (n=251)	CA first (n=253)	Total (n=685)	AHP first (n=348)	CA first (n=337)
High	94	50	44	155	76	79	249	126	123
$R^2 \geq 0.9$	(51.9%)	(51.5%)	(52.4%)	(30.8%)	(30.3%)	(31.2%)	(36.4%)	(36.2%)	(36.5%)
Moderate/low	64	40	24	250	124	126	314	164	150
R^2€ [0.7,0.9)	(35.4)	(41.2%)	(28.6%)	(49.6%)	(49.4%)	(49.8%)	(45.8%)	(47.1%)	(44.5%)
Inconsistency	23	7	16	99	51	48	122	58	64
$R^2 < 0.7$	(12.7%)	(7.2%)	(19.0%)	(19.6%)	(20.3%)	(19.0%)	(17.8%)	(16.7%)	(19.0%)
R^2	.918	.921	.914	.857	.854	.857	.880	.878	.880
χ^2	.701			.157			.779		
(p value)	(.030)*			(.924)			(.677)		
Cramer's V	.197			.018			.034		

Note. Average Rs are computed via averaging the Fisher z-transformed r

*Significant at $p<.05$

To check for any possible order effects concerning the assessed groups, internal validities (R^2) were also decomposed according to the task order (AHP first or CA first). Using chi-squared tests (see Table 26), a significant order effect was found for the offline mode, with a small effect size (χ^2 (2) =.7007, p<.05, Cramer's V=.197). In particular, a slight order effect appears to be observable for respondent with moderate and low consistency if the AHP task is completed first. Applying AHP before seems to have a positive effect on applying CA, resulting in a slight improvement of the internal validity of CA (.91 to. 92, +1.099%). This small difference could be due to a learning effect by AHP. On the other hand, no significant order effects were found concerning the assessed groups for the online mode (χ^2 (2) =.157, p>.10) as well as overall for mixture (χ^2 (2) =.779, p>.10).

To address RQ4, an ANOVA was utilized on the means of the Fisher z-transformed square roots of the R^2 values following a Gaussian distribution (Scholz, 2010). The ANOVA results (see Table 27) revealed that there was a significant difference between the survey modes regarding internal validity measured by Pearson's R (F (1, 683)=36.405, p<.01, η_p^2=.051), with the correlation coefficient being slightly higher for the offline mode (.961 vs. .927).

In the following, an ANCOVA was applied to control for pre-existing group differences. Again, demographics (gender and age), hotel knowledge and hotel involvement were used as covariates in the analysis to derive the marginal means. After adjustment for the covariates, the ANCOVA results (see Table 27) are similar to those drawn from the preceding analysis. The coefficient R (.964 vs. .925) was significantly higher for the offline mode than for the online mode, $F(1,679) = 29.497$, $p < .01$, $\eta^2 = .042$. So it can be overall concluded that the offline-conjoint analysis leads to higher values in regards to the internal validity which was measured by Pearson's R.

Table 27

Pearson's R as a measure of Internal Validity of CA

	Offline	Online	F-value (p-value)	η_p^2
Internal Validity (ANOVA)	.961	.927	36.405 (.000)*	.051
Internal Validity (ANCOVA)	.964	.925	29.497 (.000)*	.042

Note. *Levene's test was significant (p=.000) thus the assumption of equal variance was violated. However, Analysis of variance is robust to violations of the assumption of homogeneity of variances provided the ratio of the largest group variance is not more than 3 times the smallest group variance (Aurah, 2013). In the present study, the ratio was 1.649 less than the rule of thumb of 3.0.
*The Fisher z-value is not valid when r = 1 exactly, and thus approximating 1 as 0.999 is accepted as valid quantitative work (Faller, 1981). The error in r is about .001 and negligible.
*Significant at p<.001

To summarize, in case of AHP, both data collection methods lead to comparable results concerning internal validity, whereas in the case of CA, online methods generate less valid results. This result suggests that AHP compared with CA at least seems to be very robust with respect to the consistency across the two data collection modes. This might be explained by the lower cognitive strain on the data-supplying capabilities of respondents for the AHP approach than for the CA approach being more complicated and more cognitive effort being required from the respondents, since reducing respondents' cognitive load is particularly key for web-based environments, where respondents' patience tends to be low (Deutskens et al., 2004). The positive effect of ordering AHP before CA seems to confirm this explanation.

Although one measure of reliability is internal consistency, it should be cautious in using ACR and R^2 as sole measures for model evaluation. Reliability is only a necessary condition for validity. As Wittink and Walsh (1988) point out, higher reliability does not necessarily mean higher validity. It is possible that the structure of the evaluation task is such that it triggers a task simplification that results in consistent responses but does not resemble the behavior of the respondent in the market (Vriens, 1995). For example, respondents may key in on a few salient attributes and ignore the others (Gilbride and Allenby, 2004). To extent that such strategies do not mimic real market place behavior. Such task simplification behavior will affect the external validity of the conjoint results negatively (Vriens, 1995). This additional aspect of internal validity is considered by the further tests which also serve as a basis for evaluating the predictive validity of both models.

Predictive Validity

RQ5: Do both the offline and online data collection modes lead to comparable results regarding the predictive validity of AHP and CA?

RQ6: Are both AHP and CA fairly comparable in predictive performance across offline and online data collection modes?

RQ7: Do both the offline and online data collection modes moderate the differences in predictive accuracy among the AHP and CA?

The most common measure of performance used in empirical studies of conjoint analysis is predictive validity (Akaah, 1991; Mulye, 1998). Predictive validity refers to a method’s ability to predict real choices (Helm et al., 2004b). Holdout task validation is commonly used in full-profile conjoint and in choice based conjoint studies (Bakken and Frazier, 2006). The data for these tasks is kept out of the estimation step. In empirical studies the predictive validity is investigated by means of holdout choice, holdout rating or ranking tasks (Johnson, 1997; Wittink and Bergestuen, 2001). The results of Huber and Hansen (1986) showed no significant differences between the three tasks regarding predictive validity. However, McCullough (2002) suggests that holdout tasks should be choice-based to make model validation more meaningful. Thus, holdout choice tasks were used to analyze the predictive validity for both the AHP and CA, since choices more directly mirror marketplace decisions and relatively simple tasks for respondents (Desarbo et al., 1995). A typical holdout choice task consists of two or more alternatives, from which respondents choose one (Wittink and Bergestuen, 2001).

Several factors can affect the results of holdout task validation, including the choice model such as first choice, share of preference, or logic as well as the composition of the holdout task (Bakken and Frazier, 2006). For that reason, using two or more holdout tasks are recommended so that the analyst has several opportunities to determine the predictive accuracy of the conjoint results (Wittink and Bergestuen, 2001). In this regard, ten holdout choice sets containing four alternatives were constructed by random allocation of the profiles to the choice sets to limit the potentially biasing effects of the choice set composition. The holdout choice sets were constructed to be Pareto optimal; that is, no alternative was dominated by any other alternative in the choice set, as also recommended by Elrod et al. (1992).

Two measures of the validity of holdout results are commonly used (Wittink and Bergestuen, 2001). One measure is defined at the level of the individual respondent. It assesses how well the conjoint results can predict each individual’s holdout choices. The common summary measure for this is the proportion of hits, where a hit is a choice correctly predicted. The hit rate is used as an indicator for the power of predicting the real choice (Klein, et al., 2010) and is one of the criteria very often used to measure predictive validity (Vriens et al., 1998; Brusch et al., 2002; Tscheulin, 1991). Hit rates are computed by comparing the choice predicted for an individual respondent by the model using the maximum utility rule to the actual choice made by the respondent. In any case, a satisfactory predictive validity depends on high values of the hit rate. The result obtained on this measure is usually compared against what would be expected in the absence of information, that is random choices (Huber et al., 1993).

The other measure is defined at the aggregate level. In this case, the proportion of choices for each holdout alternative is compared with the proportion of predicted choices. A measure of prediction error is the deviation between holdout shares and predicted shares. To determine the quality of aggregate predictions in holdout tasks, the result can be compared against the expected result based on random choices (the minimum) (Huber et al., 1993).

Although these two summary measures tend to be positively related, they can conflict. Hit rates depend primarily on the reliability of individual models, whereas choice share estimates, by aggregating over individual estimates, depend mainly on the degree to which the models provide unbiased predictions.

Individual-Level Choice Share Prediction

The issue of comparability of the predictive validity of the AHP and CA across the two data collection modes was examined by computing the hit rate separately for each respondent. The average hit rate for each model-mode combination was then computed.

The relative hit rates of the models. Table 28 compares individual-level hit rate validations for both preference measurement methods between the two data collection modes. It can be observed from Table 28 that the results for the offline and online modes were mixed. For the offline mode, the average hit rate of CA (40.7 percent) is slightly superior to the hit rate of AHP (38.6 percent), whereas for the online mode AHP (38.4 percent) has a higher hit rate validation than CA (36.7 percent). To provide an overall indication of the predictive validities, the data were pooled across the two survey modes, namely, a mixture. For the mixture AHP has a 38.5 percent and CA a 37.7 percent hit rate, with a slight advantage for AHP. Table 28 also indicates that the offline mode reflected the highest hit rate for both AHP and CA.

As a first step, for comparison, both methods predict choices significantly better than the random model ($p \leq .001$) in any case, and the size of the improvements over a random model is quite similar for the methods and ranged from 15.6 to 20.9 percent. Thus, both methods perform reasonably well for both survey modes.

Table 28

Individual-level Choice Predictions (Before Adjustment for the Covariates)

	AHP	CA	Average across models	Experimental factors	F –value (p-value)	Partial η^2
Offline	38.6 (18.1)	40.7 (20.9)	39.7	Model type	F=.028 (p=.867)	.000
Online	38.4 (17.9)	36.7 (15.6)	37.5	Survey mode	F=2.128 (p=.145)	.003
Average between survey modes	38.5 (18.0)	37.7 (16.9)	38.6	Interaction	F=3.484 (p=.062)	.005

Note. Percentage improvement over random model in parenthesis: [100*(percent correctly predicted-percent correctly predicted by random model)/(100-perent correctly predicted by random model)] (Srinivasan and Park, 1997).

To assess the statistical significance of the differences in model performance, a two-way mixed ANOVA was performed. Between-subjects factor was experimental condition, with two levels (offline vs. online), and the within-subjects factor was model type, with two levels (AHP vs. CA). Accordingly, the four experimental conditions and their interactions served as predictor variables to test where any of these conditions moderate the differences in predictive accuracy between the models. The results of the mixed ANOVA (see Table 28) showed no significant effects for data collection mode nor for model type on the predictive accuracy: $F(1, 683)=2.128$, $p>.05$; $F(1,683)=.028$, $p>.05$, respectively. There again was no significant interaction between data collection mode and model type, $F(1,683)=3.484$, $p>.05$. Thus, the results implied the lack of major differences in model performance across the two data collection modes as well as the two models examined.

Again, these results should be taken with care due to differences in sample composition between the two survey modes. As a further analysis, a two-way mixed ANCOVA was performed with the demographic variables that were significantly different across samples as covariates (e.g., gender, age, knowledge, and involvement). The result from the ANCOVA after controlling for the covariates (see Table 29) identified a significant main effect for data collection mode ($F(1,679)=18.292$, $P<.01$). The effect of model type was not significant. ($F(1,679)=.238$, $p>.10$). The interaction between data collection mode and model type again was not significant ($F(1,679)=2.635$, $p>.05$). Thus, taking into account the differences between the offline sample and the offline sample leads to a conclusion that the two survey modes are not equivalent, with the offline mode being judged superior to online mode regarding the predictive validity of both AHP and CA.

Table 29

Individual-level Choice Predictions (After Adjustment for the Covariates)

	AHP	CA	Average across models	Experimental factors	F –value (p-value)	Partial η^2
Offline	.426	.450	.438	Model type	F=.238 (p=.626)	.000
Online	.370	.351	.360	Survey mode	F=18.292* (p=.000)	.026
Average across survey modes	.398	.401	.399	Interaction	F=2.635 (p=.105)	.004

Note. Marginal means are given in the table.
*significant at p<.01

Impact of task order on hit rates. In order to examine sequence (order) effects, the hit rates according to survey modes were further decomposed by task order (AHP first or CA first). Table 30 shows the hit rates separately for the estimated part-worths obtained from the first and the second AHP or conjoint task completed by the respondent.

For the offline mode, there is no clear sequence effect can be found for AHP and CA, whereas for the online mode both approaches lead to slight predictive improvements. For AHP, its hit rate is slightly higher (39.2% vs. .37.7%, 3.98% improvement**)** when CA is the first task, whereas CA gets the higher hit rate (37.3% vs. 36.1%, 2.96% improvement) when the AHP task is administered first. For the mixture, apparently, the task order has no virtually effect on the AHP hit rates (38.3% vs. 38.6%). On the other hand, a positive sequence effect can be recognized for CA. The CA hit rates are 38.1 percent if AHP is completed first and 37.4 percent if it is the second task (1.88% improvement in hit rates). One interpretation of this difference is that the AHP task provides a training opportunity that enhances the consistency of subsequent CA evaluations. Of course, in practice one would use either CA or AHP, and in that context only the results for the first task are relevant. Thus, AHP is substantially better if only one task is administered to each respondent. It is conceivable; however, that CA improves if a practice or warm-up task is administered first.

Table 30

Impact of Task Order on Hit Rates

	Model type					
	AHP	CA	Average across models	Experimental factors	F –value (p-value)	Partial η^2
Offline						
Task Order						
AHP first	.399	.401	.400	Model type	F=1.505 (p=.221)	.008
CA first	.371	.414	.393	Task order	F=.093 (p=.761)	.001
Average across task orders	.385	.407	.396	Interaction	F=1.238 (p=.267)	.007
Online						
Task Order						
AHP first	.377	.373	.375	Model type	F=.2.781 (p=.096)	.006
CA first	.392	.361	.376	Task order	F=.012 (p=.914)	.000
Average across task orders	.384	.367	.375	Interaction	F=1.645 (p=.200)	.003
Mixed						

Task Order						
AHP first	.383	.381	.382	Model type	F=.663 (p=.416)	.001
CA first	.386	.374	.380	Task order	F=.012 (p=.913)	.000
Average across task orders	.385	.377	.381	Interaction	F=.313 (p=.576)	.000

Note. Marginal means are given in the table

In line with these descriptive results, to check for significant differences a two-way mixed ANOVA was applied separately for each mode. In each case the within-subjects factor was model type (AHP vs. CA) and the between-subjects factor was task order (first vs. second). The four experimental conditions and their interaction served as predictors to test whether any of these conditions moderate the differences in accuracy between the models. By this test, no significant effects ($p>.05$) for model type nor for task order, and no significant interaction effects ($p>.05$) exist in the proportion of hit rates in all three cases (see Table 30). Thus, although order appears to have some slight effect on hit rates, no statistically significant effect was found.

Impact of the consistency on hit rates. A further comparison of AHP and CA with respect to the robustness against inconsistencies in choices was conducted. Respondents were also grouped by the degree of consistency in choices. Tables 31 and 32 show predictive validities separately for respondents with inconsistency, low/ moderate, and high consistencies based on CR and R^2, respectively.

In case of AHP, a positive correlation can be observed between the level of consistency and the predictive validity in all three cases (see Table 31). While the most consistent respondents with $CR \leq .1$ have considerably high hit rates, the group of respondents with moderate and low consistency ($CR \in (0.1, 0.2]$) gets lower hit rates.

Examining the group of considerably inconsistent respondents (CR >0.2) shows a further reduction in predictive power. To test the effect of the consistency in choices on the predictive validity of AHP, a one-way ANOVA was conducted for each of the data collection modes. The results of the ANOVA are reported in Table 31.

For the offline mode, very surprisingly, none of these differences between the assessed groups are significant (Welch's $F(2, 99.049)=2.177$, $p>.05$). Even for respondents with high inconsistency, the predictive validity is hardly worse than for respondents with low and moderate consistency. Thus, only a minor influence of the degree of consistency on the hit rates of AHP can be found. For both the online and mixtures; however, there were significant differences with respect to the hit rates among the groups ($F(2,501)=5.808$, $p<.01$; $F=(2,682)=8.408$, $p<.01$, respectively). Post-hoc t-tests revealed that there was no significant difference between the groups with a low/moderate (CR Є (0.1, 0.2])) and high level (CR≤.1) of consistency with respect to the hit rates produced ($p>.05$), but the group of inconsistent respondents (CR>.2) produced considerably less hits ($p<.05$)**.**

Table 31

Impact of the Consistency on Hit Rates (AHP)

Analytic hierarchy process	High consistency (CR ≤ 0.1)	Moderate/low consistency CR Є [0.1,0.2)	Considerably inconsistent (CR>0.2)	F tests (df1, df2)	p-value	Partial Eta Squared
Offline	.4174	.3886	.3310	F=2.177* (2, 99.049)	.119	*est.* ω^2=.013
Online	.4083[a]	.3879[a]	.3270[b]	F=5.808 (2, 501)	.003	η_p^2=.023
Mixed	.4105[a]	.3881[a]	.3282[b]	F=8.408 (2, 682)	.000	η_p^2=.024

Note. [a, b] significantly different groups at p<.05.
* Welch correction for non-homogeneity of variance was applied (Levene's F-test, p< .05).

In case of CA a positive correlation is even more obvious for both modes and the mixture (see Table 32). If the consistency is low (unreliable choices), the predictive validities are equally low. Inversely, for respondents with high consistency, the predictive validities are equally high. Given a certain minimum level of consistency ($R^2 > 0.7$), the power of predicting the real choice is very acceptable, but the differences in predictive validity noticeably occur for respondents with low/moderate ((R^2 Є [0.7, 0.9)) and high ($R^2 \geq 0.9$) consistencies in choices. To test the effect of different consistency groups on the predictive validity of CA, a one-way ANOVA was also performed for both modes and the mixture.

The results of the ANOVA (see Table 32) revealed that there are significant differences among the assessed groups in terms of the hit rate of individual choices: offline mode $F(2,178)=9.579$, $p<.01$; online mode Welch's $F(2,257.526)=35.553$, $p<.01$; mixture Welch's $F(2, 346.627)=49.487$, $p<.01$. Post-hoc t-tests indicated that the group of high level consistent respondents ($R^2 \geq 0.9$) produces significantly better hit rates than the other two groups for both modes and the mixture ($p<.05$). Hit rates are slightly higher for the group with low/moderate consistency ((R^2 Є [0.7, 0.9)) than for the inconsistent group, but the difference is less significant ($p>.05$).

Table 32

Impact of the Consistency on Hit Rates (CA)

Conjoint Analysis	High consistency $R^2 \geq 0.9$	Moderate/low consistency R^2€ [0.7,0.9)	Considerably inconsistent $R^2 < 0.7$	F tests (df1, df2)	p-value	Partial Eta Squared
Offline	.4697[a]	.3484[b]	.3130[b]	F=9.579 (2, 178)	.000	η_p^2=.097
Online	.4791[a]	.3265[b]	.2928[b]	F=35.553* (2, 257.526)	.000	*est.* ω^2=.121
Mixed	.4756[a]	.3309[b]	.2966[b]	F=49.487* (2, 346.627)	.000	*est.* ω^2=.124

Note. [a, b] significantly different groups at p<.05.
* Welch correction for non-homogeneity of variance was applied (Levene's F-test, p< .05).

These findings imply that as the consistency in choices increases, the predictive validities increase for the two models. However, CA is not as robust as AHP with respect to inconsistencies in choices, because CA is more sensitive to small reductions in the level of consistency. This lack of robustness is responsible for the better performance of AHP, and this can be seen as an advantage of AHP after taking into account the cost of interviewing and evaluating.

Aggregate-Level Choice Share Prediction

Although measuring predictive accuracy at the individual level constitutes an often used approach, one may argue that a more aggregate criterion, such as the accuracy of the resulting choice share predictions, would be a more relevant criterion for marketing managers (Vriens, 1995). Accordingly, additional analyses to test for predictive validity were necessary at the aggregate level. Most validation studies rely on aggregate choice share predictions based on appropriate holdout stimuli (Green and Srinivasan, 1990), although out-of-sample predictions (external validity) are often considered the silver bullet in validating different preference measurement approaches (Ding, Grewal, and Liechty, 2005; Scholz et al., 2010).

At the aggregate level the predictive validity can be measured in several ways. The Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE) are two common measures that quantify the predictive validity (Klein et al., 2010). Both are aggregate measures of how well a model predicts choices. The MAE is less sensitive to outliers and more robust over a variety of induced error distributions (Hoaglin, Mosteller, and Tukey 1983; Tukey, 1960), and thus treats all errors equally (Moore et al., 1998). In contrast, the RMSE treats larger errors more seriously than small errors, since it considers squared errors (Moore et al., 1998), and thereby is most relevant if small errors are of minor importance compared to large prediction errors. Because the two performance measures tap different aspects of predictive accuracy, they can be used together to diagnose the variation in the errors. Thus, choice shares are validated in terms of both MAE and RMSE of predicted versus actual choice shares over all alternatives and choice sets. Lower values of MAE and RMSE indicate a higher predictive validity.

As in many other comparative studies, in the present study aggregate choice shares were predicted applying the well-known first-choice model, the multinomial logit (MNL) model, and the powered (rescaled) logit model (see, e.g., Moore, 2004). Furthermore, for all choice models correlations between choice shares observed in the holdout task and the predicted shares were computed for each of the two approaches under consideration. These results are presented in Table 33.

According to Table 33 all choice models show remarkable correlations for both models (AHP and CA) and for both survey modes within the range of .580 to .802, all significant at the .01 level. In terms of the performance measures MAE and RMSE, the choice share predictions provide accurate results for both models (AHP and CA), compared to a random model. Besides, it is observable that the error measures (both MAE and RMSE) of the offline survey are slightly lower than the error measures of the online survey for both models. At first hand, this seems to support the previous findings that the data of the offline survey leads to slightly better results.

To assess the statistical significance of the differences in model performance, a two-way mixed ANOVA was performed. Between-subjects factor was experimental condition, with two levels (offline vs. online), and the within-subjects factor was model type, with two levels (AHP vs. CA). Accordingly, the four experimental conditions and their interactions served as predictor variables to test where any of these conditions moderate the differences in predictive accuracy between the models. The results of the mixed ANOVA are summarized in Table 33.

Table 33

Aggregate Choice Share Validations

	First choice model			Logit model					
	-			MNL			Powered (rescaled)		
Offline	Pearson (Kendall)	MAE	RMSE	Pearson (Kendall)	MAE	RMSE	Pearson (Kendall)	MAE	RMSE
AHP	.580 (.466)	.069	.079	.628 (.471)	.081	.093	.629 (.466)	.078	.090
CA	.788 (.582)	.053	.062	.811 (.600)	.049	.057	.798 (.593)	.051	.060
Online									
AHP	.783 (.578)	.073	.084	.800 (.616)	.085	.098	.802 (.622)	.080	.093
CA	.752 (.547)	.055	.063	.768 (.571)	.054	.062	.762 (.565)	.054	.062
Mixed									
AHP	.701 (.524)	.071	.081	.725 (.544)	.083	.095	.727 (.541)	.079	.091
CA	.766 (.553)	.054	.063	.785 (.576)	.052	.060	.777 (.564)	.053	.061
Experimental factors	-	F value (p value) (η_p^2)	F value (p value) (η_p^2))	-	F value (p value) (η_p^2)	F value (p value) (η_p^2)	-	F value (p value) (η_p^2)	F value (p value) (η_p^2)
Model type	-	3.671 (.071) (.169)	3.091 (.096) (.147)	-	20.920 (.000)* (.538)	18.998 (.000)* (.513)	-	15.335 (.001)* (.460)	13.100 (.002)* .421
Survey Mode	-	.102 (.753) (.006)	.092 (.765) (.005)	-	.280 (.603) (.015)	.382 (.544) (.021)	-	.096 (.761) (.005)	.086 (.772) (.005)
Interaction	-	.017 (.897) (.001)	.025 (.876) (.001)	-	.001 (.978) (.000)	.004 (.953) (.000)	-	.001 (.972) (.000)	.000 (.992) (.000)

Note. *significant at $p<.001$

The results for the first choice model are very similar for both model types and both survey modes. The mixed ANOVA results indicated no main effect of model type (MAE F=3.671, p=.071, η_p^2=.169; RMSE F=3.091, p=.096, η_p^2=.147) or survey mode (MAE F=.102, p=.753, η_p^2=.006; RMSE F=.092, p=.765, η_p^2=.005), and a non-significant effect of the interaction between the two (MAE F=.017, p=.897, η_p^2=.001; RMSE F=.025, p=.876, η_p^2=.001).

When the MNL model was applied to predict choice shares, the results are in favor of CA. The mixed ANOVA tests confirmed that there was a significant main effect of model type on predictive validity (MAE $F=20.920$, $p<.01$, $\eta_p^2=.538$; RMSE $F=18.988$, $p<.01$, $\eta_p^2=.513$). Neither the survey mode main effect (MAE $F=.280$, $p=.603$, $\eta_p^2=.015$; RMSE $F=.382$, $p=.544$, $\eta_p^2=.021$ nor the interaction between model type and survey mode was significant (MAE $F=.001$, $p=.978$, $\eta_p^2=.000$; RMSE $F=.004$, $p=.953$, $\eta_p^2=.000$). Thus, the results show that CA yields better choice share predictions than AHP. It was also confirmed that there are no big differences in the validity between the two data collection modes.

The mixed ANOVA results for the powered logit model (rescaled) showed a significant main effect of model type (MAE $F=15.335$, $p<.01$, $\eta_p^2=.460$; RMSE $F=13.100$, $p<.01$, $\eta_p^2=.421$). The main effect of survey mode was not significant (MAE $F=.096$, $p=.761$, $\eta_p^2=.005$; RMSE $F=.086$, $p=.772$, $\eta_p^2=.005$). The interaction between model type and survey mode also was not significant (MAE $F=.001$, $p=.972$, $\eta_p^2=.000$; RMSE $F=.000$ $p=.992$, $\eta_p^2=.000$). Again, CA produces lower MAE and RMSE (better predictive validity) than AHP. Besides, both data collection modes lead to the comparable results regarding predictive validity.

In summary, a comparison of the results of the first choice model with those of the logit model reveals slightly greater MAEs and RMSEs for the later model for AHP. More specifically, AHP validates much better with the first choice model than the logit model. Thus, it can be concluded that for AHP the first choice model is more appropriate for predicting real choice shares. On the other hand, across all measures and ways of estimating shares, there is no clear pattern for CA. Thus, CA has approximately equal validations under the max utility and logit models. More notably, there is no difference between the survey methods (offline and online) regarding predictive validity measured by the MAE and RMSE.

Table 34 provides a brief summary of the research results. In the following chapter, the research results are discussed and the conclusions are presented, followed by the implications, and directions for further research.

Table 34

Summary of Results

Research Questions	Results
RQ1: Are there any differences in the respondents' subjective evaluations of the methods in terms of (a) enjoyment, (b) difficulty and clarity, and (c) realism?	The AHP survey was rated as a more positive experience than the CA approach. The AHP survey was rated as being (a) more enjoyable, (b) less difficult and complex and (c) more realistic. Significant differences were found concerning the average survey length. AHP has considerable advantages over CA in terms of time effort and costs.
RQ2: Do AHP hotel branding results accord generally with CA (convergence)? If so, to what extent does AHP have convergent validity with CA with respect to (a) importance ratings, (b) part-worth estimations, and (c) estimated overall utilities?	The convergent validity gave mixed results: (a) Almost no convergent validity: the models result in different preference structures. Further AHP importance weights tend to show more curvature than CA importance weights (b) high convergence (c) high convergence
RQ3: Do both the offline and online data collection modes lead to comparable results regarding the internal validity of AHP?	No significant difference between the survey methods concerning internal validity
RQ4: Do both the offline and online data collection modes lead to comparable results regarding the internal validity of CA?	The validity of the data of the online-conjoint analysis is slightly low.
RQ5: Do both the offline and online data collection modes lead to comparable results regarding the predictive validity of AHP and CA?	Overall, the data gathered online leads to a slightly lower predictive validity. Nevertheless, the validity seems to be sufficient even in the case of its online form.
RQ6: Are both AHP and CA fairly comparable in predictive performance across offline and online data collection modes?	Overall, the AHP does equally as well as conjoint analysis for predictive validity.
RQ7: Do both the offline and online data collection modes moderate the differences in predictive accuracy among the AHP and CA?	No moderating effect of survey method on the models' predictive performance.

Hotel Branding Results

Table 35 presents a summary of the hotel branding results. The table illustrates the relative importance of attributes and levels obtained from both methods (AHP and CA) as well as the ranges of attribute importance weights in last row of the table.

Table 35

Relative Importance Weights and Part-worth Utilities for AHP and CA

Attributes	Offline (n=181) Relative importance		Online (n=504) Relative importance		Attribute Levels	Offline (n=181) Part-worth utility estimates		Online (n=504) Part-worth utility estimates	
	AHP	CA	AHP	CA		AHP	CA	AHP	CA
Brand Awareness	0.183 (4)	0.324 (1)	0.167 (4)	0.258 (2)	Advertising	0.199 (4)	0.104 (4)	0.144 (4)	0.196 (3)
					Top-of-mind brand	0.248 (3)	0.251 (3)	0.249 (2)	0.162 (4)
					Brand popularity	0.256 (2)	0.291 (2)	0.230 (3)	0.230 (2)
					Brand Familiarity	0.296 (1)	0.354 (1)	0.377 (1)	0.412 (1)
Brand Image	0.213 (3)	0.298 (2)	0.193 (3)	0.295 (1)	Clean	0.307 (1)	0.340 (1)	0.252 (2)	0.255 (2)
					Elegant	0.228 (3)	0.235 (3)	0.175 (4)	0.208 (3)
					Feels like home	0.189 (4)	0.090 (4)	0.244 (3)	0.143 (4)
					Good value	0.276 (2)	0.335 (2)	0.329 (1)	0.393 (1)
Perceived Quality	0.376 (1)	0.198 (3)	0.376 (1)	0.241 (3)	Error-free service	0.227 (3)	0.098 (4)	0.179 (4)	0.107 (4)
					Prompt service	0.211 (4)	0.304 (2)	0.203 (3)	0.357 (1)
					Courteous service	0.272 (2)	0.348 (1)	0.279 (2)	0.257 (3)
					Reliable service	0.290 (1)	0.250 (3)	0.339 (1)	0.279 (2)
Brand Loyalty	0.227 (2)	0.179 (4)	0.264 (2)	0.206 (4)	Friends' recommendation	0.235 (3)	0.171 (3)	0.176 (3)	0.200 (3)
					Frequent customer	0.324 (2)	0.325 (2)	0.303 (2)	0.267 (2)
					Previous experience	0.440 (1)	0.504 (1)	0.521 (1)	0.533 (1)
Range*	0.193	0.145	0.209	0.089					

Notes. The ranking of attributes and attribute levels is depicted in parentheses.
*Range= (first ranked importance weight – last ranked importance weight)

Offline Study Results

The results of the offline study indicated that the results of the AHP which showed perceived quality (37.6%) as the most important attribute college students considered when choosing a hotel brand, followed by brand loyalty (22.7%), brand image (21.3%) and finally brand awareness (18.3%), while in the CA results, brand awareness was the most important (32.4%), followed closely by brand image (29.8%), perceived quality (19.8%) and then brand loyalty (17.9%).

Within the attribute of brand awareness, brand familiarity was found to be the most important level followed by brand popularity, top-of-mind awareness and advertising in both methods. Within the brand image attribute, cleanliness and good value for money were found to be the most important levels for both methods as compared to elegant atmosphere as well as feels like home. Within the attribute of perceived quality, the AHP results indicated that reliable service was found to be the most important level and prompt service was the least. In the CA results, courteous service was found to be of the most importance to respondents. Prompt service was also of considerable importance. Interestingly, error-free service was found to be considerably less important. Within the brand loyalty attribute, the most important level was previous experience (such as satisfaction), followed by frequent guest program (such as hotel reward card), and then recommendation by family, friends, relatives, and others (such as positive word-of-mouth recommendation) in both methods.

Online Study Results

The results for the online study of the AHP method showed that perceived quality (37.6%) and brand loyalty (26.4%) were rated the top two priorities by domestic travelers and brand awareness (16.7%) was the lowest priority by the same travelers (see Table 35). In contrast, CA results indicated that brand image (29.5%) and brand awareness (25.8%) were the two most important determinants of hotel brand equity followed by perceived quality (24.1%) and brand loyalty (20.6%).

A closer examination of the AHP results indicated that within the attribute of brand awareness, brand familiarity was found to be of the most importance to respondents as compared to advertising. Within the brand image attribute, good value for money was found to be the most important level. This level was followed by cleanliness, feels like home and elegant atmosphere. Within the perceived quality attribute, reliable service was found to be of the most importance to respondents and error-free service was the least. Finally, previous experience was found to be the most important within the attribute of brand loyalty. The CA results revealed brand familiarity, good value for money, prompt service, and previous experience had the highest importance within each attribute.

In summary, the order of preferred levels for each attribute was similar for all attributes. The most preferred level for each attribute was the same in both methods: brand familiarity, cleanliness, good value for money, and past experience. Interestingly, the attribute of perceived quality was a minor exception where reliable service has the top priority in the AHP results, whereas prompt and courteous service were of greatest concern in the CA results. However, with respect to the importance of attributes the results are totally different. In the AHP results, perceived quality and brand loyalty are the most important determinants of hotel brand equity and account for more than 60 percent of the total importance. However, in the CA results, brand awareness and brand image show the top two attributes of hotel brand equity, which account for about 60 percent of the total importance in both the offline and online studies.

CHAPTER 5

DISCUSSION AND CONCLUSIONS

Chapter 5 first compares and contrasts the research findings with other relevant research and discusses the interpretations of the findings, conclusions, implications, and limitations. Conclusions about hotel branding and recommendations for future research are also presented in this chapter.

Discussion and Comparison of Previous Research

This study compared the results of AHP and CA applied to hotel branding with respect to feasibility and several validity measures. To be able to draw conclusions about the quality of the results of AHP and CA as instruments for measuring preferences, the results have to be compared with results of other studies.

Regarding the results on hotel branding, the AHP and the CA generated totally different statements with respect to the importance of attributes. The AHP results showed that perceived quality was by far the most important factor for developing hotel brand equity, followed by brand loyalty, brand image and then brand awareness. However, the CA results revealed brand awareness and brand image were aspects of vital importance when building hotel brand equity but perceived quality and brand loyalty, respectively, were less effective on the hotel brand equity. Practical differences in attribute preferences were clearly observed between the AHP and the CA. In order to assess the factual quality of attribute importance weights (attribute preferences), this study verified the present results by considering previous empirical studies in the field of hotel branding. To date, there have only been a small number of studies that have explored branding in a hotel context (Kim, Kim, and An, 2003; Kayaman and Arasli, 2007; Kim and Kim, 2005; Bailey and Ball, 2006; So and King, 2010; Kim and Kim, 2007; Sun and Ghiselli, 2010; Nel, North, Myburg, and Hern, 2009; Zhou and Jiang, 2011; Baldauf, Cravens, and Binder, 2003; Kim, Jin-Sun, and Kim, 2008).

The results of this study are comparable to the results of previous studies. Kim, Kim, and An (2003) and Kim and Kim (2005) identified four dimensions of brand equity including brand loyalty, brand awareness, perceived quality, and brand image to examine the underlying dimensions of brand equity as well as how they affect financial performance of firms in the luxury hotel sector. The results indicated that in sequence of degree of significance, perceived quality, brand loyalty, and brand image were important components of customer-based hotel brand equity, where perceived quality was most important and brand awareness was least significant for establishing brand equity in luxury hotels. The authors further indentified brand loyalty as having the strongest influence on the hotel firms' performance, whereas brand image was found to have no significant effect on financial performances of hotels. Similar to Kim and Kim's (2005) study on luxury hotels, in a study by Kayaman and Arasli (2007) perceived quality, brand loyalty, and brand associations were found to be the core components of customer-based brand equity for five-star hotels, where perceived quality was found to be the most influential factor on brand equity and brand awareness was not a key dimension of hotel brand equity. This result was consistent with the study by Bailey and Ball (2006) who stated that having a brand-name (brand awareness) alone was not a guarantee of success within the hotel industry. This was further confirmed by the findings of So and King

(2010) that brand awareness was not a significant contributor to hotel brand equity. Kim and Kim (2007) found that brand loyalty, perceived quality, and brand associations were significant antecedents of overall brand equity of mid-priced hotels. In particular, perceived quality was found to be the most important determinant of mid-scale hotel brand equity followed by brand loyalty and then brand image, while brand awareness, a seemingly important source of brand equity, did not exert a significant influence on building brand equity of mid-priced hotels. This was further validated and supported by Sun and Ghiselli (2010) who found that perceived quality was the strongest predictor when determining brand equity in the lodging industry and suggested that the source of generating value of brand equity was perceived service quality. The authors further proposed that preserved quality plays a central role for building the four dimensions of brand equity in the hotel industry. The result was similar to the one in the study on three hotel categories (low, medium and high priced) by Nel, North, Myburg and Hern (2009), who reported that perceived quality was found to significantly affect a hotel's performance. In a similar vein, Zhou and Jiang (2011) showed that perceived quality and brand loyalty had a more positive effect on the customers' perceived value than brand awareness/brand association (image) and further highlighted that perceived quality was a direct determinant of revisit intentions in the budget/economy hotel segment and brand loyalty was an important dimension of brand equity amongst the sample visitors. The result was consistent with the value hotel chain study by Baldauf, Cravens and Binder (2003). In the research on mid-priced hotels by Kim, Jin-Sun, and Kim (2008) among four brand equity dimensions, perceive quality displayed the most powerful and dominant effect on customers' perceived value followed by brand loyalty, while brand awareness and brand association (image) did not affect perceived value in the mid-scale hotel industry and further were not direct influential factors on hotel revisit intentions. Insofar, the AHP results were largely consistent with the previous studies by awarding high importance to perceived quality and brand loyalty and lower importance to brand awareness and brand image.

Comparing the AHP and the CA, the results have demonstrated some similarities and differences in the obtained outcomes. The feasibility of the AHP and the CA was assessed by measuring the administration time, and by asking the respondents how they evaluated each task with respect to difficulty and clarity, enjoyment, and realism. Overall, respondents rated the AHP survey more positively than the CA survey. They felt that AHP's presentations were clearer, the survey was less difficult and boring, and that the format elicited more realistic answers. Interestingly, the CA is not rated as being more realistic despite the holistic approach of considering all attributes simultaneously. The results also show that the AHP significantly reduces the survey length compared with the CA since the AHP survey required less time to complete than the CA survey. This is confirmed by the findings of Scholl et al. (2005) that AHP showed a better applicability than CA in terms of motivation and time.

To test for the predictive validity, this study compared the AHP and the CA in terms of their ability to predict choices. Overall, the individual and aggregate-level analyses strongly suggest that for the current empirical application, the predictive validity of AHP and CA are almost equal. In terms of individual hit rates, the two approaches predicted about equally well on average. This finding is consistent with many studies (e.g., Mulye, 1998; Scholl et al., 2005; Meißner and Decker, 2009; Kallas et al., 2011; Ijzerman et al., 2012) which showed that AHP and CA have equal predictive validity. AHP is equivalent to CA with respect to aggregate choice share predictions, while Meißner and Decker (2009) found AHP even outperformed CA in choice share prediction. The two measures are useful together because they measure very different properties of preference elicitation methods. Hit rates depend primarily on the reliability of individual models, whereas choice share estimates, by aggregating over individual estimates, depend mainly on the degree to which the models provide unbiased predictions (Hagerty, 1986; Huber et al., 1993). Contrary to Hagerty (1986) and Moore et al. (1998) who found that there was a relatively weak relationship between aggregate and individual-level validation, in the current study still both AHP and CA validated well at both levels. It is gratifying that in the present study the results are relatively invariant across the two criteria.

In order to examine the impact of task order on hit rates, this study rotated task order by putting either AHP or CA first. Though not statistically significant at the 5% level, there seems to be a tendency that CA is likely to be more effective when it is preceded by the AHP task that familiarizes respondents with the attributes and their levels. The same logic holds true for the AHP, such that it leads to slightly better results when it is used after CA. These results imply knowledge of the methods and familiarity with the technique can play an important role in obtaining better results as mentioned by Helm et al. (2004b). As they suggested, it seems to be advantageous to explain the evaluation tasks to the respondents prior to the final experiment by some sort of warm-up exercise. However, as Hollings et al. (1998) have advised, the warm-up task should be not too difficult and promote their motivation to participate.

A further analysis investigated the correlation between the inconsistency of the respondents and the predictive power of a method. In this case, a weak but significant relationship was found indicating that increasing consistency improves the predictive validity for both methods. However, with respect to the robustness towards different consistency levels, AHP seems to be more robust than CA in that the latter is more vulnerable to small reductions in the level of consistency than the former. This is supported by empirical studies (Helm et al., 2003). More notably, AHP appears to be robust against judgment 'errors', such as inconsistent answers. These results are in line with Scholz, Meibner, and Wagner (2006)'s findings which demonstrated the superior robustness of the AHP in comparison to CA in the case of fuzzy and ambiguous preference judgments. In this perspective, AHP may have a slight edge in the case of ambiguous statements.

The convergent validity was measured thorough Spearman's coefficient of rank order correlation comparing the utilities of the holdouts computed by the part-worths of AHP and CA, and the part-worths as well as attribute importance weights of AHP and CA. The results of this study showed good holdout and attribute-level convergence between the methods. This finding is consistent with the conjectures of Muyle (1998), Malvinas et al. (2005), and Meißner and Decker (2009). Mulye (1988) found that AHP converged better with CA ranking than with CA rating based on predicting the holdout and attribute-level convergence. Malvinas et al. (2005) mentioned that CA produces relatively same ranks as AHP. Meißner et al. (2009) found that the resulting preference structure between AHP and CA prove to be similar on the aggregate level.

The convergent validity of both methods is less than it seems to be especially when considering the attribute importance weights. Further analyses showed that non-convergent results of the two methods with respect to attribute importance weights estimated at the individual level as well as at the aggregate level. This evidence of a lack of convergent validity, which is concordance with findings reported by Helm et al. (2003, 2004a), Scholl et al. (2005); Meißner et al. (2008), Ijzerman et al. (2008), van Til et al. (2008), and Ijzerman et al. (2012), is an important finding because a possible explanation for the mixed results of previous research could be a failure to include a compensatory process in the assumed evaluation model. For example, one explanation for previous studies indicating a lack of convergent validity is that the assumed model was compensatory; however, the actual evaluative process used by respondents may have corresponded to non-compensatory simplifying heuristics where typically only a small number of attributes are taken into consideration (Todd, 2007). This problem is avoided in the AHP approach based on hierarchical structuring where the respondents have to evaluate each relevant attribute even if the attribute is less important to them (Mulye, 1998). This could be considered as one of the advantages of the AHP over the conjoint approach.

It is important to note that the attributes on the AHP and CA were listed in the same order. However, one potential cause of the lack of convergent validity could be order bias if the potential effects of order are different in the two measurement methods. For example, Johnson (1989) indicates that the relative importance of attributes derived from the part-worths is influenced by the order in which the attributes appear in the full profiles. That is, if the same attribute appears first for every respondent, it tends to receive more attention than if it appears later. Other researchers have found evidence of order bias across a variety of measurement methods. Unfortunately the order within profiles in this research was not randomized during the research because as Chrzan (1994) points out, in some product/service category a variation of the order or attributes may cause irritation among respondents. Thus, the amount of influence of the different effects in order is not known in this study. A possible topic for future research is to examine the degree to which order of stimulus presentation affects AHP and CA in different ways and consequently, reduces the convergent validity of the two measurement methods.

As a further important result, the average difference in derived importance weights between AHP and CA is significantly different. Comparisons of differences in attribute importance weights were found in the studies by van Til et al. (2008) and Ijzerman et al. (2008). These authors suggest that differences in the attribute importance weights between the methods can be attributed to the elicitation procedure itself rather than to faults in judgment of the respondents. AHP’s importance weights are more distinctive than those of CA. Thus AHP seems to overcome the well-known problem of “flat” attribute importance weights that have been frequently reported in the literature on CA (see, e.g., Orme, Alpet, and Christensen, 1997).

Another important issue in the present study is the effect of the data collection method. This study compared the internal and the predictive validity of AHP and CA across two data collection modes, e.g., paper-based questionnaire, and web-based survey. Regarding internal validity, both AHP and CA lead to comparable good results and show very high values for both survey modes. For AHP there are no known comparable studies of alternative data collection methods. Thus, the results obtained by AHP cannot be directly compared with previous research. Nonetheless, AHP compared with CA at least seems to be very robust with respect to the consistency between the two data collection modes. As the present study and other empirical studies demonstrated, this can be attributed to the fact that AHP is easier to answer, requires less time, and the questions are clearer, thus it demands less cognitive effort from respondents. Consequently, the respondents may have been more confident and more consistent in the evaluation task. In the case of CA, the data gathered via the offline method leads to a higher internal validity. This finding supports that of Melles, Laumann, and Holling (2000) and contradicts that of Klein et al. (2010). One reason for this finding could be that Interviewer-led administration of the survey may improve the quality of data because the interviewer can recognize that more explanation is needed, can more fully explain the task, and can answer questions (without leading the respondent). Thus, the presence of an interviewer may influence the data positively.

Regarding predictive validity, the two data collection methods can produce different results. Although the predictive validity of AHP and CA is definitely satisfying for both survey methods, the offline method was judged slightly superior to the online method on the basis of predictive validity. In case of CA, this finding contradicts the findings of Klein, Nihalani, and Krishnan (2010) that the validity of the data of the online-conjoint analysis is slightly higher. One possible reason for the discrepancy between this result and those with Klein et al. (2010) is differences in task complexity. These latter authors used nine stimuli described on two attributes with three levels and two attributes with two levels, whereas this study used 16 stimuli described on three attributes with four and one attribute with three levels, which can be classified as being relatively complex. Thus, the ranking task was more complex, increasing the need for simplifying heuristics. Payne (1976) as well as Billings and Marcus (1983) suggest that a relatively large number of options induces a shift from compensatory to non-compensatory processing with the goal of reducing the number of relevant alternatives as quickly as possible. Some authors also report more use of non-compensatory strategies when the number of attributes increases (Biggs, Bedard, Gaber, and Linsmeier, 1985; Sundström, 1987). Such situations again are referred to as information overload (Verins, 1998), thereby reducing the predictive validity of the sample. As Netzer and Srinivasan (2011) point out, reducing respondents' cognitive load is particularly important for web-based environments, where respondents' patience tends to be low (Deutskens et al., 2004). Information overload could otherwise have impaired predictive validity (Green and Srinivasan, 1990; Lines and Denstadli, 2004).

Another reason for the inconsistent results in terms of the validity of different data collection methods is caused by the misunderstanding of the evaluation task (Helm et al., 2004b). According to Helm et al. (2004b), the knowledge of the preference measurement methods influences the validity of the results positively. Contrary to the online survey without the help with an interviewer, in the paper-based survey, the questionnaire was done with the help of an interviewer sitting beside the interviewees and thus they had the chance to ask an interviewer in case of any questions, thereby making them better able to understand the preference measurement methods. In this case, again the interviewer could possibly affect the results in a positive way. This is confirmed by the empirical studies of Tscheulin (1991, 1992) and Helm et al. (2004b) that both AHP and CA get favorable results when only those respondents were considered which understood the evaluation tasks. The results suggest that respondent training and experience can improve the quality of results. So it is important to include some sort of warm-up task (e.g., an initial presort of the cards in the case of CA) prior to a preference measurement method as also recommended by Helm et al. (2004b). This can be achieved by allowing respondents to study how methods work and how selections between choice sets should be conducted. In the case of AHP, it should be noted that the offline method generates more valid results concerning predictive validity, although no significant difference was found regarding internal validity between the offline and online survey. This might be caused by inconsistent response behavior of the respondents, whose answering behavior for earlier questions is not the same as their choice decision (Klein et al., 2010). Such a change in the answering behavior is supported by empirical studies (Riley et al., 1997) and has been confirmed within this study.

Despite the observed incongruence some interesting patterns were observed at the aggregate level. For both AHP and CA, aggregate choice share predictions are very consistent across the data collection modes. Hagerty (1986) suggested one possible explanation for his results: that for individual-level predictions noise in part-worth estimates is quite important, whereas for aggregate-level predictions the noise tends to cancel itself out across respondents. From this perspective, in the present study individual differences seem to balance out in the aggregate and both data collection methods lead to comparable aggregate results.

Summary and Conclusions

Conjoint analysis is widely recognized as a useful marketing research tool which can provide invaluable information for product design, market segmentation, pricing decisions, and completive positioning (Kim et al., 2004; Orme, 2010). It can also measure brand equity, which is an especially critical issue for many managers (Orme, 2010). Conjoint analysis quickly became the most broadly used and powerful survey based technique for measuring and predicting consumer preference. The AHP is a relatively new technique and has been overlooked by most researchers in marketing so far (Mulye, 1998; Meißner and Decker, 2009). There has been some CA research in the field of hospitality marketing. An influential case study was published by Wind et al. (1989) regarding a successful application of conjoint analysis to help Marriot design its new Courtyard hotels. Their sophisticated study contributed important guidelines and examples for using CA in market segmentation, product improvement, and service positioning. Nonetheless, relative to the amount of research in hospitality industry every year CA research is still very spare since it requires a considerable amount of time and money. Cost considerations often make a CA study impractical.

One of the primary limitations of CA is the number of attributes that can be handled. For example, increasing the number of attributes and attribute levels may raise the number of stimuli, respondents are asked to evaluate, which may result in task overload. Because respondents are limited with respect to their willingness and capacity to process information, it is important to take into account the complexity of the stimuli. Again, if information processing capacity and willingness limits are ignored, respondents may engage in undesired task simplification strategies such as ignoring the less important attributes or levels to reduce their information load, making the part-worth less reliable. Indeed, CA studies involving many attributes and attribute levels often occur in practice (Francois and MacLachlan, 1997). Because such studies can cause respondent fatigue and lack of cooperation it is important to design data collection tasks that reduce those problems.

The AHP has been found to offer promising features to overcome some of the shortcomings of CA in measuring consumer preferences regarding hotel branding. As the AHP evaluation task is based on direct paired comparisons of individual attributes and levels it is possible to add more attributes each having a larger number of levels. The AHP approach implements an extremely easy evaluation task for respondents because pair-wise comparisons are easy to make, discuss, justify, and agree on (Dyer and Forman, 1992). AHP should be the preferred method if the complexity of the evaluation task is high or the respondents' motivation is low. The results show the AHP offers great advantage over the CA in terms of reducing interview time and fatigue. AHP requires less time to complete the questionnaires. This is an important finding respondents may become less fatigued by an AHP questionnaire and cognitive efforts are reduced. The AHP promises to be a suitable tool in a situation where respondent willingness to participate in a study, interview time, and respondent fatigue are major considerations. Given this result, future applications of CA for measuring customers' preference structures seem to be at least questionable because of the practical advantages of AHP.

Both methods ask for a certain level of consistency in the respondents' answers. AHP seems to be more robust than CA in the case of ambiguous and fuzzy statements, whereas the latter task is very vulnerable to judgment 'errors', such as inconsistent answers. The robustness of the AHP constitutes a further important result of this empirical analysis.

A detailed analysis reveals that AHP is equivalent to CA with respect to predictive accuracy. This is very surprising news since the majority of today's market research firms use CA to perform market segmentation and market share predictions (Meißner and Decker, 2009). Despite the similar outcomes of the AHP and CA in this study, the models generate totally different results in particular when considering the importance of attributes, e.g. there is almost no convergent validity. Thus, an important result is that although the methods result in different preference structures, their predictive validity is very similar.

Regarding the choice of a data collection method, the study results show that for both AHP and CA data gathered through paper-based surveys with the help of an interviewer lead to a slightly higher internal validity (with AHP being a minor exception: equal internal validity) as well as predictive validity. Therefore, this study confirmed and supported the assumption (e.g. Klein et al., 2010) that the presence of an interviewer influences the data positively, particularly when the complexity of the task is high. This implies that it seems to be advantageous to explain some relevant methodical aspects of AHP and CA respectively to the respondents prior to the formal judgment task. It can be further assumed that especially in the case of a non-personal interview setting (e.g., an online survey without the help of an interviewer), some type of warm-up task (e.g., a warm-up card sort or attribute importance task) prior to starting the evaluation should have a positive effect as also suggested by Helm et al. (2004b).

In summary, since the main objectives of AHP and CA focus not on prescribing how a consumer should make a decision but on predicting their choice by analyzing the trend in their preferences (Mulye, 1998), both methods seem to be applicable when targeting at predicting choice as well as solving multiattribute design problems. Rather, AHP seems to be a promising alternative to CA in many cases where the task is usually boring, complex, and frustrating for even the most highly motivated respondents. The resulting cost of running the AHP is advantageous. The two methods did produce somewhat different patterns of attribute importance. Whether the differences in attribute values observed here are meaningful enough to warrant the additional time and effort remains unclear and will require additional research. The choice of a data collection method for both AHP and CA is still a practical one. There are no big differences in the validity between offline and online data collection methods, e.g., altogether the data quality of both surveys is very satisfactory. Rather it can be shown that the validity of the data collected via the offline method is only slightly higher.

Managerial Implications

Consumer brand preferences are influenced by a strong brand equity. Hotel brand management can be successful if management can understand and manage brand equity correctly. This success can produce strong attributes that can influence how consumers make their brand preferential choices. This study focused on the importance of these brand equity on consumer brand preferences. Thus, the current study examined the brand equity dimensions (attributes) and their sub-dimensions (attribute levels) that customers perceived to be important in hotel branding using the AHP and the CA.

The findings from the AHP and the CA in this study will allow hotel management to identify and prioritize hotel brand equity components requiring attention and to act to ensure improvements in overall brand equity. When developing marketing and brand strategies, hotel management should be aware of the importance customers place on each of the primary components and on each of the sub-components. This approach can help brand managers allocate their resources based on the relative importance of the hotel brand equity attributes of their particular target segments. However, the different priorities of attribute importance weights (attribute preferences) obtained from the two methods could lead to very different allocations of marketing effort. The AHP results could lead to substantial effort on improving perceived quality and brand loyalty and little attention to brand awareness and brand image. Inversely, the CA results could lead to emphasis on brand awareness and brand image, but de-emphasis on perceived quality and brand loyalty. Since the AHP and the CA have equal predictive validity but yielded totally different results, a natural follow-up question is: Which method should be used?

In the current application to the hotel branding setting the AHP might generate more valid results than the CA with respect to the importance of attributes. Because the AHP results were nearly identical with previous study results by awarding high importance to perceived quality, the initial verdict based on the previous study results would favor the AHP. A second conclusion was based on external validity, that is, the example of InterContinental Hotels Group (IHG)'s Holiday Inn brand family, comprising Holiday Inn, Holiday Inn Express, Holiday Inn Club Vacations and Holiday Inn Resort. The company conducted a survey that included 18,000 consumers to find out what guests were looking for in their hotel stay, and used those findings in order to address the changes that needed to be made to the brand, and then began a re-branding process that completed in 2011. Top priorities included décor, service and overall quality changes. To meet the required service and quality levels the company created a new service promise/culture called "Stay Real." The “Stay Real” imitative trains the Holiday Inn staff to make sure they provide a good service to their customers. In other words, all employees had to be retrained in being personable and responding to guests' issues- to further ensure staff develops the behaviors and skills to best serve guests by treating them as real people and consistently delivering the real, genuine service for which Holiday Inn is known (InterContinental Hotels Group, 2007). The company suggested that improving perceived quality is one of their top priorities since a differentiated lodging experience cannot be delivered through imagery and product alone (InterContinental Hotels Group, 2007). The Holiday Inn case reaffirms the importance of perceived quality and adds support to the argument that perceived quality is central to building hotel brand equity. In this perspective, the AHP hotel branding results might seem closer to marketplace reality.

Consequently, as Bailey and Ball (2006) also stated, developing positive perceptions of quality are vital parts of hotel brand management. Based on the AHP results, brand managers may allocate their marketing resources to properly handle each of the dimensions of brand equity in order to build a strong brand or brand equity and, as such, the focus should be on (1) enhancing perceived quality, (2) strengthening brand loyalty, (3) improving brand image and (4) creating brand awareness.

Enhancing Perceived Quality

Since perceived quality was identified as having the strongest influence on hotel brand equity, it is evident that hotel practitioners need to pay special attention to the perceived quality component of hotel brand equity. These results imply that the hotel may increase its competitive advantages over others by substantially improving the service performance in order of importance with respect to (1) reliable, (2) courteous, (3) prompt, and (4) error-free service.

As the perceived quality received is based on customers' expectations, hotel firms are required to provide quality services to meet customers' expectations (Yoo et al., 2000). As the study results indicated, of the four sub-components of hotel service quality, reliability was the most important factor of perceived service quality, and in the hospitality industry, this factor refers to sincerely solving guests' problems and complaints in a fair manner (Markovic and Raspor, 2010). This finding was consistent with the previous findings by Berry, Parasuraman, and Zeithaml (1994) and Markovic and Raspor (2010) research conducted in hotel settings. The research suggested that reliability was the core of quality service. Some of the biggest issues with customers trusting a brand are related to the way in which the hotel employees deal with problems and complaints. Problem-solving and handling complaints can negatively influence the customer's hotel experience, thus resulting in decreased trust in the brand and negative brand associations (Mackay et al., 2013). When guests have specific problems with a hotel, such as an unclean room, poor temperature control, inaccuracy of hotel reservations, stay inconveniences, or billing errors, their problems and complaints should be resolved in a proper, timely and professional manner. Otherwise, customers lose confidence in the hotel's ability to do what it promises dependably and accurately. In fact, Simmerman (1992) stressed that proper complaint handling would retain or even build customer loyalty since such handling can reflect the reliability of hotel services. Friendliness from the staff and sincere apologies do not compensate for unreliable service (Berry et al., 1994). Although most customers appreciate an apology, the apology does not erase the memory of that service (Berry et al., 1994). If a pattern of service failure develops, customers conclude the hotel firm cannot be counted on, friendly and apologetic or not (Berry et al., 1994).

Since reliability (handling complaints of customers) appears to be a key driver for evaluating a service in this study, particular attention should be paid to the training of service providers' problem solving skills. A training program should focus on teaching how to properly treat upset and frustrated customers and how to quickly react to various service failure situations. Some hotel training programs use videotaped scenarios of service failures to show employees potential problems and the appropriate solutions (Reid and Bojanic, 2009). More importantly, hotel employees should be trained to know their guests' needs before they ask, so that the hotel firms can anticipate and avoid possible service failures. Of course, hotel companies cannot prevent all customer complaints, and thus it is critical for a hotel organization to possess a well-managed, good recovery and complaint system in order to provide quality service (Johnston, 2004). Hotel managers should develop service recovery strategies or plans such as specific monetary compensation guidelines and need to train their employees to use them and to offer proper level of compensations (such as price discounts, free meals, free one night stay, refunds, coupons, or room upgrades) depending on different levels of complaints and situations of the customers (Kim et al., 2009). According to Reid and Bojanic (2009), the Ritz-Carlton allows its employees to spend up to $1,000 to take care of dissatisfied customers.

The participants further indicated that a courteous staff was the second important factor in evaluating quality services during their customers' hotel stay. In this study, courtesy involved politeness and friendliness of the hotel staff and other contact personnel, as well as, well-dressed, clean, neat and professional staff appearance. Since the hotel service requires frequent interaction with customers, courteousness of the contact personnel is an important service quality indicator. Since hotel customer perceptions are mainly influenced by the employee's behavior and attitudes, employee selection and training is one of the critical issues for offering good service and building a strong brand (Sun and Ghiselli, 2010). In particular, front desk agents have an important impact on consumers' experiences of services and their performance affects customers' perceptions and they are the ones who most influence customers' evaluation of a hotel's service (Li and Krit, 2012). It is important to select those employees carefully and train them well.

In the following, it is important for hotels to respond to customers' requests, be error-free (the knowledge of staff to answer guests, as well as their ability to convey trust and confidence) and in a prompt manner (service without delay). Since providing error-free or accurate information can enhance brand trust (Khan and Tabassum, 2010), hotel employees must be competent and must possess the required skills and knowledge. Hotel employees need to pay special attention the hotel's facilities (e.g. business center, swimming pools, gift shops, banquet rooms, conference room, and any recreational facilities) making sure that if a guest asks about the hotel's products or services, they can introduce them accurately. In particular, front desk agents should be familiar with the hotel's information booklet and tour guide information (e.g. local tourist spots and historic sites), menus placed in rooms or restaurants, and availability of complimentary items. Hotel managers, in turn, have a crucial role in ensuring that knowledge and training are in place to enable front desk staff to deliver the services brand values to all the hotel's customers. Furthermore, hotel staff should react and deal with guests' requests promptly in order to satisfy their different needs (Li and Krit, 2012). Hotel firms should guarantee prompt services by assigning more employees during rush hour periods. For example, part-time employees can be used on an on-call basis during an unusually busy day.

In order to monitor service experience in the eyes of the customer, hotel managers may need to consider the use of mystery shoppers with the purpose of identifying elements of service experience that require management attention for improvement. So and King (2010) also suggest that hotel firms may consider introducing employee programs to allow service providers to experience the hotel services as a guest and to see how their work behavior influences customer experience. Customer comment cards, including questions about the knowledge and confidence of the staff, courtesy and friendliness of the staff, the timelines of the service, and the neat staff appearance, can be also used to encourage customers to discuss problems that they had with the service.

In order to gain a strong completive advantage, it is suggested that a hotel firm differentiate its service delivery process through its people with respect to reliability (solving guests' problems and complaints), courtesy and friendliness/neat staff appearance (empathy/tangibles), promptness (responsiveness) and error-free (assurance) and should select its employees carefully and train them well. More attention should be paid to the training programs and education for them to cultivate their abilities to solve customers' problems efficiently and effectively. All employees need to be courteous, friendly, and respectful and maintain a neat, clean and well-groomed appearance. They must service customers with consistency and accuracy, and make an effort to understand their customers and respond quickly to customer requests and problems.

Not only personal differentiation but also service differentiation can play an integral role in building a strong brand. In fact, hotel companies provide many services but most of these become routine and are indistinguishable from competitors. A hotel company must differentiate its service offerings from those of competitors (Kotler et al., 2006). For example, Sheraton, Shangri la, and other hotels provide an in-room check-in service. Seoul's Hotel Shilla offers 24-hour room service to international business travelers. Marriott is setting up hotel rooms for high-tech travelers who need accommodations that will support computers, fax machines, and email. Perceived quality can be enhanced and supported by creating memorable and enjoyable customer experiences. According to Kotler et al. (2006), the MGM hotel and Casino in Las Vegas is trying to create memorable and differentiating customer experiences by using wake-up calls with recorded voices of celebrities who have performed there. This provides a differentiating extra to an otherwise routine service. These examples illustrate how hotel companies can differentiate themselves on service. Benchmarking is another way of enhancing perceived service quality by identifying and correcting important weaknesses compared to the competitors, ultimately to enhance hotel brand equity.

In today's completive hotel industry, the hotel's survival depends greatly on its ability to provide superior service which generates a strong brand. All aspects of service quality, including service efficiency, understandability, helpfulness, politeness and friendliness, and appearance should be maintained and consistently reviewed to see whether any improvements are required (Chu and Choi, 2000). Hotel practitioners should devote more resources to staff training. Furthermore, hotel practitioners should ensure that all employees are required to become involved in setting quality standards, and all employees should realize that maintaining service quality is part of their jobs (LeBlanc and Nguyen, 1996). Hotel managers should monitor perceived service quality of the hotel on a regular basis to enhance employee’s service skills, and increase the overall level of service and through training make the standards perfect (Li and Krit, 2012). These efforts can contribute to managing, building, and ultimately strengthening hotel brand equity.

Strengthening Brand Loyalty

The second most important dimension for strengthening hotel brand equity is brand loyalty. Previous studies (e.g. Atilgan, Aksoy and Akinci, 2005) found brand loyalty to have the strongest and the most influential effect on brand equity. For example, in Turkey's beverage industry, only brand loyalty was relevant to building overall brand equity and the results from this study provided further evidence of this in the hotel industry. Brand loyalty has several important strategic benefits to the hospitality firms, such as gaining high market share and new customers, supporting brand extensions, reducing marketing costs, and strengthening brand to the competitive threats (Atilgan, Aksoy and Akinci, 2005). Managers should concentrate their efforts primarily on brand loyalty which, if increased, will contribute positively to their hotel firm’s brand equity.

The results of this study indicated that participants rated past experience (such as satisfaction) as being the most important factor to their overall hotel experience, which suggest customer experience has a major impact on brand loyalty. This finding coincides with a recent survey conducted by the Deloitte study (2013) which found that customer experience had more of an impact on their decision to revisit than hotel loyalty programs to encourage and reward frequent guests. Only 19% of those who responded to the survey said that a loyalty program was very important when choosing a hotel brand. These results imply that hotel loyalty programs were not a priority to build customer loyalty, although issuing hotel reward cards has been a major tool for building loyalty in the hotel industry for more than a decade and also provide valuable information for the marketing department of hotels. Rather, improving customer satisfaction was a top priority of all mangers working in the hotel industry and requisite for loyalty. This result was consistent with others (e.g., Reichheld and Aspinall, 1993) who have indicated that 90 percent of customers who change service providers were satisfied with their previous service provider. In other words, depending on hotel reward programs such as, free rooms or upgrades, coupons or any price discounts and special online rates to build customer loyalty may have worked in the past, but today's more informed and sophisticated customers are more influenced by their experiences such as satisfaction with the service and products (or amenities) that a hotel provides. To build a brand loyalty, hotel firms must meet or exceed customer expectations on a consistent basis in order to satisfy them (Reid and Bojanic, 2009). Customers who are satisfied with their hotel stay are more likely to become repeat customers, and to spread favorable word-of-mouth publicity (Fornell, 1992). Hotel practitioners should pay special attention to and make every effort to improve customer satisfaction in order to reinforce their behavior and increase the possibility of a return visit. This means that hotels should focus on attracting and retaining loyal guests by improving guest experiences rather than simply issuing hotel reward cards to encourage guests to return. For guests to return and to differentiate among hotels, a hotel should establish consistent quality and also add amenities that add value to a hotel stay. For example, customers may respond better to comfortable mattresses and extra fancy linens, variety and quality of sports and recreational facilities or room's décor than they will to a free basic breakfast buffet.

Since brand loyalty is built through experience such as customer satisfaction, the more positive hotel experiences the consumer has with the brand, the more loyal he or she is likely to become. As a result, investments in satisfaction programs through guest feedback surveys and in the design of relationship communication strategies that aid in creating and informing consumers about the responsive attitudes and behaviors of the brands are ways of building brand loyalty. For example, many hotel firms utilize customer feedback cards or customer satisfaction surveys to obtain information on customer perceptions of the quality of hotel, leading to invaluable information on how to improve existing strategies. Furthermore, hotel managers can use communication methods such as social media to keep their brand relationship with customers so that they can obtain feedback. Consequently, it is more important for hotel companies to manage the hotel customer experience, not only by listening to customer feedback from all communication methods, but also by analyzing conversations to extract valuable business insights and using those insights to improve the customer experience with their hotels. By improving the guest experience, hotel firms should keep in touch with customers and provide satisfaction and reinforcement to current and existing customers (Tepeci, 1999). More importantly, since guests feel a certain degree of brand loyalty especially after their first visit, treating them like loyal guests from the very beginning can be beneficial (Hochgraefe et al., 2012). If a first-time guest is rewarded with a pleasant experience, the foundation for repeat patronage can be successfully established (Reid and Bojanic, 2009).

This study's results suggest that to build customer loyalty top management should improve hotel mangers' understanding of customer satisfaction which is built over time from experience with service and product. This understanding of customer satisfaction and acting on what customers' desire and need can contribute to repeat business and positive word-of-mouth advertisements (Tepeci, 1999), eventually leading to increased sales and profits. All of this combined help nurture hotel brand equity. For hotels to survive and prosper in today's highly competitive hotel industry with numerous brands, repeat business is critical. However, hotel practitioners should bear in mind that building brand loyalty could be a difficult job because even completely satisfied customer may not lead to repeat business for a variety of reasons (Bowie and Buttle, 2004; Ukpebor and Ipogah, 2008). One possible reason could be result of the fact that consumers give more attention to other factors such as the lowest price available or best rewards offered at that time, or simply liking to stay at different hotels when they are making their hotel choices (Bowie and Buttle, 2004). It may be a wise strategy for hotel marketers to identify those who are heavy users of hotel accommodation, who typically account for a large amount of total sales volume, and establish a long term relationship with the those patrons who are likely to become loyal customers by continually fulfilling their needs and wants via relationship marketing (retention marketing).

Improving Brand Image

The study results showed that brand image was the third core component of hotel brand equity although the importance of brand image was smaller than perceived quality and brand loyalty. The implications this has on hotel managers are that they must maintain or strengthen their effort upon the brand image attributes such as (1) good value for money, (2) cleanliness, (3) feel like home (comfort), and (4) elegant atmosphere, depending on their segmentation targets.

The results of this study imply that price can be used as one of the most important differentiators for enhancing hotel image to position a brand as a good value for the money. This findings support Zhou and Jiang's (2011) argument that perceived value (such as value for money) is one of the critical success factors for hotels since in hotels customers perceive a relationship between price and quality in terms of brand images. In the hotel industry, the link between price and quality in different product classes is strong (Bowie and Buttle, 2004). Consumers looking to be pampered in a luxurious environment expect to pay higher prices including prestige prices. In this case, hotels seeking to position themselves as luxurious and elegant will enter the market with a high price, high quality, and exclusive image that will support this position (Bowie and Buttle, 2004). On the other hand, consumers looking for basic products expect to pay lower prices; in this case, the positioning focus might be a standard quality at a lower price, implying better value for money in the economy product class (Bowie and Buttle, 2004). However, it is important to note that the perceived image of a hotel company may not necessarily reflect the true or real state of the product or service offered (Bowie and Buttle, 2004).

Keller (1993, p. 3) describes brand image as "perceptions about a brand as reflected by the brand associations held in consumer memory." This perception influencing image may differ from the actual attributes. This could be a result of the customer's individual encounter and possibly the impact of differences in perception, misrepresentation or recognition (Ukpebor and Ipogah, 2008). For example, mid-market hotels with falling standards but still maintaining a medium pricing strategy do not represent good value for money, or old established hotels which are no longer as luxurious as they used to be, and whose facilities no longer match the price charged, theses hotels will eventually either have to reinvest in their facilities or reduce their prices (Bowie and Buttle, 2004); otherwise, it will damage hotels' brand image related to perceived value. As customers recognize the poor value for money, the image of the business will rightly suffer. Hotel marketers should keep in mind that prices must accurately reflect the property's desired position, be consistent across the range of products offered, and match the target market's expectation of quality and value (Bowie and Buttle, 2004). The hotel company should adopt a more appropriately balanced strategy. In other words, hotel revenue managers have to be meticulous about setting prices and take into consideration perceived value in their hotel survey (Zhou and Jiang, 2011) so as to create and hold a positive image and perception in the mind of customers. The difference between adjacent product classes (e.g. a four-star hotel and a three star hotel) can be virtually indistinguishable (Bowie and Buttle, 2004). This can lead to customer confusion, as the relative value for money between competing offers is not transparent (Bowie and Buttle, 2004). More research and analysis should determine what consumers can afford to pay and what they are willing to pay.

Second, it is not surprising to find that the effect of atmosphere, which is most significantly affected by cleanliness, followed by feel like home (comfort) and finally elegance of the hotel, on the guest's perceived image is significant. This is due to how critical atmosphere is to image building. Since cleanliness and comfort are paramount to creating a positive brand image in the hotel industry, hotel practitioners should concentrate on these aspects from their customers' point of views. Given that those image attributes are abstract, it is important to know how hotel marketers can position their intangible products positively in the minds of customers. One way of doing this is to tangibilize the intangible, in other words, to provide tangible evidence that reinforces the position the hotel company is aiming to attain (Bowie and Buttle, 2004). In particular, cleanliness of the room and lobby is the first thing guests see. Since the first impression is crucial, hoteliers must acknowledge the concern of the customers for cleanliness and allocate resources adequately to the housekeeping area. It is highly suggested that hotel firms should employ a mystery auditor to evaluate how closely cleaning standards are being followed to ensure that the unit is maintaining a continuously high level of cleanliness and hygiene. In this way, hotel cleanliness will prevent the likelihood of a shock to the hotel’s image (such as a boycott, negative word-of-mouth or bad reviews).

The study results also suggest that hotels should strive to make their hotel room feel comfortable and familiar to home, not a temporary place to reside. This sends a meaningful message to hoteliers, in that resources should be directed to improving and maintaining the quality of rooms, including room design/décor, quality and comfort of a bed/bed linen/pillows, quality and sufficiency of fixtures, and lighting and temperature control. For instance, in order to reflect a clean, comfortable image and a contemporary look and feel, Holiday Inn redesigned the building interiors (e.g. refreshed guest room, including comfortable beds and efficient bathrooms as well as upgraded lobby areas) and exteriors (e.g. exterior hotel lighting and landscaping). The friendliness of the staff can also make the guest feel at home. A good staff image could help the hotel to shape the brand image since the hotel staff has direct interaction with customers. In order for guests to perceive a hotel as a home, full of warmth, comfort and human kindness, hotel staff should cultivate an environment in which guests are treated as family, not just as the objects of service (Li and Krit, 2012). Simple procedures like using the name of the guests and taking note of preferences such as a grocery list are important. Many hotels have tried to do their part to make rooms seem more like home to differentiate themselves from competitors. For example, the Residence Inn, Marriott's extended-stay brand, has a complimentary grocery-shopping service to make the guests stay more convenient and feel like home.

Even when competing offers look the same, hotel customers may perceive a difference based on company or brand image (Kotler et al., 2006) and hotel companies need to work to establish images that differentiate them from competitors. According to Kotler et al. (2006), atmosphere is appreciated through senses such as vision, hearing, smell, taste and touch, so one of the strategies that can be used by the hotel marketers is using sensory marketing (branding), as many hotel firms use custom-made “signature scents” in their lobbies to create unique associations/image. For example, Omni Hotels infuses its lobbies with a lemongrass and green tea scent with a feeling of cleanliness, relaxing guests and creating an elegant environment. Starwood's Westin brand uses a white tea fragrance in all its lobbies worldwide to express the lifestyle feeling of the brand, emphasizing to the image of the brand. Other international hotel chains such as the

Sheraton, Sofitel, InterContinental and Mandarin Oriental are all using carefully selected fragrance as well as customized music to create a luxurious feel and perfect first impression with their guests. These are good examples of how the hotel firms' sensory marketing, such as customized music and scent selections, can be used to engage guests in a complete sensory experience and to establish a distinctive and unique image, and ultimately to differentiate itself in the marketplace.

Brand image can be strengthened by advertising and promotional activities. Advertising and promotions should be used to create and reinforce an image for the target public (Reid and Bojanic, 2009). To establish and maintain a specific image in the mind of the consumer, Hampton Inns advertising campaign emphasizes high-quality accommodations, friendly and efficient service, and clean, comfortable surroundings. Many hotels in the Hyatt's product class were working on differentiating their image, specifically through attractive physical features such as spacious atrium-style lobbies (Bowie and Buttle, 2004). This product differentiator was consistently used in Hyatt's promotional campaigns, and succeeded in positioning Hyatt, in the minds of consumers as a more exotic, grand, majestic and distinctive image than its competitors (Bowie and Buttle, 2004). Hotel companies may need to change their image through their logos or symbols without changing their name (Kotler et al., 2006). Advertising and promotional campaigns should also carry these logos or symbols effectively to provide a differentiated and familiar image to a hotel (Kim and Kim, 2004). For instance, all Holiday Inn hotels changed its logo and brand signage to enhance the customer's impression of the hotel's brand image and match the company's leading edge image- to further create a new, clean and more contemporary brand image. The new logo, green ‘H,’ was created to fit the chain’s consumer image, evolving the iconic script logo, energizing the signature color green and eliminating the current shield shape.

To create a successful and differentiated brand image and in turn, brand equity, hotel marketers should communicate, position and differentiate their brand on all four elements of brand image in this study identified. To enhance the customer’s perception of the hotel’s brand image and to build a lasting, positive image, a hotel company should convey a singular or distinctive message that communicates the product’s major benefits and positioning using advertising and promotional activities. Also, it is important to continuously use advertising and promotions through various communication channels in a strategic way to improve the hotel’s brand image. The multiple communication channels are especially important as brand image involves brand association.

Consequently, hotels should engage in more communication with customers in order to enhance the customer’s impression of the hotel’s brand image and should build up long-term relationships with customers, focus more on the differentiation strategy of a hotel’s image, to distinguish it from other hotels.

Creating Brand Awareness

Among the four dimensions that were identified in this study, brand awareness seems to be the least influential dimension when it comes to consumer’s perception of brand in the hotel industry, as such, it appears to be a noncritical dimension of hotel brand equity. Nonetheless, brand awareness plays a crucial role in marketing communication efforts of a hotel firm for all promotion strategies because the brand is not likely to be considered or purchased, unless they are aware of it (Peter and Olson, 2001). In other words, the higher level of awareness of a brand, the more likelihood there is of this brand being considered when customers purchase (Hoyer, 1990; Nedungadi, 1990). The level of brand equity is determined by the level of brand awareness which plays an important role in brand equity.

The study findings indicated that to create brand awareness, brand familiarity was the most important factor for selecting a hotel brand, which could be supported by Park and Lessing's (1981) argument where brand familiarity has been as an important factor in consumer decision-making. This factor was followed by top-of-mind awareness, brand popularity and finally advertising. These results imply that customers rarely select a hotel brand as a simple reaction to the stimulus of advertising; they select the brand because it is familiar. In other words, customers are much more likely to think of or recall hotel brands that they have used before. For this reason, popular hotel brands with higher market shares have a distinctive advantage (Peter and Olson, 2001). Because they are used by more consumers, these brands are more likely to be activated in evoked sets and included in more consumers' consideration sets (Hoyer and Brown, 1990). In contrast, unfamiliar and low market shared hotel brands are at a disadvantage because they are much less likely to be included in consumers' evoked sets and thereby be considered as choice alternatives (Peter and Olson, 2001). As the study results also indicate, hotel brands with heavy top-of-mind awareness are more likely to be included in the evoked set of choice alternatives that come to mind during decision making processes. To be successful in the hotel business, a brand must include the consideration sets of at least some consumers because the brand is not likely to be considered or purchased, unless consumers are able to recall the brand name (Peter and Olson, 2001).

Hotel marketers should develop strategies to increase the likelihood that a brand will be activated from consumers' memories and included in their evoked sets of choice alternatives. One marketing strategy to increase the probability of inclusion and being chosen from the evoked set is the repetitive and expensive advertising campaigns through various communication channels, such as newspapers, magazines, tourist board publications, broadcast media (radio, cinema and television), banners or pop-up on search engines and websites, SMS (text messaging), social media, email promotions and celebrity/star endorsements. Advertising repetition and the heavy expenditures create brand familiarity, promote customers to make quick purchase decisions (due to heavy top-of mind awareness), and enhance brand popularity.

A hotel company's distribution strategy plays a critical role in increasing brand awareness because it can influence whether a hotel brand is in consumers' consideration sets and it keeps customers thinking about the brand. Prominent brand-name signs (buses, taxis, trains, gas stations, and billboards) remind consumers of the brand name and maintain brand awareness, thereby enhancing the likelihood that consumers will encounter the brand at the time of the decision, which increases its chances of entering consumers' consideration sets (Peter and Olson, 2001). Kim and Kim (2005) suggest that brand awareness can be improved through charity involvement and sponsored activities (such as sports, arts, cultural activities, or other kinds of public events).

To create brand awareness, hotel marketing managers need to ensure that their customers are so familiar with the brand that they will be able to immediately recognize and/or recall it. Hotel marketing managers should have promotional communication strategies in place to constantly remind customers of their brand, thereby maintaining the brand's presence in the customers' minds. In fact, intense and successive promotional activities exist within the hotel industry, such as advertising. Increasing brand awareness through increasing investment or developing promotional channels is essential when hotels attempt to differentiate themselves from competitors.

In summary, the measurement and management of brand equity have become top priority marketing issues in recent years, as evidence by the growing literature on the subject. Several empirical studies in the literature supported the positive relationship between consumer-based brand equity constructs and brand preference and purchase intention (Cobb-Walgren et al., 1995; Agarwal and Rao, 1996; Vakratsas and Ambler, 1999; Mackay, 2001; Prasad and Dev, 2000; Myers, 2003; de Chernaony et al., 2004; Chen and Chang, 2008; Chen and Liu, 2009; Mishra and Datta, 2011; Moradi and Zarei, 2011; Tolba and Hassan, 2009). Furthermore, Forgacs (2003) suggests that branded hotels outperform non-branded properties on performance indicators such as average price, level of occupancy, revenue per available room, revenue per available customer and return on investment. Such linkages have been empirically validated by Kim, Kim and An (2003) and Kim and Kim (2005), who establish a positive relationship between brand equity and financial performance in the hotel industry. Thus, it is imperative to know how much equity a brand commands in the market as building strong brand equity is a very successful strategy for differentiating a product/service from its competitors (Aaker, 1991). In other words, hotel practitioners need to understand how brand equity, as an indicator of brand strategy success, can be measured, and how it can be built (So and King, 2009).

The results of this study show that different hotel brand equity dimensions contribute to overall equity in different ways, and that an order exists among the four dimensions. Since marketing and brand managers often have limited resources (such as money and time) to implement branding strategies, the findings can help them prioritize and allocate resources across the dimensions. According to the results of this study, two managerial implications for hotel practitioners can be derived here. First, managers should concentrate their efforts primarily on perceived quality and brand loyalty, which have high importance in the dimension of brand equity. In the highly competitive hotel industry, brand managers are advised to develop strategies that provide superior quality to the customers that will positively contribute to their firm's brand equity and then make every effort to improve and keep their loyalty and gain their repeat business. Although the size of impact is smaller than perceived quality and brand loyalty, a strong brand image is a core requirement for brand owners because the brand name distinguishes a product/service from the competitors' products/services (Kayaman and Arasli, 2007). In other words, although the brand awareness in this study is not a critical factor of hotel brand equity, it serves as a foundation for brand image since brand image is memories relating to a hotel brand (Aaker, 1991). As a result, the second implication is that when concentrating on creating perceived quality and brand loyalty, marketing and brand managers should not undervalue the effects of brand image and brand awareness to hotel brand equity and balance their efforts to improve brand awareness and brand image.

Although brand equity cannot be built in short term, it can be built in long term through carefully designed marketing activities (Yew Leh and Lee, 2011). There should be a continuous effort by hotel marketers to enhance customer-based hotel brand equity. Otherwise, old familiar brands die and even well established brands can wear out over time, as a result of poor management of brand, overextension and lack of investment in developing brand equity and values (Kim and Kim, 2004; Ukpebor and Ipogah, 2008). There has recently been a proliferation of new hotel brands, as the products and facilities are relatively homogenous. The plethora of hotel brands and products has led to consumer confusion over the lack of perceived differences between competing brands. This would imply that hotel brand managers should bear in mind that how hotels position themselves and differentiate themselves from competitors is critical to their success when trying to build a strong brand. It is highly suggested that developing more effective brand differentiation strategies is the most important task of hotel brand management, due to the lack of differentiation and customer confusion stemming from the recent brand explosion. Therefore, differentiation among hotels should be receiving more attention.

Methodological Implications

Since AHP and CA are completely different approaches but get similar results, the results found in this study have experimental design implications. If one chooses to use either AHP or CA, the results provide some suggestions on how its performance can be improved.

First, a theoretical drawback of the use of CA is that the number of stimuli that is required for part-worth estimations on an individual level increases dramatically with an increasing number of attributes. The number of stimuli necessary for reliable value estimates can be daunting for the respondent, especially if many stimuli or alternatives are presented simultaneously. In that case a practical alternative could be the use of pair-wise comparisons of AHP.

Second, practitioners are assured of getting reliable results from the use of fractional factorial designs that help reduce the number of profiles in a study (Muyle, 1998). However, even with fractional factorial designs, full-profile data collection procedures often involve comparing a number of stimuli with multiple attributes and levels. To reduce the task difficulty in conjoint studies and respondents' fatigue in profile evaluations an alternative to the full-profile approach is based on Clatworthy (1973)'s balanced incomplete block designs (BIBD). For example, in this study a balanced incomplete block design can be employed to assign 20 profiles to four different blocks and each profile can be then used five times in random location of different cards in the survey using BIB test. The BIBD is an application that is widely used in the design of statistical experiments. Since the limit of ranking as a human being is "within five fingers of one hand" (Won and Jung, 2001), by extracting less than five among more than 10 variables, respondents are able to rank items easily and the exercise can be repeated to complete the design. When calculated by a special computer program, the effect is the same as ranking 10 (Won and Jung, 2001). In this regard, the BIBD can be an effective method in dealing with the problem of information overload, thereby reducing respondents' burden and making it possible to increase the number of attributes studied. Even in the case of AHP, large numbers of required paired comparisons are almost unavoidable. In such cases, it is advisable to limit the number of paired comparisons to a manageable size by applying the method of Incomplete Pair-wise Comparison (ICP) developed by Harker (1987). The AHP allows for missing pairs thus reducing the number of needed paired comparisons and reducing the respondents' burden.

Third, in order to improve the quality of data, one may even consider eliminating the inconsistent respondents based on multiple criteria to test the reliability and validity of the CA. The validity of the conjoint model has to be evaluated by checking the value of Pearson's as well as Kendall's tau (Orme et al., 1997, Sorenson and Bogue, 2005). Pearson's r and Kendall's tau values are recommended to be (>.80) and (>.70), respectively. This helps to improve the quality of the input data and thereby the validity of the part-worths (Daiber and Hemsing, 2005). In the case of AHP it may be sometimes necessary to remove the most inconsistent respondents so as to improve judgment accuracy based on consistency ratio (>.20) as a guide in their comparisons .

The study results also offer useful perspectives to consider when choosing between offline and online data collection methods. Although the data gathered online leads to a slightly lower internal validity (not the case for AHP) and predictive validity, it cannot be said that the difference in the validity is based solely on the different data collection methods, since it could be possible that the presence of the interviewer is the cause for the difference in validity (Klein et al., 2010). The presence of an interviewer might have a positive impact on the validity of the gathered data because it facilitates a higher level of control and help. Problems can also arise in cases where interviewer biases have to be expected. Nevertheless, the internal validity and predictive validity seems to be sufficient enough even in the case of its online form. This implies that the lack of an interviewer does not necessarily result in a lower validity. In general, the advantages of online surveys are widely discussed in the literature including their relatively low cost, ease of administration, and geographical flexibility, while holding true for an online version of AHP as well as an online CA. Pictures and diagrams can also be included in the questionnaire. In fact, the choice of a special data collection method should not depend on statistical criteria but on the purpose of the investigation as advised by Klein et al. (2010). Once the researcher has determined the appropriate vehicle for interviewing a given population further recommendations can be given.

First, if only a small sample size is needed to predict individual choices (e.g. small boutique hotels where the number of respondents small) and the target market segment for a new hotel product is very small, a survey with the help of an interviewer might be conducted. However, if managers usually need to predict market shares (where respondents will not purchase a specific alternative such accommodation demands for all markets served by a large hotel chain) and conduct market simulations to guide managerial decisions, there is no reason of not collecting the data online, since a bigger sample size can be obtained easier and faster by offering flexibility in the data collection process and online surveys are usually cheaper than paper surveys (Klein et al., 2010).

Second, the practicality of the alternative data collection methods would differ as a function of the size of the stimulus design (e.g., the number of attributes and their levels). Again, reducing respondents' cognitive load is particularly important for Web-based environments, for which respondents' patience tends to be low (Deutskens et al., 2004). In this regard, online data collection methods may appear better suited to smaller designs than lager designs. There are no suggestions concerning the amount of stimuli for an online conjoint analysis (Lines and Denstadli, 2004). However, it is evident that the complexity and the challenge increase with the number of stimuli (Klein et al., 2010) as in the cases of using conjoint analysis in traditional paper-based surveys. Although the notion of what qualifies as smaller design is unclear (e.g., Melles et al., 2000; Akaah, 1991; Green, 1984), this study of hotel branding suggests that a full-profile conjoint can be successfully implemented via the Internet for a small study including four attributes (or up to 16 stimuli), each attribute being conceptualized at three or four levels. However, in the case of AHP more attributes and attribute levels are possible because it may have an advantage since less complex paired comparisons are required than when using the common full-profile approach in CA.

Finally, if the present study is any indicator, the predictability of AHP and CA seems to be enhanced if they involve the combined use of a paper-based questionnaire and an online survey. According to Cerro (1988) and Stahl (1988), the combined use of different data collection methods may enhance study participation as well as model validity. However, the review article by Sethuraman, Kerin, and Cron (2005) cautions against merging responses obtained from online and offline data collection methods, because in their study the two data collection methods yielded different attribute preferences in a conjoint analysis task. Of course, to provide for better understanding of the conditions that favor the use of combination methods, further research is warranted.

Practical Implications

This study empirically compared AHP with CA in the domain of customer-based hotel brand equity. Many studies have suggested that brand equity should be an important research domain because of its strong association with marketing strategy and hospitality firms' sustainable competitive advantage (Keller, 2003; Pappu et al., 2005; Tasci et al., 2007). In general, brand equity has been accepted as the primary source of capital for many hotel industries (Bendixen et al., 2004). Brand equity has primarily focused on exploring customer-based brand equity and is widely acknowledged as an indicator of measuring the effectiveness of branding strategies. Nonetheless, an instrument to evaluate brand equity from a customer preference perspective has been lacking (Lassar, 1995). More practical measurement applications in the areas of hotel, restaurant, theme park, club, convention center, and tourism organizations need to be reported in the brand literature (Kim, 2008). Besides, many tourism organizations such as state destination management organizations (DMOs) and convention and visitors bureaus (CVBs) urgently need to know best practices and more innovative measurements regarding destination branding (Kim, 2008). More methodology research measuring brand equity is necessary to advance the knowledge base on branding in the hospitality and tourism discipline (Kim, 2008).

Understanding how to measure brand equity is an importance issue facing hospitality brand managers. It is important for hospitality firms to measure accurately their brand equity in order to manage and leverage it properly. There are many different descriptions and definitions for both brand equity and its measurement methods (Yoo and Donthu 2001). Further, there are doubts with regard to whether the methods are capable of yielding credible and sensitive criteria (Aaker, 1996). Most of the researches on the causal relationships between brand equity and other related variables and on the construction of brand equity adopt quantitative survey-based studies as measuring tools (Prasad and Dev, 2000; Yoo and Donthu, 2001; Washburn and Plank, 2002; de Mortanges and van Riel, 2003; Kim et al., 2008). However, quantitative data analysis is criticized on grounds that it limits the utilization of real customer value (preferences) in evaluating the importance levels of the brand equity attributes (sub-dimensions of brand equity) for different customers as well as the differences in the importance levels of these brand equity attributes in different customers (Hsu et al., 2012). In other words, the traditional approach measures the importance of attributes one at a time and only compares their relative importance in aggregate although the relative importance of brand equity factors will be different across individuals. It also ignores the trade-offs that affect and characterize choice. Conjoint methods have emerged as a response to the shortcomings of other traditional methods. Conjoint analysis makes it possible to measure relative values of things considered jointly which might be unmeasurable taken one at a time (Kim et al., 2004) and enables an attribute hierarchy to be established (Peral et al., 2012). The application of conjoint methodology to consumer choice problems produces stronger results than those obtained from scale rating techniques (Huber, 1987), because it sheds light on the trade-offs that occur in the decision choice (Won and Bravo, 2009).

For this reason, conjoint analysis has been a popular research tool for modeling consumer preferences among mutliattirbute alternatives (Akaah, 1991).

Despite its popularity, there are some problems with this application of conjoint analysis. Because conjoint tasks often are complex and cognitively burdensome to respondents, potential measurement error may be a serious concern in conjoint studies (Lloyd, 2003). The conjoint method loses its appeal when a large number of attributes have to be considered, due to limitations of human cognitive capacity in sorting out a large number of profiles (Mulye, 1998). This may add incremental costs to the conduct of CA studies. Thus, it may be difficult to apply conjoint analysis in situations where the research budget is very low or where the time to conduct the study is rather limited. Consequently, the development and evaluation of preference measurement techniques that accommodate large numbers of attributes without cognitively overburdening respondents is an important and prolific research area in marketing (Bradlow, 2005).

Hospitality practitioners urgently need alternative methods that can handle many attributes in a given time span. Several preference measurement approaches try to limit information load by reducing the number of attributes shown simultaneously to the respondent in the evaluation task. Many newly developed CA variants (e.g., Adaptive Conjoint Analysis) have been developed in the last few years to deal with these problems, but none of them have proved to be dominant (Helm et al., 2004a). Empirical studies that compare Adaptive Conjoint Analysis (ACA) with full-profile conjoint analysis and self-explicated approaches respectively find a slightly poorer or at best the same internal validity of ACA (Green and Srinivasan, 1990; Green et al., 1991; Agarawal and Green, 1991; Green et al., 1993; Huber et al., 1993). Although it is probable that hybrid designs (Green, 1984) and adaptive methods employing interactive computers (Johnson, 1987) help to reduce the task difficulty and boredom factors, it is likely that full-profile designs will continue to be used for some time in marketing applications of conjoint analysis (Francois and MacLachlan, 1997). Although different data collection methods used in studies (e.g., full-profile, trade-off, and graded paired-comparisons, with rating or ranking of stimuli) have different degrees of task difficulty, it is apparent that the task is never easy (Francois and MacLachlan, 1997).

The study results revealed that AHP has some appealing features with respect to savings of time and costs in data collection, as well as motivational aspects and is equivalent to CA with respect to predictive accuracy. AHP is a possible alternative to CA for estimating preferences and almost as accurate as CA. AHP promises to be a cost effective method compared to CA. This can be seen as an advantage of AHP taking into account the costs of a survey, thereby gathering more information in a relatively short time and possibly reducing the cost of a study. In order to give a practical recommendation on the selection of the methods to hotel practitioners, this is an important finding which suggests AHP as a good alternative to CA for measuring consumer preferences regarding hotel branding because of its practical advantages in terms of ease, time effort and costs. The proposed AHP approach can be applied as an effective design method to connect the attributes of brand equity and managerial strategies that allow the practitioners of hotel firms to better understand customer preferences and evaluate and utilize their brand equity accordingly. Consequently, it provides a real guide for hotel organizations to manage marketing resource to enhance brand equity.

Providing a better technique is a critical issue in today's booming hospitality industry, since customer’s wants and needs keep changing continuously and rapidly. It is a challenge for hospitality companies to improve their existing products and develop the new ones. Increasing innovation expenditures are focusing hospitality marketers to look for faster and more precise methods for measuring consumer preferences. Furthermore, a hospitality company has to be able to develop new products conforming customer preferences in a relatively short time. Particularly in new hotel product development, probably the most important field of preference measurement in marketing, the number of possible product alternatives can be tremendous and unavoidable. AHP is capable of handling a larger number of product attributes and seems to be more feasible because of its advantages in terms of interview length and fatigue. AHP might be a good alternative to CA when evaluation tasks are complex, such as new hotel product development.

Theoretical Contributions

This study offers contributions to the hotel brand literature and the industry in several ways. To date there has been a small number of studies in hotel brand literature and focus largely on the main effects of brand equity dimensions on hotel brand equity. However, the concept of brand equity is multidimensional and very complex, requiring different types of measurement methods (Keller, 2003). In other words, measuring brand equity and proper brand equity management are important aspects of building a strong brand. Particularly, the personal decision process implied by the hierarchical brand equity model is absent. This study measures personal preferences regarding hotel branding by adopting a hierarchical approach using the AHP and the CA. It also contributes to the literature by providing the empirical test for the use of hierarchical brand equity model.

Customers tend to make trade-offs among the various choice factors since customers choose a hotel brand based on various factors and hotel brand choice is a multiattribute decision. The use of the AHP or the CA helps to overcome some of the weaknesses of traditional methods that fail to capture the trade-offs that affect and characterize choice. In this regard, it provides a better approach for assessing brand equity in the hotel industry from a customer preference perspective. The adoption of the AHP or the CA will stimulate future research on predicting consumer preferences and choice and could contribute to further research studies in the area of the hospitality and tourism academia.

Second, based on the AHP results, this study reaffirms the importance of perceived quality. Indicative of the results, more resources should give priority to perceived quality. The results of the study add support to research that contends that perceived quality plays a central role for building hotel brand equity. Third, the present study relies on a sample of actual hotel customers with diverse backgrounds, which may contribute to hotel industry practitioners as well.

Finally this study makes a significant contribution to consumer preference measurement research by quantifying the performance of two alternative measurement approaches. This study tested a compositional approach, AHP, and a decompositional approach, CA, and found that while each has different theoretical advantages, the experimental success of AHP in this study suggests that the compositional approach may be the best practical alternative. While both were of good internal and predictive validity the practical advantages of brevity, ease of completion and enjoyment by respondents favored the compositional approach. The results of the study also suggest that offline and online data collection methods are both satisfactory.

Significance of the Study

Some advantages of this study are that the use of the AHP or the CA not only removes the limitations of the traditional methods found in the extant literature but also allows brand managers to parcel out each brand equity component into its respective sub-components and estimate the relative importance of each of the primary components of brand equity and the sub-components.

Using the results provided by the AHP or the CA, hotel managers can tailor marketing mix strategies based on the order of importance of the primary components and sub-components of hotel brand equity from a customer preference perspective. This is a useful and effective way for hotel brand management to identify which brand equity components and sub-components can be improved to enhance hotel brand equity but also, constrained by limited resources, which brand equity component and sub-component should be given top priority. So the end results give opportunity to hotel practitioners to develop detailed brand equity strategies for their firms. By using either the AHP or the CA, the practitioners will be better able to prioritize tasks, allocate their resources, and develop tailored marketing strategies for their target segments.

Limitations and Future Research

This study has several limitations that need to be addressed in future research.

First, a comparison of AHP and CA demonstrated that there is almost no convergent validity with respect to attribute importance weights. More research is needed to identify why relative importance weights between the two approaches differ. It is interesting to note that importance weights estimated here by AHP are most similar to previous findings.

Second, hotel brand equity attributes and their levels were selected, based on the literature review. However, there may be some other attributes and levels of hotel brand equity that have not been identified in this study. Future studies are needed to explore and identify more attributes and levels of hotel brand equity from other sources.

Third, as Melles et al. (2000) point out, data collection methods that enhance simplification strategies (such as concentrating on a few attributes only, when the respondents' motivation is low or the complexity of the task is high) cannot be valid in predicting choices that are made through a complex trade-off. It is interesting to note that respondents in an online study tend to complete surveys more quickly than those in the offline group. It can be assumed that they must be employing simplification strategies to answer so quickly. These may have been learned as professional panelists. Unfortunately, this study could not test this assumption because it is difficult to know which respondents apply non-compensatory decision strategies and which do not and thus leave it for further research.

Fourth, the study showed that the data gathered online leads to a slightly lower internal and predictive validity. However, even though offline and online data collection methods are both satisfactory, gathering data without the help of an interviewer is not recommended. Rather more empirical work is needed concerning the influence of an interviewer (Klein et al., 2010).

Fifth, Melles et al. (2000) cautioned that the suitability of conjoint analysis over the Internet depends on the number of attributes in the design. Nonetheless, little is known about how many attributes to present in a stimulus and how many judgments to be made and thus leave this aspect for further research.

Sixth, the study results suggest that including some warm-up tasks may have a positive effect on the predictive effectiveness of AHP and CA before respondents make their evaluations, but because there have been no systematic studies comparing the impact of different warm-up tasks, further research is warranted.

Seventh, the empirical comparisons in this research were limited to CA in its basic form and to AHP. Alternative versions of AHP and CA should be applied and compared for a large of number of attributes. A compositional approach, AHP might share some weaknesses of this model class, for example, intercorrelations between attributes and levels might harm the elicitation of accurate preferences (Green and Srinivasan, 1990). Given that interactions occur principally within individual preference structures, it is easy to question whether more sophisticated AHP methods, particularly the fuzzy AHP, really lead to substantial improvements in preference measurement compared with CA. This issue requires further research.

Finally, this study did not randomize subjects between offline and online, so differences between the two groups might have arisen by sampling biases and should be replicated under randomized conditions.

REFERENCES

Aaker, D. A. (1991). *Managing brand equity*. New York: Free Press.

Aaker, D. A. and Jacobson, R. (1994). Study shows brand building pays off for Stockholders. *Advertising Age, 65*(30), 18.

Aaker, D.A. (1996). *Building strong brands*. New York: The Free Press.

Agarwal, M. K. and Green, P. E. (1991). Adaptive conjoint analysis versus self-explicated models: some empirical results. *International Journal of Research in Marketing*, *8*, 141-146.

Agarwal, M. K. and Rao, V. R. (1996). An empirical comparison of consumer-based measures of brand equity. *Marketing Letters*, *7*(3), 237-47.

Ahmad, Z. and Hashim, R. (2010). Customer's brand equity and customer loyalty: a study on hotel's conference market. *World applied sciences journal*, *10* (special issue of tourism & hospitality), 115-120.

Ajzen, I. and Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Akaah, I. P. and Korgaonkar, P. K. (1983). An empirical comparison of predictive validity of self-explicated Huber-hybrid, traditional conjoint, and hybrid conjoint models. *Journal of Marketing Research*, *20*(May), 187-197.

Akaah, I. P. (1991). Predictive performance of self-explicated, traditional conjoint, and hybrid conjoint models under alternative data collection modes. *Journal of the Academy of Marketing Science*, *19*(4), 309-314.

Ambler, T. and Styles, C. (1996). Brand development versus new product development: Towards a process model of extension decisions. *Journal of Brand Management*, *6*(1), 13-20.

Anderson, J. C. and Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*(3), 411–423.

Anderson, N. H. (1981). *Foundations of information integration theory*, New York: Academic Press.

Atilgan, E., Aksoy, S., and Akinci, S. (2005). Determinants of the brand equity – a verification approach in the beverage industry in Turkey. *Marketing Intelligence & Planning*, *23*(3), 237-248.

Aurah, C. M. (2013). The effects of self-efficacy beliefs and metacognition on academic performance: A Mixed method study. *American Journal of Educational Research, 1*(8), 334-343.

Barwise, P. (1993). Brand equity: Snark or Boojum? *International Journal of Marketing Research, 10* (March), 93-104.

Baek, S., Ham, S., and Yang, I. (2006). A cross-cultural comparison of fast food restaurant selection criteria between Korean and Filipino college students. *Hospitality Management*, *25*, 638-698

Baglione, S. L. (1994). *A Comparison of the predictive power of the analytic hierarchy process and conjoint analysis for estimating multi-attribute decision processes* (Unpublished doctoral dissertation). University of South Carolina, Columbia, South Carolina.

Baier, D. and Säuberlich, F. (1997). Kundennutzenschätzung mittels individueller Hybrid Conjointanalyse. *Zeitschrift für betriebswirtschaftliche Forschung*, *11*, 951-972.

Backhaus, K., Erichson, B., Plinke, W., and Weiber, R. (2003). *Multivariate Analysemethoden* (10th ed.). Berlin et al.: Springer.

Bakken, D. and Frazier, C. L. (2006). *Conjoint analysis: Understanding consumer decision making*. Thousand Oaks, CA: Sage Publications, Inc.

Bagozzi, R. P. and Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, *16*, 74-94.

Baldauf, A., Cravens, K. S., and Binder, G. (2003). Performance consequences of brand equity management: Evidence from organizations in the value chain. *Journal of Product & Brand Management*, *12*(4), 220-236.

Bamert, T. and Wehrli, H. P. (2005). Service quality as an important dimension of brand equity in Swiss services industries. *Managing Service Quality*, *15*(2), 132-141.

Baumgartner, H. and Homburg, C. (1996). Applications of structural equation modeling in marketing and consumer research: A Review. *International Journal of Research in Marketing*, *13*, 139-61.

Bailey, R. and Ball, S. (2006). An exploration of the meanings of hotel brand equity. *Service Industries Journal*, *26*(1), 15-38.

Berry, L. L. (2000). Cultivating service brand equity. *Journal of the Academy of Marketing Science, 28*(1), 128-137.

Berry, L. L., Parasuraman, A., and Zeithaml, V. A. (1994). Improving service quality in America: Lessons learned. *Academy of Management Executive*, *8*(2), 32-52.

Bettman, J. R., Capon, N., and Lutz, R. J. (1975). Cognitive Algebra in Multiattribute Attitude Models, *Journal of Marketing Research*, *12*(May), 151-164.

Bettman, J. R., Luce, M. F., and Payne, J. W. (1998). Constructive consumer choice processes. *Journal of Consumer Research*, *25*, 187-217.

Bettman, J. R. and Park, C. W. (1980). Effects of prior knowledge and experience and phase of choice process on consumer decision processes: A Protocol Analysis. *Journal of Consumer Research*, *7*(December), 234-248.

Bendixen, M., Bukasa, K. A., and Abratt, R. (2004). Brand equity in the business-to-business market. *Industrial Marketing Management*, *33*, 371-380.

Benaim, C., Perennou D. A., Pelissier J. Y., and Daures J. P. (2010). Using an analytical hierarchy process (AHP) for weighting items of a measurement scale: A pilot study. *Rev Epidemiol Sante Publique*, *58*(1), 59-63.

Bennett, J. and Russell B. (2001). *The choice modelling approach to environmental valuation*. Cheltenham, UK and Northampton, MA: Edward Elgar.

Bendixen, M., Bukasa, K. A., and Abratt, R. (2004). Brand equity in the business-to-business market. *Industrial Marketing Management*, *33*, 371–380.

Billings, R. S. and Marcus, S. A. (1983). Measures of compensatory and noncompensatory models of decision behavior: Process tracing versus policy capturing. *Organizational Behavior and Human Performance*, *31*, 331-352.

Biggs, S. F., Bedard, J. C., Gaber, B. G., and Linsmeier, T. J. (1985). The effects of task size and similarity on the decision behavior of bank loan officers. *Management Science*, *31*, 970-987.

Bian, X. and Moutinho, L. (2011). The Role of brand image, product involvement, and knowledge in explaining consumer purchase behavior of counterfeits: Direct and indirect Effects. *European Journal of Marketing*, *45*(1/2), 191-216.

Bollen, K. (1989). *Structural equations with latent variables*. New York: Wiley.

Bonilla, T. (2010). *Analysis of consumer preferences toward 100% fruit juice packages and labels* (Master's thesis). Louisiana State University, Baton Rouge, Louisiana.

Bodin, L. and Gass, S. I. (2003). On teaching the analytic hierarchy process. *Computers and Operations Research*, *30*, 1487-1497

Boo, S. Y., Busser, J. A., and Baloglu, S. (2009). A model of customer-based brand equity and its application to multiple destinations. *Tourism Management*, *30*(2), 219-231.

Bonoma, T. V. and Johnson, W. J. (1979). Decision making under uncertainty: A direct measurement approach. *Journal of Consumer Research*, *16*(September), 177-191.

Bowie, D. and Buttle, F. (2004). *Hospitality marketing: An introduction*. Elsevier Butterworth-Heinemann, Oxford.

Brusch, M., Baier, D., and Treppa, A. (2002). Conjoint analysis and stimulus presentation: A comparison of alternative methods. In K. Jajuga, A. Sokolowski, and H.H. Bock (Eds.), *Classification, Clustering, and Analysis* (pp. 203-10). Springer, Berlin.

Bradlow, E. T. (2005). Current issues and a 'Wish List' for conjoint analysis. *Applied Stochastic Models in Business and Industry*, *21*(4-5), 319-23.

Brucks, M. (1986). A typology of consumer knowledge content. *Advances in Consumer Research*, *13*(1), 58-63.

Bravo, G. A., Won, D., and Ferreira, M. (2009). Attributes, trade-offs and choice: A conjoint analysis in sport management programs. *Sport Management Education Journal*, *3*(1), 19-33.

Büschken, J. (1994). Conjoint-Analyse – Methodische Grundlagen und Anwendungen in der Marktforschungspraxis. In T. Tomczak, & S. Reinecke (Eds.), *Marktforschung* (pp. 72-89), St. Gallen: Thexis.

Burmann, C., Jost-Benz, M., and Riley, N. (2009). Towards an identity-based brand equity model. *Journal of Business Research*, *62*, 390-397.

Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema*, *20*(4), 872-82.

Capon, N. (2008). *Managing marketing in the 21st Century*. Wessex Pr.

Cattin, P. and Weinberger, M. G. (1980). Some validity and reliability issues in the measurement of attribute utilities. *In advances in consumer research*, 7, J.C. Olson, ed., Ann Arbor, Association for Consumer Research, 780-783.

Carmone, F. J., Green, P. E., and Jain, A. K. (1978). Robustness of conjoint analysis: Some Monte Carlo results. *Journal of Marketing Research*, *15*(2), 300-303.

Cattin, P. and Wittink, D. R. (1982). Commercial use of conjoint analysis: A survey. *Journal of Marketing*, *46*(3), 44-53.

Carroll, J. D. (1972). Individual Differences and Multidimensional Scaling. In R.N. Shepard, A.K. Romney, and S.B. Nerlove (Eds.), *Multidimensional scaling: theory and applications in the behavioral sciences: Vol. I. Theory* (pp. 105-155). New York: Seminar Press.

Carroll, J. D. and Chang, J. J. (1967). *Relating preference data to multidimensional scaling via a generalization of Coombs' unfolding model*. Paper presented at the Annual meeting of the Psychometric Society, Madison, WI.

Celsi, R. L. and Olson, J. C. (1988). The Role of involvement in attention and comprehension processes. *Journal of Consumer Research*, *15*(9), 210-224.

Cerro, D. (1988). Conjoint analysis by mail, *Proceedings of the Sawtooth Software Conference on perceptual mapping*, Sun Valley, pp. 139-143.

Chang, H. H. and Liu, Y. M, (2009). The impact of brand equity on brand preference and purchase intention in the service industries. *The Service Industries Journal*, *29*(12), 1687-1706.

Chen, C. and Chang, Y. (2008). Airline brand equity, brand preference, and purchase intentions-The moderating effects of switching costs. *Journal of Air Transport Management*, *14*, 40-42.

Cheung, G. W. and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233-255.

Chrzan, K. (1994). Three kinds of order effects in choice-based conjoint analysis. *Marketing Letters*, *5*(2), 165-172.

Chu, R.K.S and Choi, T. (2000). An Importance-Performance Analysis of hotel selection factors in the Hong Kong hotel industry: A comparison of business & leisure travelers. *Tourism Management*, *21*(4), 363-377.

Clatworthy, W. H. (1973). *Tables of two associates partially balanced designs: National bureau of standards, Applied Mathematics Series 63*. Washington, DC: U.S. Department of Commerce.

Cote, J.R., Netemeyer, R., and Bentler, P. (2001). Structural equation modeling-improving model fit by correlating errors. *Journal of Consumer Psychology*, *10*(1, 2), 87-88.

Cobb-Walgren, C. J., Ruble, C. A., and Donthu, N. (1995). Brand equity, brand preference, and purchase intent. *Journal of Advertising*, *24*(3), 25-40.

Cook, M., Angus, A., Gottberg, A., Smith, R., and Longhurst, P. (2007). Promoting sustainable resource use through product service systems. In: *CIWM Conference, Waste: A Global Resource*, Paignton, Torbay, 12-15 June 2007.

Daiber, A. and Hemsing, W. (2005). Online Conjoint: Eine bewährte Methode im neuen Gewand. *Planung & Analyse, 32*(1), 47-52.

Darmon, R. Y., and Rouziès, D. (1994). Reliability and internal validity of conjoint estimated utility functions under error-free versus error-full conditions. *International Journal of Research in Marketing*, *11*(5), 465-476.

Dawes, R. M. and Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, *81*, 95-106.

de Chernatony, L. and Segal-H., S. (2003). The Criteria for successful services brands. *European Journal of Marketing*, *37*(7/8), 1095-1118.

de Chernatony, L. and McDonald, M. (2003). *Creating powerful brands*. London: Elsevier.

de Chernatony, L. D. and Segal-Horn, S. (2001). Building on services' characteristics to develop successful services brands. *Journal of Marketing Management*, *17*(7/8), 645-669.

de Chernatony, L., Daury, S., and Segal-Horn, S. (2005). Using triangulation to assess and identify successful services brands. *Service Industries Journal*, *25*(1), 5-21.

de Chernatony, L., Harris, F. and Christodoulides, G. (2004). Developing a brand performance measure for financial service brands. *The Service Industries Journal*, *24*(2), 15-33.

Deloitte (2013). *A Restoration in Hotel Loyalty: Developing a blueprint for reinventing loyalty programs*. Retrieved from https://www.deloitte.com/assets/Dcom-UnitedStates/Local%20Assets/Documents/Consumer%20Business/us_thl_customerloyaltyPOV_012313.pdf

de Mortanges, C. P. and van Riel, A. (2003). Brand equity and shareholder value. *European Management Journal, 21*(4), 521-527.

Denstadli, J. M., Lines, R., and Ortúzar, J. de D. (2012). Information processing in choice-based conjoint experiments: A process-tracing study. *European Journal of Marketing*, *46*(3), 422-446.

Deutskens, E., Ruyter, K. D., Wetzels, M., and Oosterveld, P. (2004). Response rate and response quality of internet-based surveys: An experimental study. *Marketing Letters*, *15*(1), 21-36.

Desarbo, W. S., Ramaswamy, V., and Cohen, S. H. (1995). Market segmentation with choice-based conjoint analysis. *Marketing Letters*, *6*, 137-147.

Diamantopoulos, A., and Siguaw, J. A. (2000). *Introducing LISREL*. London: SAGE Publications, Inc.

Ding, M., Grewal, R., and Liechty, J. (2005). Incentive-aligned conjoint analysis. *Journal of Marketing Research*, *42*(1), 67-82.

Dubas, K. M. and Strong, J. T. (1993). Course design using conjoint analysis. *Journal of Marketing Education*, *15*(1), 31-36.

Duffy, B., Smith, K., Terhanian, G., and Bremer, J. (2005). Comparing data from online and face-to-face surveys. *International Journal of Market Research*, *47*, 615-639.

Dyer, R. F. and Forman, E. H. (1992). Group decision support with the Analytic Hierarchy Process. *Decision Support Systems*, *8*, 99-124.

Dyer, J. S., Fishburn, P. C., Steuer, R. E., Wallenius, J., and Zionts, S. (1992). Multiple criteria decision making, multiattribute utility theory: The next ten years. *Management Science*, *38*(5), 645-654.

Enders, C. K. (2010). *Applied Missing Data Analysis*. New York: Guilford Press.

Elrod, T., Louviere, J. J. and Davey, K. S. (1989). *How well can compensatory models predict choice for efficient choice sets?* School of Business, Vanderbilt University, Nashville, TN.

Elrod, T., Louviere, J. J., and Davey, K. S. (1992). An empirical comparison of ratings-based and choice-based conjoint models. *Journal of Marketing Research*, *29*, 368-377.

Ernst, O. and Sattler, H. (2000). Multimediale versus traditionelle Conjoint-Analysen: Ein empirischer Vergleich alternativer Produktpräsentationsformen. *Marketing-Zeitschriftfür Forschung und Praxis*, *22*, 161-172.

Farquhar, P. H. (1989). Managing brand equity. *Marketing Research*, *1*(3), 24-33.

Faller, A. J. (1981). An average correlation coefficient, *Journal of Applied Meteorology*, *20*, 203-205.

Felser, G. (1997). *Werbe- und Konsumentenpsychologie*. Stuttgart: Schäffer-Poeschel.

Finan J. S. and Hurley W. J. (1999). Transitive calibration of the AHP verbal scale. *European Journal of Operational Research*, *112*(2), 367-372.

Fishbein, M. (1963). An Investigation of relationships between belief about an object and the attitude toward the object, *Human Relations*, *16*, 233-240.

Fishbein, M. (1967). *Attitude and prediction of behavior, in readings in attitude theory and measurement*. New York: John Wiley.

Fishbein, M. and Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.

Flynn, L. R. and Goldsmith, R. E. (1999). A Short, reliable measure of subjective knowledge. *Journal of Business Research*, *46*, 57-66.

Forgacs, G. (2003). Brand asset equilibrium in hotel management. *International Journal of Contemporary Hospitality Management*, *15*(6), 340-342.

Forman, E. and Selly, M. A. (n.d.). Decision by objectives, Retrieved from http//mdm.gwu.edu/Forman/DBO.pdf

Fornell, C. and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39-50.

Fornell, C. (1992). A national customer satisfaction barometer: The Swedish experience. *Journal of Marketing*, *56*, 6-21.

Foxall, G. R. (1983). *Consumer choice. London*: The Macmillan Press Ltd.

Freitag, J. (2008). Performance of branded, independents in downturns, *Hotel and Motel Management*, *223*(19), 24.

Fricker, S., Galesic, M., Tourangeau, R., and Yan, T. (2005). An Experimental comparison of web and telephone surveys. *Public Opinion Quarterly*, *69*(3), 370-392.

Francois, P. and MacLachlan, D. L. (1997). *Bridging designs for conjoint analysis: The Issue of attribute importance*. Working paper (September 1997), 2-39.

Fukui, S., Davidson, L. J., Holter, M. C., and Rapp, C. A. (2010). Pathways to Recovery (PTR): Impact of peer-led group participation on mental health recovery outcomes. *Psychiatric Rehabilitation Journal*, *34*(1), 42-48.

Furr, R. M. (2011). *Scale Construction and Psychometrics for Social and Personality Psychology*. London, UK: Sage Publications.

Gescheider, R. D. (1976). *Psychographics: Method and theory*, Hillsdale, NJ: Lawerence Erlbaum.

Gil, J. M. and Sa´nchez, M. (1997). Consumer preferences for wine attributes: A conjoint approach. *British Food Journal*, *99*(1), 3-11.

Gilbride, T. J. and Allenby, G. M. (2004). A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science*, *23*(3), 391-406

Gold, A. H., Malhotra, A., and Segars, A. H. (2001). Knowledge management: An organizational capabilities perspective. *Journal of Management Information System*, *18*(1), 185-214.

Goldberg, S. M., Green, P. E., and Wind, Y. (1984). Conjoint analysis of price premiums for hotel amenities. *Journal of Business*, *57*(1), S111-S132.

Goldberger, A. S. (1964). *Econometric theory*, New York: John Wiley & Sons, Inc.

Green, P. E. (1974). On the design of choice experiments involving multifactor alternatives. *Journal of Consumer Research*, *1*, 61-68.

Green, P. E. and Wind, Y. (1975). New way to measure consumers' judgments. *Harvard Business Review*, *53*,107-117.

Green, P. E., and Carmone, F. J. (1977). *A BIB/logit approach to conjoint analys*is. Working paper, Wharton School, University of Pennsylvania,

Green, P. E., Carroll, J. D. and Carmone, F. J. (1978). Some new types of fractional factorial designs for marketing experiments. In J.N. Sheth (Eds.), *Research in marketing* (pp. 99-122). Greenwich: JAI.

Green, P. E., Helsen, K., and Shandler, B. (1988). Conjoint internal validity under alternative profile presentations. *Journal of Consumer Research*, *15*(December), 392-7.

Green, P. E., Krieger, A. M., and Wind, Y. (2001). Thirty years of conjoint analysis: Reflections and prospects. *Interfaces*, *31*(3), S56-S73.

Green, P. E, Tull, D. S. and Albaum, G. (1988). *Research for marketing decisions* (5th ed.). Prentice-Hall, Englewood Cliffs, NJ.

Green, P. E. and Rao, V. R. (1971). Conjoint measurement for quantifying judgmental data. *Journal of Marketing Research*, *8*(August), 355-363.

Green, P. E. and Srinivasan, V. (1990). Conjoint analysis in marketing: New developments with implications for research and practice. *Journal of Marketing*, *54*, 3-19.

Green, P. E. and Srinivasan, V. (1978). Conjoint analysis in consumer research: Issues and outlook. *The Journal of Consumer Research*, *5*(2) 103-123.

Green, P. E. (1984). Hybrid models for conjoint analysis: An expository review. *Journal of Marketing Research*, *21*(May), 155-169.

Green, P. E., Krieger, A. M., and Agarwal, M. (1991). Adaptive conjoint analysis: Some caveats and suggestions. *Journal of Marketing Research*, *28*(May), 215-222.

Green, P. E., Krieger, A. M., and Agarwal, M. (1993). A cross validation test of four models quantifying multiattributes preferences. *Marketing Letters*, *4*(4), 369-380.

Grant, D., Teller, C., and Teller, W. (2005). "Hidden" opportunities and benefits in using web based business-to-business survey. *International Journal of Market Research*, *47*, 641-666.

Hagerty, M. R. (1986). The cost of simplifying preference models. *Marketing Science*, *5*, 298-319.

Harker, P. T. (1987). Incomplete pairwise comparisons in the analytic hierarchy process. *Mathematical Modelling*, *9*(11), 837-848.

Hair, J. F. Jr., Anderson, R. E., Tatham, R. L., and Black, W. C. (1992). *Multivariate data analysis with readings*. New York: MacMillan.

Hair, Jr., J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2010*). Multivariate data analysis: A global perspective* (7th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Hair, J., Joseph F., Anderson, R. E., Tatham, R. L., and Black, W. C. (1995). *Multivariate data analysis: With readings* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Hansen, F. (1972). *Consumer choice behavior: A cognitive theory*, New York: Free Press.

Hauser, J. R. and Urban, G. L. (1979). Assessment of attribute importances and consumer utility functions: von neumann-morgenstern theory applied to consumer behavior. *Journal of Consumer Research*, *5*(March), 251-262

Hensel-Börner, S. and Sattler, H. (2000). Validity of customized and adaptive hybrid conjoint analysis. In:R. Decker & W. Gaul (Eds.), *Classification and Information Processing at the Turn of the MILLENIUM*. Proceedings of the 23rd Annual Conference of the Gesellschaft für Klassifikation, 10-14 May 1999, Bielefeld. Springer, Berlin, pp. 320-329.

Helm, R., Manthey, L., Scholl, A., and Steiner, M., (2003). *Empirical evaluation of preference elicitation techniques from marketing and decision analysis.* Jenaer

Schriftenzur Wirtschaftswissenschaft. Retrieved from http://pubdb.wiwi.uni-jena.de/pdf/JSW02-2003.pdf

Helm, R., Scholl, A., Manthey, L., and Steiner, M. (2004a). Measuring customer preferences in new product development: Comparing compositional and decompositional methods. *International Journal of Product Development*, *1*(1), 12-29.

Helm, R., Steiner, M., Scholl, A., and Manthey, L. (2004b). *A comparative empirical study on common methods for measuring preferences*. Jenaer Schriften zur Wirtschaftswissenschaft. Retrieved from http://www.wiwi.uni-jena.de/Papers/wp-sw2004.pdf

Helm, R. and Bühren, C. and Pabst, R. (2011). Using warm-up stages before measuring preferences for innovative products. *International Journal of Product Development*, *15*(4), 227-248.

Hobbs, J. E. (1996). A transaction cost analysis of quality, traceability and animal welfare issues in UK beef retailing. *British Food Journal*, *98*(6), 16-26.

Hoaglin, D. C., Mosteller, F. and Tukey, J. W. (1983). *Understanding robust and exploratory data analysis*. John Wiley &Sons, Inc., New York.

Hochgraefe, C., Faulk, S., and Vieregge, M. (2012). Links between Swiss hotel guests' product involvement and brand loyalty. *Journal of Hospitality Marketing & Management*, *21*(1), 20-39.

Holling, H., Melles, T., and Reiners, W. (1998). *How many attributes should be used in a paired comparison task?: An empirical examination using a new validation approach*. Technical report, Westfalische Wilhelms-Universitat Munster, Germany.

Howard, J. A. (1977). *Consumer behavior: Application of theory*. New York: McGraw-Hill.

Howard, J. A. and Sheth, J. N. (1969). *The theory of buyer behavior*. New York: John Wiley & Sons.

Hoyer, W. D., and Brown, S. P. (1990). Effects of brand awareness on choice for a common, repeat-purchase product. *Journal of Consumer Research*, *17*(2), 141-148.

Hsu, T., Hung, L. and Tang, J. (2012). An analytical model for building brand equity in hospitality firms. *Annals of Operations Research*, *195*(1), 355-378.

Huang, P. (2010). *The impact of rebranding on guest satisfaction and financial performance: A case study of Holiday Inn Singapore Orchard City Centre* (Master's thesis). University of Nevada, Las Vegas, Nevada.

Huber, G. P. (1974). Multiattribute utility models: A review of field and field-like studies. *Management Science*, *20*(10), 1393-1402.

Huber, J., Wittink, D. R., Fiedler, J. A., and Miller, R. (1993). The effectiveness of alternative preference elicitation procedures in predicting choice. *Journal of Marketing Research*, *30*(February), 105-114.

Huber, J. C., Wittink, D. R., Fiedler, J. A., and Miller, R. L. (1991). An empirical comparison of ACA and full profile judgments. *Sawtooth Software Conference Proceedings*, 189-202.

Huber, J. and Hansen, D. (1986). Testing the impact of dimensional complexity and affective differences of paired concepts in adaptive conjoint analysis. In M. Wallendorf & P. Anderson (Eds.), *Advances in consumer research* (No.14, pp. 159-163). Provo, UT: Association for Consumer Research.

Huber, J. (1987). Conjoint analysis: How we got here and where we are. *Proceedings of the Sawtooth Software Conference on Perceptual Mapping, Conjoint Analysis, and Computer Interviewing* (pp. 237-252). Sequim, WA: Sawtooth Software, Inc.

Hu, C. and Hiemstra, S. J. (1996). Hybrid conjoint analysis as a research technique to measure meeting planners’ preferences in hotel selection. *Journal of Travel Research*, *35*(2), 62-69.

Huber, J. (1997). *What we have learned from 20 years of conjoint research: When to use self explicated, graded pairs, full profiles or choice experiments*. Sequim, WA: Sawtooth Software, Inc.

Hyun, S. S. (2009). Creating a model of customer equity for chain restaurant brand formation. *International Journal of Hospitality Management*, *839*, 1-11.

Ijzerman, M. J., van Til, J. A., and Snoek, G. J. (2008). Comparison of two multi-criteria decision techniques for eliciting treatment preferences in people with neurological disorders. *Patient*, *1*(4), 265-272.

Ijzerman M. J., van Til J. A., and Bridges J. F. (2012). A comparison of analytic hierarchy process and conjoint analysis methods in assessing treatment alternatives for stroke rehabilitation. *Patient*, *5*(1), 45-56.

InterContinental Hotels Group (2007). *IHG making the first major change to the Holiday Inn logo in more than 50 years; will replace more than 11,000 signs around the* *world by 2010*. Retrieved from http://www.hotel-online.com/News/PR2007_4th/Oct07_HIRelaunch.html

Jaccard, J., Brinberg, D., and Ackerman, L. J. (1986). Assessing attribute importance: A comparison of six methods. *Journal of Consumer Research*, *12*(March), 463-468.

Jensen, R. E. (1983). Aggregation (composition) schema for eigenvector scaling of criteria priorities in hierarchical structures. *Multivariate Behavioral Research*, *18*(1), 63-84.

Ji, P. and Jiang, R. (2003). Scale transitivity in the AHP. *Journal of the Operational Research Society*, *54*(8), 896-905.

Johnson, R. M. and Orme, B. K. (1996). How many questions should you ask in choice-based conjoint studies? *Sawtooth Software Research Paper Series*, Sequim (WA).

Johnson, R. (1974). Tradeoff analysis of consumer values. *Journal of Marketing Research*, *11*(May), 121-127.

Johnson, R. (1997). Including holdout choice tasks in conjoint studies. *Sawtooth software research paper series*, Sequim.

Johnson, C. R. (1980). Constructive critique of a hierarchical prioritization scheme employing paired comparisons. *Proceedings of the International Conference of Cybernetics and Society of the IEEE*, Institute of Electrical Engineers, Cambridge MA, 373-378.

Johnson, E. J. and Meyer, R. J. (1984). Compensatory choice models of noncompensatory processes: The effect of varying context, *Journal of Consumer Research*, *11*(June), 528-541.

Johnson, R. M. (1989). Assessing the validity of conjoint analysis. *Sawtooth Software Conference Proceedings*, Ketchum, ID: Sawtooth Software, June, 273-280.

Johnson, R. M. (1987). Adaptive conjoint analysis. *Sawtooth Conference on Perceptual Mapping, Conjoint Analysis, and Computer Interviewing*, Ketchum, ID: Sawtooth Software, 253-266.

Johnston, J. (1972). *Econometric methods* (2nd ed.). New York: McGraw-Hill Book Co.

Johnston, R. (2004). Towards a better understanding of service excellence. *Managing Service Quality Journal*, *14*(2/3), 129-133.

Jung, J. and Sung, E. Y. (2008).Consumer-based brand equity: Comparisons among South Koreans in the USA and South Koreans in Korea. *Journal of Fashion Marketing and Management*, *12*(1), 24-35.

Kallas, Z., Lambarraa, F., and Gil, J. M. (2011). A stated preference analysis comparing the analytical hierarchy process versus choice experiments. *Food Quality and Preference*, *22*(2), 181-192

Kamenetzky, R. D. (1982). The relationship between the analytic hierarchy process and the additive value function. *Decision Sciences*, *13*, 702-713.

Kara, A. (1993). *An experimental investigation of the predictive validity of selected hybrid conjoint models and adaptive analytic hierarchy in the context of consumer choice* (Doctoral dissertation). Florida International University, Miami, Florida.

Karpetrovie, S. and Rosenbloom, E. S. (1999). A quality control approach to consistency paradoxes in AHP. *European Journal of Operational Research*, *119*, 704-718.

Kay, M. (2006). Strong brands and corporate brands. *European Journal of Marketing*, *40*(7/8), 742-760.

Kayaman, R. and Arasli, H. (2007). Customer based brand equity: Evidence from the hotelindustry. *Managing Service Quality*, *17*(1), 92-109.

Kempf, D. S. and Smith, R. E. (1998). Consumer processing of product trial and the influence of prior advertising: A structural modelling approach. *Journal of Marketing Research*, *35*(3), August, 325-338.

Keller, K. L. (1993). Conceptualizing, measuring, and managing consumer-based brand equity. *Journal of Marketing*, *57*(1), 1-22.

Keller, K. L. (2000). The brand report card. *Harvard Business Review*, *78*(1), 147-55.

Keller, K. L. (2001). Building customer-based brand equity. *Marketing Management*, *10*(2), 14-19.

Keller, K. L. and Lehmann, D. R. (2006). Brands and branding: Research findings and future priorities. *Marketing Science*, *25*(6), 740-759.

Keller, K. L. (1998). *Strategic brand management: building, measuring, and managing brand equity*. New Jersey: Prentice Hall.

Keller, K. L. (2003). *Strategic brand management: building, measuring and managing brand equity*. Upper Saddle River: Prentice-Hall.

Khan, P. I. and Tabassum, A. (2010-2011). Service quality and customer satisfaction of the beauty-care service industry in Dhaka: A Study on high-end women‘s parlors. *Journal of Business in Developing Nations*, *12*, 33-58.

Kim, H. B., and Kim, W. G. (2005). The relationship between brand equity and firms' performance in luxury hotels and chain restaurants. *Tourism Management*, *26*(4), 549-560.

Kim, H. B., Kim, W. G., and An, J. A. (2003). The effect of consumer-based brand equity on firms' financial performance. *Journal of Consumer Marketing*, *20*(4), 335-351.

Kim, K. H., Kim, K. S., Kim, D. Y., Kim, J. H., and Kang, S. H. (2008). Brand equity in hospital marketing. *Journal of Business Research*, *61*, 75-82.

Kim, T., Kim, W. G., and Kim, H. B. (2009). The effects of perceived justice on recovery satisfaction, trust, word-of-mouth, and revisit intention in upscale hotels. *Tourism Management*, *30*, 51-62.

Kim, W. G., and Kim, H. J. (2007). Determinants of mid-scale hotel brand equity, *FIU Hospitality and Tourism Review*, *25*(1), 61-69.

Kim, W. G. (2008). Branding, brand equity, and brand extensions. *In Handbook of hospitality marketing management* (pp. 87-119). Oxford, UK: Butterworth-Heinemann.

Kim, W. G., DiMicelli, P., and Kang, J. (2004). Using conjoint analysis to measure brand equity of restaurant chains. *FIU Hospitality Review*, *22*(2), 28-41.

Kim, W. G., Jin-Sun, B., and Kim, H. J. (2008). Multidimensional customer-based brand equity and its consequences in mid-priced hotels. *Journal of Hospitality & Tourism Research*, *32*(2), 235-254.

Kim, W. G. and Kim, H. B. (2004). Measuring customer-based restaurant brand equity: Investigating the relationship between brand equity and firms' performance. *Cornell Hotel and Restaurant Administration Quarterly*, *45*(2), 115-131.

Kim, Y. S. and Kim, C. W. (2006). A study on the scale development of hotel brand equity using ZMET. *International Journal of Tourism Sciences*, *7*(1), 1-28.

Klein, A., Nihalani, K., and Krishnan, K. S. (2010). A comparison of the validity of interviewer-based and online-conjoint analyses. *Journal of Management and Marketing Research*, *2*(1), 1-15.

Konecnik, M. and Gartner, W. C. (2007). Customer-based brand equity for a destination. *Annals of Tourism Research*, *34*(2), 400-421.

Koo, H. Y. H., and Koo, L. C. (2010). Empirical examination of AHP and conjoint analysis on casino attributes in Macau. *Conference Proceedings of an International Conference on Public Welfare and Gaming Industry*, Beijing 9-11 December, pp. 327-350.

Kotler, P., Bowen, J., and Makens, J. (2006). *Marketing for Hospitality and Tourism* (4th ed.). Upper Saddle River, NJ: Prentice-Hall.

Kotler, P. (2003). *Marketing management*. New Jersey: Prentice Hall.

Krapp, A. and Sattler, H. (2001). Rethinking Preference Measurement. In: Breivik, E.; Falkenberg, A.W.; Gronhaug, K. (Hrsg.), *Conference Proceedings of Rethinking European Marketing – 30th EMAC Conference*, Bergen.

Krieger, A. M. and Green, P. E. (1988). *On the generation of Pareto optimal, conjoint profiles from orthogonal main effects plans*. Working Paper, Wharton School, University of Pennsylvania, (August).

Kroeber-Riel, W. and Weinberg, P. (1996). *Konsumentenverhalten*. München: Franz Vahlen.

Kruskal, J. (1965). Analysis of factorial experiments by estimating monotone transformations of data. *Journal of the Royal Statistical Society, Series B*, *27*, 251-263.

Kruskal, J. B. and Carmone, F. J. Jr. (1969). MONANOVA: A Fortran-iv program for monotone analysis of variance. *Behavioral Science*, *14*, 165-166.

Lawson, S. R., Roggenbuck, J. W., and Hal, T. E. (2006). A conjoint analysis of preference heterogeneity among day and overnight visitors to the Okefenokee wilderness, *Journal of Leisure Research*, *38*(4), 575-600.

Lassar, W., Mittal, B., and Sharma, A. (1995). Measuring customer-based brand equity. *Journal of Consumer Marketing*, *12*(4), 11-19.

Lancaster, K. (1991). *Modern consumption theory*. Aldershot: Edward Elgar.

LeBlanc, G. and Nguyen, N. (1996). An examination of the factors that signal hotel image to travelers. *Journal of Vacation Marketing*, *3*(1), 32-42.

Lee, J. S. and Back, K. L. (2008). Attendee-based brand equity. *Tourism Management*, *29*, 331-344.

Lewis, R. C., Ding, S., and Geschke, U. (1991). Using trade-off analysis to measure consumer choices: The full profile method. *Hospitality Research Journal*, *15*, 75-92.

Leian, Ti. B. and Widdows, R. (1999). Product knowledge and product involvement as moderators of the effects of information on purchase decisions: A case study using the perfect information frontier approach. *Journal of Consumer Affairs*, *33*(1), 165-186.

Leigh, T. W., MacKay, D. B., and Summers, J. O. (1981). On alternative experimental methods for conjoint analysis. *Advances in Consumer Research*, *8*, 317-322.

Li, X., and Krit, J. (2012). Service is power: Exploring service quality in hotel's business. *International Business Research*, *5*(5), 35-48.

Lines, R. and Denstadli, J. M. (2004). Information overload in conjoint experiments. *International Journal of Market Research*, *46*(3), 297-310.

Lin, L. Y. and Chen, C. S. (2006). The influence of the country-of-origin image, product knowledge and product involvement on consumer purchase decisions: An empirical study of insurance and catering services in Taiwan Graduate School of Management Sciences, Aletheia University, Taipei, Taiwan. *Journal of Consumer Marketing*, *23*(5), 248-265.

Lloyd, A. J. (2003). Threats to the estimation of benefit: are preference elicitation methods accurate? *Health Economics*, *12*(5), 393-402.

Louviere, J. J. (1988a), *Analyzing decision making: Metric conjoint analysis*. Newbury Park: Sage Publications Inc.

Louviere, J. J. (1988b). Conjoint analysis modeling of stated preferences: A review of theory, methods, recent developments and external validity. *Journal of Transport Economics and Policy*, *10*, 93-119.

Low, G. S., and Lamb Jr., C. W. (2000). The measurement and dimensionality of brand associations. *Journal of Product and Brand Management*, *9*(6), 350-368.

Lund, D. B., Malhotra, N. K., and Smith, A. K. (1988). Field validation study of conjoint analysis using selected mail survey response rate facilitator. *Journal of Business Research*, *16*, 351-368.

MacCallum, R. C., Browne, M.W., and Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*(2), 130-49.

Mackay, N., Spies, H., Williams, C., Jansen van Rensburg, L. R., and Petzer, D. J. (2012). The influence of service brand equity on the strength of brand relationships in the fast food industry. *Southern African Business Review*, *17*(2), 67-92

Mackay, M. M. (2001). Application of brand equity measures in service markets. *The Journal of Services Marketing*, *15*(3), 210-221.

Madigan, M. P, Troisi, R., Potischman, N., Brogan D., Gammon, M. D., Malone, K. E., and Brinton, L. A. (2000). Characteristics of respondents and non-respondents from a case-control study of breast cancer in younger women. *International Journal of Epidemiology*, *29*(5), 793-798.

Majláth, M. (2009). Evaluation of environmentally friendly product attribute -results of an empirical research. *Proceedings of the MEB 7th International Conference on Management, Enterprise and Benchmarking*, Budapest, June 5-6, 2009, 201-212.

Malik, M. E., Naeem, B., and Nasir, A.M. (2011). Hotel service quality and brand loyalty. *Interdisciplinary Journal of Contemporary Research in Business*, *3*(8), 620-629.

Malvinas, F., Mangkoesubroto, K., Suryadi, K., and Yudhistira, T. (2005). *Development of customer decision model in selecting product concept based on conjoint-analysis like AHP (CALAHP).* Communication presented at the International symposium on the Analytic Hierarchy Process. Hawaii: Honolulu.

Marsh, H. W. and Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept first- and higher-orderfactor models and their invariance across groups. *Psychological Bulletin*, *97*, 562-582.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*, 519-530.

Markovic, S. and Raspor, S. (2010). Measuring perceived service quality using SERVQUAL: A case study of the Croatian hotel industry. *Journal of management*, *5*(3), 95-209.

McCullough, D. (2002). A user's guide to conjoint analysis. *Marketing Research*, *14*(2), 18-22.

McFadden, D. (1976). Quantal choice analysis: A survey. *Annals of Economic and Social Measurement*, *5*, 363-390.

McFadden, D. (1980). Econometric models of probabilistic choice. *Journal of Business*, *53*(July), S13-19.

McFadden, D. (1986). The choice theory approach to market research. *Marketing Science*, *5*(4), 275-297.

Meißner, M. and Decker, R. (2009). *An empirical comparison of CBC and AHP for measuring consumer preferences*. International Symposium of Analytical Hierarchy Process, Pittsburgh.

Meißner, M., Scholz, S. W., and Decker, R. (2008). AHP versus ACA- An empirical comparison. In C. Preisach et al. (Eds.), *Data analysis, machine learning, and applications* (pp. 447-454), Berlin: Springer.

Melles, T., Laumann, R., and Holling, H. (2000). Validity and reliability of online conjoint analysis. *In Proceedings of the Sawtooth Software Conference*, Sequim, WA: Sawtooth Software, *8,* 31-40.

Meeds, R. (2004). Cognitive and attitudinal effects of technical advertising copying: the roles of gender, self-assessed and objective consumer knowledge. *International Journal of Advertising*, *23*(3), 309-335

Meade, A. W., Johnson, E. C., and Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology*, *93*(3), 568-592.

Meyers, L., Gamst, G., and Guarino, A. (2006). *Applied multivariate research: Design and interpretation*. Thousand Oaks, CA: Sage Publications, Inc.

Milfont, T. L. and Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, *3*(1), 111-121.

Mistry, B. (1998). Life and soul of the brand. *Marketing*, *26*(March), 47-49.

Mishra, S., Umesh, U., and Stem, D. (1989). Attribute importance weights in conjoint analysis: Bias and precision. *Advances in Consumer Research*, *16*, 605-611.

Mishra, P. and Dutta, B. (2011). Perpetual asset management of customer based brand equity- The PAM Evaluator. *Current Research Journal of Social Science*, *3*, 34-43.

Mola, F. and Jusoh, J. (2011). Service quality in Penang hotels: A gap score analysis. *World Applied Sciences Journal* (Special Issue of Tourism & Hospitality), *12*, 19-24.

Moradi, H. and Zarei, A. (2011). The impact of brad equity on purchase intention and brand preference- the moderating effects of country of origin image. *Australian Journal of Basic and Applied Sciences, 5* (3), 539-545.

Moreau, C. P., Lehmann, D. R., and Markman, A. B. (2001). Entrenched Knowledge Structures and Consumer Response to New Products. *Journal of Marketing Research*, *38*(1), 14-29.

Moore, W. L. (2004). A cross-validity comparison of rating-based and choice-based conjoint analysis models. *International Journal of Research in Marketing*, *21*(3), 299-312.

Moore, W. L. and Holbrook, M. B. (1990). Conjoint analysis on object with environmentally correlated attributes: The questionable importance of representative design. *Journal of Consumer Research*, *16*(March), 490-497.

Moore, W. L., Gray-Lee, J., and Louviere. J. J. (1998). A cross-validity comparison of conjoint analysis and choice models at different levels of aggregation. *Marketing Letters*, *9*, 195-207.

Mulhbacher, H. and Botschen, G. (1988). The use of trade-off analysis for the design of holiday travel packages. *Journal of Business Research*, *17*, 117-131.

Mulye, R. (1998). An empirical comparison of three variants of the AHP and two variants of conjoint analysis. *Journal of Behavioral Decision Making*, *11*, 263-280.

Myers, C. A. (2003). Managing brand equity: a look at the impact of attributes. *The Journal of Product & Brand Management*, *12*(1), 39-51.

Nedungadi, P. (1990). Recall and consumer consideration sets: Influencing choice without altering brand evaluations. *Journal of Consumer Research*, *17*, 263-276.

Nehis, L., Seaman, B., and Montgomery, D. B. (1976). *A PL1 program for trade-off analysis*. Cambridge, MA: Marketing Science Institute.

Nel, J. D.W., North, E. J., Myburgh, T. and Hern, L. (2009). A comparative study of customer-based brand equity across selected South African hotels. *International Retail and Marketing Review*, *5*(1), 15-24.

Netzer, O. and Srinivasan, V. (2011). Adaptive self-explication of multiattribute preferences. *Journal of Marketing Research*, *48*(1), 140-156

Netemeyer, R. G., Krishnan, B., Pullig, C., Wang, G., Yagci, M., Dean, D., Ricks, J., and Wirth, F. (2004). Developing and validating measures of facets of customer-based brand Equity. *Journal of Business Research*, *57*(2), 209-224.

Newell, A. and Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Nunnally, J. and Bernstein, I. (1994). *Psychometric theory*. New York: Mcgraw-Hill.

Oh, H. (2001). Revisiting importance-performance analysis, *Tourism Management*, *22*, 617-627.

Olshavsky, R. W. and Acito, F. (1980). An information processing probe into conjoint analysis. *Decision Sciences*, *11*(July), 451-470.

O'Neil, J. W. and Xiao, Q. (2006). The role of brand affiliation in hotel market value. *Cornell Hotel and Restaurant Administration Quarterly, 47*(3), 210-223.

Orme, B. (1996). Which conjoint method should I use? *Research Paper Series*, Sawtooth Software, Inc.

Orme, B. K., Alpert, M. I., and Christensen, E. (1997). Assessing the validity of conjoint analysis-continued. *Sawtooth Software Research Paper Series*, Sequim.

Orme, B. K. (2002a). Formulating attributes and levels in conjoint analysis. *Sawtooth Software Research Paper Series*, Sequim.

Orme, B. (2002b). Interpreting Conjoint Analysis Data, pp.1-4, *Research Paper Series*, Sawtooth Software Inc., Orem, UT.

Orme, B. (2006). *Getting started with conjoint analysis: Strategies for product design and pricing research*. Madison, WI: Research Publishers LLC.

Orme, B. (2010). *Getting started with conjoint analysis: Strategies for product design and pricing research* (2nd ed.). Madison, Wis.: Research Publishers LLC.

Orme, B. K. and King, W. C. (1998). *Conducting full-profile conjoint analysis over the internet.* Research Paper Series, Sawtooth Software, Inc, *1*, 1-14.

Pappu, R., Quester, P. G., and Cooksey, R. W. (2005). Consumer-based brand equity: improving the measurement-empirical evidence. *The Journal of Product & Brand Management*, *14*(3), 143-154.

Parasuraman, A., Zeithaml, V. A., and Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality, *Journal of Retailing*, *64*(1), 14-40.

Park, C. W., and Lessing, V. P. (1981). Familiarity and its impact on consumer decision biases and heuristics. *Journal of Consumer Research*, *8*(September), 223-230.

Payne, J. W. (1976). Task complexity and contingent processing in decision processing: An information search and protocol analysis. *Organizational Behavior and Human Decision Processes, 16*, 366-387.

Park, C. W. and Moon, B. J. (2003). The relationship between product involvement and product knowledge: Moderation. *Psychology & Marketing*, *20*(11), 977.

Peral, B. P., Rey. J. R-B., Ramos, A. F. V., and Sevilla, U. D. (2012). A study of consumer preferences for e-retailers' attributes: an application of conjoint analysis. *The International Journal of Management Science and Information Technology* (IJMSIT), *1*(3), 37-67.

Pecchia, L., Bracale, U., Melillo, P., Sansone, M., and Bracale, M. (2009). AHP for Health Technology Assessment. A case study: prioritizing care approaches for patients suffering from chronic heart failure. In: Y. Tjader. *Proceedings of the 10th International Symposium on the Analytic Hierarchy/Network Process.* Pittsburgh, Pennsylvania, USA, July 29-August 1, 2009, pp. 1-8, Pennsylvania.

Pecchia, L., Bath, P. A., Pendleton, N., and Bracale, M. (2009). AHP and risk management: a case study for assessing risk factors for falls in community-dwelling older patients. *In ISAHP 2009 Symposium*. Pittsburgh, Pennsylvania, USA, July 29-August 1, 2009, pp. 1-15.

Pecchia, L., Bath, P. A., Pendleton, N., Jackson, S., Clarke, C., Briggs, P., McInnes, L., Angelova, M., and Bracale, M. (2011). The use of analytic hierarchy process for the prioritization of factors affecting wellbeing in elderly. In: *ISAHP2011*. Sorrento (ITALY), 15-18 Giugno 2011, pp. 1-4.

Peter, J. P. and Olson, J. C. (2001). *Consumer Behavior and Marketing Strategy – International Edition* (6th ed.). New York: McGraw-Hill.

Pignone, P., Brenner, A., Hawley, S., Sheridan, S., Lewis, C., Jonas, D., and Howard, K. (2012). Conjoint analysis versus rating and ranking for values elicitation and clarification in colorectal cancer screening. *Journal of General Internal Medicine*, *27*(1), 45-50.

Ping, R. A., Jr (2004). On assuring valid measures for theoretical models using survey data. *Journal of Business Research*, *57*, 125-41.

Polit, D. F. and Beck, C. T. (2008). *Nursing research: Generating and assessing evidence for nursing practice* (8th ed.). Philadelphia: Wolters Kluwer Health.

Pöyhönen, M. and Hämäläinen, R. (2001). On the convergence of multiattribute weighting methods. *European Journal of Operational Research*, *129*(3), 106-122.

Prasad, K., and Dev, C. S. (2000). Managing hotel brand equity: A customer-centric framework for assessing performance. *The Cornell Hotel and Restaurant Administration Quarterly*, *41*(3), 22-31.

Punj, G. N. and Staelin, R. (1978). The choice process for graduate business schools, *Journal of Marketing Research*, *15*(November), 588-598.

Rao, V. R. and Winter, F. W. (1977). Application of the multivariate probit model for market segmentation and product design. *Working Paper 338*, School of Business Administration, University of Illinois at Urbana.

Ratchford, B. T. (1975). The new economic theory of consumer behavior: An interpretive essay. *Journal of Consumer Research*, *2*(September), 65-78.

Renaghan, L. M. and Kay, M. Z. (1987). What meeting planners want: the conjoint analysis approach. *The Cornell Hotel and Restaurant Administration Quarterly*, *28*(1), 67-76.

Reid, R. and Bojanic, D. (2009). *Hospitality Marketing Management* (5th ed.). New York: John Wiley & Sons, Inc.

Reichheld, F. F., and Aspinall, K. (1993). Building high-loyalty business system. *Journal of Retail Banking*, *15*(4), 21-29.

Riley, D., Ehrenberg, A. S. C., Castleberry, S. B., Barwise, T. P., and Barnard, N. R. (1997). The variability of attitudinal repeat-rates. *International Journal of Research in Marketing*, *14*(5), 437-450.

Rosenberg, M. J. (1956). Cognitive structure and attitudinal effect. *Journal of Abnormal and Social Psychology*, *53*, 367-372.

Rothschild, M. L. (1975*). The Effects of political advertising on the voting behavior of a low involvement electorate* (Doctoral dissertation). Stanford University, CA.

Saltzman, A. and MacElroy, W. H. (1999). Disk-based mail surveys: A longitudinal study of practices and results. *In Proceedings of the Sawtooth Software Conference* (No.7, pp. 43-53). Sequim, WA: Sawtooth Software.

Saaty, T. L. (1977). A Scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, *15*(3), 234-281.

Saaty, T. L. (1980). *The Analytic Hierarchy Process*. McGraw-Hill, New York.

Saaty, T.L. (1990). An exposition of the AHP in reply to the paper ‘Remarks on the Analytic Hierarchy Process’ *Management Science*, *36*, 259-268.

Saaty, T. L. (1994). *Fundamentals of decision making and priority theory with the AHP*. RWS Publications, Pittsburgh, PA, U.S.A.

Saaty, T. L. (1996). *Decision making with dependence and feedback: The Analytic Network Process*. RWS Publications, Pittsburgh, PA, U.S.A.

Saaty, T. L. (2003). Decision-making with the AHP: Why is the principal eigenvector necessary? *European Journal of Operational Research*, *145*(1), 85-91.

Salo, A. A. and Hämäläinen, R. P. (1997). On the measurement of preference in the Analytic Hierarchy Process. *Journal of Multi-Criteria Decision Analysis*, *6*(6), 309-319.

Sattler, H., Hensel-Börner, S., and Krüger, B. (2001). Die Abhängigkeit der Validität von Conjoint-Studien von demographischen Probanden-Charakteristika. *Zeitschrift für Betriebswirtschaft*, *71*(7), 771-787.

Sattler, H. and Nitschke, T. (2003). Ein empirischer Vergleich von Instrumenten zur Erhebung von Zahlungsbereitschaften. *Zeitschrift für betriebswirtschaftliche Forschung*, *55*, 364-381.

Schillewaert, N. and Meulemeester, P. (2005). Comparing response distributions of offline and online data collection methods. *International Journal of Market Research*, *47*, 163-178.

Scholl, A., Manthey, L., Helm, R., and Steiner, M. (2005). Solving multiattribute design problems with analytic hierarchy process and conjoint analysis: An empirical comparison. *European Journal of Operational Research*, *164*(3), 760-777.

Scholz, S. W., Meissner, M. and Wagner, R. (2006). Robust Preference Measurement: A Simulation Study of Erroneous and Ambiguous Judgement's Impact on AHP and Conjoint Analysis. In: H.-O. Haasis, H. Kopfer and J. Schönberger (Eds.): OR *Proceedings 2005*, Springer, Berlin, 613-618.

Scholz, S. W., Meissner, M., and Decker, R. (2010). Measuring consumer preferences for complex products: A compositional approach based on paired comparisons. *Journal of Marketing Research*, *47*, 685-698.

Scholz, M. (2008). From consumer preferences towards buying decisions - conjoint analysis as preference measuring method in product recommender systems. *Proceedings of the 21st Bled eConference*, 2008, Bled, Slovenia.

Scholz, M. (2010). Implications of consumer information behaviour to construct utility-based recommender systems-A prototypical study. *Proceedings of the 23rd Bled eConference*, 2010, Bled, Slovenia.

Schoner, B. and Wedley, W. C. (1989). Ambiguous criteria weighted in AHP: Consequences and solutions, *Decision Sciences*, *20*, 462-475.

Sethuraman, R., Kerin, R. A., and Cron. W. L. (2005). A field study comparing online and offline data collection methods for identifying product attribute preferences using conjoint analysis. *Journal of Business Research*, *58*, 602-610.

Shafir, E., Simonson, I., and Tversky, A. (1993). Reason-based choice. *Cognition*, *49*(1-2), 11-36.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99-118.

Simoes, D. and Dibb, S. (2001). Rethinking the brand concept: new brand orientation. *Corporate Communications*, *6*(4), 217-24.

Silk, A. J. (2006). *What is marketing?* Boston, MA: Harvard Business School Press.

Simon, M. (2011). Brands in context. *Journal of Advertising Research*, *51*, 189-194.

Simmerman, S.J. (1992). Improving customer loyalty. *Business and Economics Review*, *38*(3), 3-6.

Sinha, A., Ashill, N.J., and Gazley A. (2008). Measuring customer based brand equity using Hierarchical Bayes methodology. *Australasian Marketing Journal*, *16*(1), 3-19.

Sloot, L. M., Verhoef, P. C., and Franses, P. H. (2005). This impact of brand equity and the hedonic level of a product on consumer stock out reactions. *Journal of Retailing*, *81*(1), 14-35.

Smith, D. C. and Park, C. W. (1992). The effects of brand extensions on market share and advertising efficiency. *Journal of Marketing Research*, *29*(3), 296-313.

Smith, S. M. and Albaum, G. (2005). *Fundamentals of marketing research*. Thousand Oaks, CA: Sage Publications.

So, K. K. F. and King, C. (2010). When experience matters: Building and measuring hotel brand equity: The customers' perspective. *International Journal of Contemporary Hospitality Management*, *22*(5), 589-608.

Sorenson, D. and Bogue, J. (2005). A conjoint-based approach to concept optimisation: Probiotic beverages. *British Food Journal*, *107*(11), 870-883.

Sönmez, H. and Hacıköylü, B. (2012). Determining the importance levels of the criteria considered for giving food and accommodation grant to university students through analytic hierarchy process (AHP) and conjoint analysis (CA). *African Journal of Business Management*, *6*(14), 4742-4751.

Srinivasan, V. and Shocker, A.D. (1973). Linear programming techniques for multidimensional analysis of preferences. *Psychometrika*, *38*(3), 337-369.

Srinivasan, S. V. and Park, C. S. (1997). Surprising robustness of self-explicated approach to customer preference structure measurement. *Journal of Marketing Research*, *34*, 286-291.

Strebinger, A., Hoffmann, S., Schweiger, G., and Otter, T. (2000). Zur Realitätsnähe der Conjointanalyse. *Marketing Zeitschrift fürForschung und Praxis*, *22*, 55-74.

Steckel, J. H., DeSarbo, W. S., and Mahajan, V. (1990). On the Creation of feasible conjoint analysis experimental designs. *Decision Sciences*, *22*(2), 435-444.

Steenkamp, J. E. M. and Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, *25*(1), 78-107.

Stillwell, W. G., Barron, F. H., and Edwards, W. (1983). Evaluating credit applications: A validation of multiattribute utility weight elicitation techniques. *Organizational Behavior & Human Decision Processes*, *32*(1), 87-108.

Stahl, B. (1988), Conjoint Analysis by Telephone. In: Sawtooth Software Inc. (Hrsg.): *Sawtooth Software Conference Proceedings.* Ketchum, ID: Sawtooth Software Inc., S131-138.

Sun, L. and Ghiselli, R. F. (2010). Developing a conceptual model of brand equity in the hotel industry based on Aaker's Perspective. *Journal of Quality Assurance in Hospitality & Tourism*, *11*(3), 147-161.

Sundström, G. A. (1987). Information search and decision making: The effects of information displays. *Acta Psychologica*, *65*, 165-179

Tasci, A. D. A., Gartner, W. C., and Cavusgil, S. T. (2007). Conceptualization and operationalization of destination image. *Journal of Hospitality & Tourism Research*, *31*, 194-223.

Tashchian A., Tashchian, R. O., and Slama, M. E. (1982). The impact of individual differences on the validity of conjoint analysis. *Advances in Consumer Research*, *9*(1), 363-366.

Tabachnick, B., and Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). SanFrancisco, CA: Allyn and Bacon.

Tepeci, M. (1999). Increasing brand loyalty in the hospitality industry. *International Journal of Contemporary Hospitality Management*, *11*(5), 223-229.

Todd, P. M. (2007). How much information do we need? *European Journal of Operational Research*, *177*(3), 1317-1332.

Tolba, A. H. and Hassan, S. S. (2009). Linking Customer based brand equity with brand market performance: A managerial approach. *Journal of Product and Brand Management*, *18*(5), 356-366.

Tscheulin, D.K. (1991). Ein empirischer Vergleich der Eignung von Conjoint-Analyse und Analytic Hierarchy Process (AHP) zur Neuproduktplanung. *Zeitschrift für Betriebswirtschaft*, *61*(11), 1267-1279.

Tscheulin, D. K. (1992). *Optimale Produktgestaltung-Erfolgsprognose mit Analytic Hierarchy Process und Conjoint-Analyse*. Gabler, Wiesbaden.

Tscheulin, D. K. and Blaimont, C. (1993). Die Abhängigkeit der Prognosegüte von Conjoint-Studien von demographischen Probanden-Charakteristika. *Zeitschrift für Betriebwirtschaft*, *63*, 839-847.

Tull, D. S. and Hawkins, D. I. (1993). *Marketing research: Measurement and method* (6th ed.). Englewood Cliffs: Prentice-Hall.

Tukey, J. W. (1960). A survey of sampling from contaminated distributions. Contributions to Probability and Statistics, Olkin and Others, *Stanford University Press*, 448-485.

Ukpebor, P. and Ipogah, B. (2008). *A study to indicate the importance of consumer based-brand equity on consumer perception of brand* (Master thesis). Blekinge Tekniska Hogskola (BTH).

Vakratsas, D. and Ambler. T. (1999). How advertising works: What do we really know? *Journal of Marketing*, *63*(1), 26-43.

van Til, J. A., Dolan, J. G., Stiggelbout, A. M., Groothuis, K. C., and Ijzerman, M. J (2008). The use of multi-criteria decision analysis weight elicitation techniques in patients with mild cognitive impairment: A pilot study. *Patient*, *1*(2), 127-135.

van Ittersum, K., Pennings, J. M. E., Wansink, B., and van Trijp, H. C. M. (2007). The validity of attribute-importance measurement: A review. *Journal of Business Research*, *60*(11), 1177-1190.

Voeth, M. (2000). *Nutzenmessung in der Kaufverhaltensforschung*. Wiesbaden.

Vriens, M. (1995). *Conjoint analysis in marketing: Developments in stimulus representation and segmentation methods* (Unpublished doctoral dissertation). Groningen University.

Vriens, M., Loosschilder, G. H., Rosbergen, E., and Wittink, D. R. (1998). Verbal versus realistic pictorial representation in conjoint analysis with design attributes. *Journal of Product Innovation Management*, *15*, 455-467.

Wang, H., Wei, Y. and Yu, C. (2008). Global brand equity model: Combining customer-based with product market outcome approaches. *Journal of Product and Brand Management*, *17*(5), 305-316.

Washburn, J. H., and Plank, R. E. (2002). Measuring brand equity: An evaluation of a consumer-based brand equity scale. *Journal of Marketing Theory and Practice*, *10*(1), 46-61.

Wind, J., Green, P. E., Shifflet, D., and Scarbrough, M. (1989). Courtyard by Marriott: designing a hotel facility with consumer based marketing models. *Interfaces*, *19*(1), 25-47.

Wiley, J. B. (1977). Selecting Pareto optimal subsets from multiattribute alternatives. *Advances in Consumer Research*, *5*, 171-174.

Wilkie, W. L. (1994). *Consumer behavior* (3rd ed.). New York: John Wiley & Sons, Inc.

Wind, Y. and Saaty, T. L. (1980). Marketing applications of the analytic hierarchy process. *Management Science*, *26*(7), 641-658.

Wittink, D. and Cattin, P. (1981). Alternative estimation methods for conjoint analysis – a Monte Carlo study. *Journal of Marketing Research*, *18*(1), 101-106.

Wittink, D. and Cattin, P. (1989). Commercial use of conjoint analysis: An update. *Journal of Marketing*, *53*(3), 91-96.

Wittink, D. R. and Walsh, J. W. (1988). *An interpretation of conjoint reliability findings*. Working Paper, Johnson Graduate School of Management, Cornell University, Ithaca, NY, USA.

Wittink, D. R. and Bergestuen, T. (2001). Forecasting with conjoint analysis. In J.S. Armstrong (Eds.), *Principles of forecasting*: *A handbook for researchers and practitioners* (pp.147-170). Boston and Dordrecht: Kluwer.

Won, T. Y. and Jung, S. W. (2001). *Hangul SPSS 10K statistical research analysis*. SPSS Academy, Seoul, Korea.

Wong, K. K., and Lam, C. (2001). Predicting hotel choice decisions and segmenting hotel consumers: a comparative assessment of a recent consumer based approach. *Journal of Travel and Tourism Marketing*, *11*(1), 17-33.

Won, D. and Bravo, G. A. (2009). Course design in sport management education: Addressing students' perspectives through conjoint methodology. *Journal of Hospitality, Leisure, Sport and Tourism Education*, *8*(2), 83-96.

Wright, P. (1975). Consumer choice strategies: simplifying vs. optimizing. *Journal of Marketing Research*, *12*, 60-67.

Yew Leh, F. C., and Lee, G. C. (2011). Customer-based brand equity: a literature review. *Journal of Arts Science & Commerce*, *2*(1), 33-42.

Yoo, B., and Donthu, N. (2001). Developing and validating a multidimensional consumer-based brand equity scale. *Journal of Business Research*, *52*(1), 1-14.

Yoo, B., Donthu, N. and Lee, S. (2000). An examination of selected marketing mix elements and brand equity. *Journal of the Academy of Marketing Science*, *28*(2), 195-211.

Yudhistira, T. (2002). *Product development methodology using fuzzy quality function deployment* (Master thesis). Industrial Engineering Department - ITB, Bandung, Bahasa, Indonesia.

Zahedi, F. (1986). The analytic hierarchy process-a survey of the method and its applications. *Interfaces*, *16*(July-August), 96-108.

Zhou, Y. and Jiang, J. (2011). The impacts of customer-based brand equity on revisit intentions: An empirical study on business and leisure traveler at five shanghai budget hotels. *Research Journal of International Studies*, *22*, 110-119.

Zhou, K. Z., Brown, J. R., and Dev, C. S. (2008). Market orientation, competitive advantage, and performance: a demand-based perspective. *Journal of Business Research*, *61*, 1-8.

Zeithaml, V. A. (1988). Consumer perceptions of price, quality and value: a means-end model and synthesis of evidence. *Journal of Marketing*, *52*, 2-22.

Zinkhan, G. M. and Locander, W. B. (1988). ESSCA: A multi-dimensional analysis tool for marketing research. *Journal of the Academy of Marketing Science*, *16*, 36-46.

APPENDIX A

THE ANALYTIC HIERARCHY PROCESS

Estimation of Relative Weights

The estimation of the relative weights of decision elements based on the eigenvalue method. After the relative weights of n elements have been identified for a level, they can be represented in an n x n matrix $A = (a_{ij})$, $(i, j=1,2,3,..,n)$.

$$A = \left[\begin{array}{ccccc} w_1/w_1 & w_1/w_2 & w_1/w_3 & \ldots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & w_2/w_3 & \ldots & w_2/w_n \\ w_3/w_1 & w_3/w_2 & w_3/w_3 & \ldots & w_3/w_n \\ . & . & . & . & . \\ . & . & . & . & . \\ w_n/w_1 & w_n/w_2 & w_n/w_3 & \ldots & w_n/w_n \end{array}\right] \begin{array}{l} 1 \\ 2 \\ 3 \\ . \\ . \\ n \end{array}$$

An estimate of the ratio between element *i* and *j* (w_i / w_j) is denoted as a_{ij}. The matrix *A* has all postive values and satisfies the reciprocal property $a_{ij} = 1/a_{ij}$, which is referred to as a postive reciprocal matrix (Saaty, 1990). The weights of elements $(w_1, w_{2,} w_3 ,.., w_n)$ can be estimated by solving the following matrix equation by determining the eigenvector associated with the maximum eighenvalue.

$$Aw = \begin{bmatrix} w_1/w_1 & w_1/w_2 & w_1/w_3 & \ldots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & w_2/w_3 & \ldots & w_2/w_n \\ w_3/w_1 & w_3/w_2 & w_3/w_3 & \ldots & w_3/w_n \\ . & . & . & . & . \\ . & . & . & . & . \\ w_n/w_1 & w_n/w_2 & w_n/w_3 & \ldots & w_n/w_n \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ . \\ . \\ w_n \end{bmatrix} = n \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ . \\ . \\ w_n \end{bmatrix} = nw$$

where n is the number of elements (maximum eigenvalue), and $w = (w_1, w_2 , w_3 \ldots. w_n)^T$ is the vector of actual relative weights (right eigenvector of matrix A).

It is assumed that A is known, but not w. Therefore, the relative weights of matrix A cannot be produced accurately. In this case, the above matrix equation for w can be solved by replacing n with $\lambda_{\max}$

$$Aw = \lambda_{\max} w$$

where $\lambda_{\max}$ is the largest eigenvalue of A. The eigenvectors of A corresponding to λ are the nonzero solutions of

$$(A - \lambda I)w = 0$$

where I is the identity matrix.

In order to solve the above equation, the eigenvalues of A need to be computed by solving the following equation:

$$\det(A - \lambda I) = |A - \lambda I| = 0$$

Estimation is the same for each level of a multi-level hierarchy.

Evaluation of Consistency

If a_{ik} is the exact estimate of the elements' weights, A is consistent because it satisfies the condition $a_{ik} = a_{ij} \times a_{jk}$ for all elements *i*, *j*, *k*. For a consistent positive reciprocal matrix, the largest eigenvalue $(\lambda_{\max})$ has a value of n, since the sum of the eigenvalues of a positive matrix is equal to the sum of the diagonal elements (trace of the matrix) (Forman & Selly, n.d.). Such perfect consistency, however, is not attainable in real world situations. If a_{ij} is an imperfect estimate of the ratio between *i* and *j*, then deviations between the consistent and the inconsistent matrix occur, causing changes in the eigenvalues (Saaty, 1996). $\lambda_{\max}$ is always greater than or equal to n. The closer $\lambda_{\max}$ is to n, the more consistent the judgments. Thus, the difference between λ_{max} and n is used as a measure of the consistency. Saaty (1996) has developed the consistency index (CI) as:

$$CI = \frac{(\lambda_{max} - n)}{(n-1)}$$

Inconsistency is inherent in the judgment process (Saaty, 1994). Therefore, perfect consistency is not expected and required by AHP. However, inconsistency may be considered a tolerable error in measurement only if it is small enough.

In order to determine an acceptable level of consistency, Saaty (1994) developed a random index (RI) table for matrices of the order from 1 to 10 (see Table 36).

Table 36

Random Inconsistency Indices (RI)

Size (order) of Matrix	1	2	3	4	5	6	7	8	9	10
Random Index (RI)	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.46	1.49

Source: Saaty (1980).

The value of CI is then compared to the value of RI. The ratio of CI to the average RI for the same order matrix yields the contingency rate (CR) which is used as a measure of the overall consistency of the matrix.

$$CR = \frac{CI}{RI}$$

In general, a CR value of 0.1 or less is considered acceptable (Saaty, 1996). If a CR value is over 0.1 then it is recommended that values assigned to the pair-wise comparison matrix be reevaluated to resolve the inconsistency in pair-wise comparisons.

For a multilevel hierarchy, an aggregate *CR* can be obtained by considering the *CI* and *RI* of each level. For instance, the *CI* of a three-level hierarchy *M* can be calculated as

$$M = CI^2 + W^2 \begin{bmatrix} CI_1^3 \\ CI_2^3 \\ \vdots \\ CI_m^3 \end{bmatrix}$$

where CI^2 is the *CI* value of the second level and CI_i^3 is the *CI* value for the *i*th element of the third level for *m* elements. Similarly, the *RI* for a three-level hierarchy $\bar{M}$ can be calculated as

$$\bar{M} = RI^2 + W^2 \begin{bmatrix} RI_1^3 \\ RI_2^3 \\ \vdots \\ RI_m^3 \end{bmatrix}$$

where RI^2 is the *RI* value for the number of elements in the second level and RI^3 is the *RI* value for the number of elements in the third level. Finally, the consistency ratio of the hierarchy (CRH) can be calculated as

$$CRH = \frac{M}{\bar{M}}$$

As with CR, consistency in judgment increase as CRH decreases.

Synthesis of Relative Weights

Relative weights of various levels obtained from the previous step are aggregated to produce a vector of composite weights that serve as ratings of decision alternatives (or selection choices) in achieving the general objective. In order to compute the priority for each alternative, the composite relative weight vector of elements at *k*th level with respect to that the first level may be computed from (Saaty 1980; Zahedi 1986),

$$C[1,k] = \prod_{i=2}^{k} B_i$$

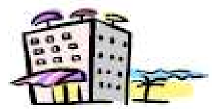
where C[1,*k*] is the vector of composite weights of elements at level k with respect to the element on Level 1, and B, is the n_{i-1} by n_i , matrix with rows consisting of estimated *w* vectors, n_i represents the number of elements at level *i*.

Consequently, the composite weighted priorities of the elements at a certain level are obtained by multiplying the priorities of attributes at that level by the priority of their corresponding attribute at the level above. This composite priority vector is then used to weight the priorities of elements at the level below and this process continues through to the lowest attribute level. Finally, the priority for each alternative is calculated by multiplying the composite priority of each attribute by the alternative's preference priority with respect to the attribute, and adding them.

APPENDIX B

SURVEY INSTRUMENT

"Hotel Preference Study"



Thank you for participating in this survey. Your answers will help us better understand your preferences when choosing a hotel. All of your answers are completely confidential. **There are total of *eight parts* in the survey for you to complete. The survey takes approximately 15~20 minutes. Some parts may look similar, but each part was designed to measure different aspects of your preferences. So, please respond to all the questions.**

Below are the definitions of the ***four factors of hotel selection*** and ***their*** **components** used in this survey. Please read the definitions before beginning.

Hotel Selection Factors	Components	Descriptions of the Components
Brand-Name Recognition: Customers' ability to recognize a brand's name	**Advertising**	The extent to which you are exposed to hotel ads
	Top-of-Mind brand	The name of a hotel that comes first to your mind
	Brand popularity	The extent to which you feel the brand is popular and used by others
	Brand familiarity	The extent to which you are familiar with the hotel brand
Hotel Brand Image: The total impression of a brand in an individual's mind	**Clean Image**	A very clean and orderly image
	Elegant atmosphere	A luxurious, stylish, prestigious, and suitable place for high class clientele
	Feels like home	A comfortable, quiet, and restful image
	Good value for money	The room rate and other fees for using the facilities at the hotel are reasonable to pay
Hotel Service Quality: Customers' opinion of service	**Error-free service**	Staff of the hotel has knowledge and confidence to answer guests.
	Prompt service	Service without delay (e.g. Promptness of check-in and check-out)
	Courteous service	Friendly, polite, and respectful service with neat, clean and appropriately groomed appearance
	Reliable service	Handling of complaints and problems sincerely
Hotel Brand Loyalty: The faithfulness of customers' to a particular brand	**Friends' recommendation**	Positive comments about the hotel brand from other people
	Frequent customer	The hotel brand would always be my first choice compared to other hotels (e.g. frequent guest program)
	Previous experience	The hotel brand that never disappoints me, guarantees satisfaction, and always meets my expectations

Part I. We are going to ask you to compare the components of each factor. <u>**Please check the box for which factor is more important.**</u>

For example, if you consider *Advertising* is somewhat important than *Top-of-Mind Brand* <u>with respect to Brand-Name Recognition in selecting a hotel</u>, you might respond like this: (See example below)

Component A	Extremely Important	Much More Important	More Important	Somewhat Important	Equally Important	Somewhat Important	More Important	Much More Important	Extremely Important	Component B
Advertising	☐	☐	☐	☒	☐	☐	☐	☐	☐	Top-of-Mind Brand

Which of the following components of <u>**Brand-Name Recognition**</u> is *more important*?

Component A	Extremely Important	Much More Important	More Important	Somewhat Important	Equally Important	Somewhat Important	More Important	Much More Important	Extremely Important	Component B
Advertising	☐	☐	☐	☐	☐	☐	☐	☐	☐	Top-of-Mind Brand
Advertising	☐	☐	☐	☐	☐	☐	☐	☐	☐	Brand Popularity
Advertising	☐	☐	☐	☐	☐	☐	☐	☐	☐	Brand Familiarity
Top-of-Mind Brand	☐	☐	☐	☐	☐	☐	☐	☐	☐	Brand Popularity
Top-of-Mind Brand	☐	☐	☐	☐	☐	☐	☐	☐	☐	Brand Familiarity
Brand Popularity	☐	☐	☐	☐	☐	☐	☐	☐	☐	Brand Familiarity

Which of the components of <u>**Hotel Brand Image**</u> is *more important*?

Component A	Extremely Important	Much More Important	More Important	Somewhat Important	Equally Important	Somewhat Important	More Important	Much More Important	Extremely Important	Component B
Clean Image	☐	☐	☐	☐	☐	☐	☐	☐	☐	Elegant
Clean Image	☐	☐	☐	☐	☐	☐	☐	☐	☐	Feels Like Home
Clean Image	☐	☐	☐	☐	☐	☐	☐	☐	☐	Good value
Elegant	☐	☐	☐	☐	☐	☐	☐	☐	☐	Feels Like Home
Elegant	☐	☐	☐	☐	☐	☐	☐	☐	☐	Good value
Feels Like Home	☐	☐	☐	☐	☐	☐	☐	☐	☐	Good value

Which of the components of **<u>Hotel Service Quality</u>** is *more important*?

Component A	Extremely Important	Much More Important	More Important	Somewhat Important	Equally Important	Somewhat Important	More Important	Much More Important	Extremely Important	Component B
Error-Free Service	☐	☐	☐	☐	☐	☐	☐	☐	☐	**Prompt Service**
Error-Free Service	☐	☐	☐	☐	☐	☐	☐	☐	☐	**Courteous Service**
Error-Free Service	☐	☐	☐	☐	☐	☐	☐	☐	☐	**Reliable Service**
Prompt Service	☐	☐	☐	☐	☐	☐	☐	☐	☐	**Courteous Service**
Prompt Service	☐	☐	☐	☐	☐	☐	☐	☐	☐	**Reliable Service**
Courteous Service	☐	☐	☐	☐	☐	☐	☐	☐	☐	**Reliable Service**

Which of the components of **<u>Hotel Brand Loyalty</u>** is *more important*?

Component A	Extremely Important	Much More Important	More Important	Somewhat Important	Equally Important	Somewhat Important	More Important	Much More Important	Extremely Important	Component B
Friends' Recommendation	☐	☐	☐	☐	☐	☐	☐	☐	☐	**Frequent Customer**
Friends' Recommendation	☐	☐	☐	☐	☐	☐	☐	☐	☐	**Previous Experience**
Frequent Customer	☐	☐	☐	☐	☐	☐	☐	☐	☐	**Previous Experience**

Part II. We will ask you to compare two factors regarding a hotel choice. Please check the box for which factor is more important.

Which of the two factors is more important in selecting a hotel

Factor A	Extremely Important	Much More Important	More Important	Somewhat Important	Equally Important	Somewhat Important	More Important	Much More Important	Extremely Important	Factor B
Brand-Name Recognition	☐	☐	☐	☐	☐	☐	☐	☐	☐	Hotel Brand Image
Brand-Name Recognition	☐	☐	☐	☐	☐	☐	☐	☐	☐	Hotel Service Quality
Brand-Name Recognition	☐	☐	☐	☐	☐	☐	☐	☐	☐	Hotel Brand Loyalty
Hotel Brand Image	☐	☐	☐	☐	☐	☐	☐	☐	☐	Hotel Service Quality
Hotel Brand Image	☐	☐	☐	☐	☐	☐	☐	☐	☐	Hotel Brand Loyalty
Hotell Service Quality	☐	☐	☐	☐	☐	☐	☐	☐	☐	Hotel Brand Loyalty

Part III. Please rate this survey by *checking the box* that best fits your opinion <u>when you compared the four hotel selection factors and their components (Part I & II).</u>
(1 being the *lowest* and 9 being the *highest*).

Satisfaction:										
How much did you enjoy this survey?										
Boring	1☐	2☐	3☐	4☐	5☐	6☐	7☐	8☐	9☐	Interesting
Information content:										
How realistic is this form of questioning in choosing a hotel?										
Unrealistic	1☐	2☐	3☐	4☐	5☐	6☐	7☐	8☐	9☐	Realistic
Decision Making:										
How difficult was it to respond to the questions asked?										
Difficult	1☐	2☐	3☐	4☐	5☐	6☐	7☐	8☐	9☐	Easy

Part IV. (a) There are 16 cards and each card has a different combination of four components of the hotel selection factors. Please *rank* the cards *in order* of preference in choosing a hotel (1 - Most Preferred, 16 - Least Preferred).

Hotel Choice 1	Hotel Choice 2	Hotel Choice 3	Hotel Choice 4	Hotel Choice 5	Hotel Choice 6	Hotel Choice 7	Hotel Choice 8
Brand Familiarity	Advertising	Advertising	Brand Popularity	Top-of-Mind Brand	Brand Popularity	Brand Familiarity	Brand Familiarity
Elegant	Good value	Elegant	Good value	Clean Image	Elegant	Good value	Feels Like Home
Courteous Service	Reliable Service	Prompt Service	Prompt Service	Prompt Service	Reliable Service	Error-Free Service	Prompt Service
Previous Experience	Previous Experience	Friends' Recommendation	Frequent Customer	Previous Experience	Friends' Recommendation	Friends' Recommendation	Friends' Recommendation

Hotel Choice 9	Hotel Choice 10	Hotel Choice 11	Hotel Choice 12	Hotel Choice 13	Hotel Choice 14	Hotel Choice 15	Hotel Choice 16
Advertising	Brand Popularity	Top-of-Mind Brand	Advertising	Top-of-Mind Brand	Top-of-Mind Brand	Brand Popularity	Brand Familiarity
Clean Image	Clean Image	Feels Like Home	Feels Like Home	Elegant	Good value	Feels Like Home	Clean Image
Error-Free Service	Courteous Service	Reliable Service	Courteous Service	Error Free Service	Courteous Service	Error Free Service	Reliable Service
Friends' Recommendation	Friends' Recommendation	Friends' Recommendation	Frequent Customer	Frequent Customer	Friends' Recommendation	Previous Experience	Frequent Customer

(Note: Please make sure you have ranked all 16 cards *without duplicate card number)*

Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Card #																

(b) Please *rank* four cards below (1- the *most preferred* to 4 - the *least preferred*).

Hotel Choice 17	Hotel Choice 18	Hotel Choice 19	Hotel Choice 20
Top-of-Mind Brand	Brand Popularity	Brand Popularity	Advertising
Clean Image	Clean Image	Elegant	Feels Like Home
Error-Free Service	Courteous Service	Courteous Service	Error-Free Service
Friends' Recommendation	Previous Experience	Friends' Recommendation	Friends' Recommendation

Rank	1	2	3	4
Card #				

Part V. Please rate this survey by *checking the box* that best fits your opinion <u>when you ranked the 16 cards in order of preference (Part IV).</u>
(1 being the *lowest* and 9 being the *highest*).

Satisfaction:										
How much did you enjoy this survey?										
Boring	1☐	2☐	3☐	4☐	5☐	6☐	7☐	8☐	9☐	**Interesting**
Information content:										
How realistic is this form of questioning in choosing a hotel?										
Unrealistic	1☐	2☐	3☐	4☐	5☐	6☐	7☐	8☐	9☐	**Realistic**
Decision Making:										
How difficult was it to respond to the questions asked?										
Difficult	1☐	2☐	3☐	4☐	5☐	6☐	7☐	8☐	9☐	**Easy**

Part VI. There are **10 additional tests** and each test has **another set of different combinations** of the four components of **hotel selection factors**. Please check ***one box you mostly prefer*** when choosing a hotel.

For example, if you prefer card #4 the most, you would mark an "X" in the box: (See below)

<EXAMPLE>			
Card 1 ☐	Card 2 ☐	Card 3 ☐	Card 4 ☒
Brand popularity	Advertising	Advertising	Advertising
Clean image	Elegant	Elegant	Clean image
Prompt service	Error-free service	Reliable service	Courteous service
Previous experience	Friends' recommendation	Frequent customer	Previous experience

TEST 1			
Card 1 ☐	Card 2 ☐	Card 3 ☐	Card 4 ☐
Brand familiarity	Advertising	Top-of-mind brand	Brand familiarity
Feels like home	Feels like home	Feels like home	Elegant
Error-free service	Courteous service	Prompt service	Courteous service
Previous experience	Friends' recommendation	Frequent customer	Frequent customer

TEST 2			
Card 1 ☐	Card 2 ☐	Card 3 ☐	Card 4 ☐
Brand popularity	Brand familiarity	Top-of-mind brand	Advertising
Good value	Feels like home	Feels like home	Clean image
Error-free service	Prompt service	Courteous service	Prompt service
Frequent customer	Previous experience	Frequent customer	Friends' recommendation

TEST 3			
Card 1 ☐	Card 2 ☐	Card 3 ☐	Card 4 ☐
Brand familiarity	Brand familiarity	Advertising	Brand familiarity
Good value	Clean image	Elegant	Feels like home
Prompt service	Prompt service	Reliable service	Error-free service
Previous experience	Previous experience	Friends' recommendation	Frequent customer

TEST 4			
Card 1 ☐	Card 2 ☐	Card 3 ☐	Card 4 ☐
Brand popularity	Advertising	Advertising	Advertising
Clean image	Elegant	Elegant	Clean image
Prompt service	Error-free service	Reliable service	Courteous service
Previous experience	Friends' recommendation	Frequent customer	Previous experience

TEST 5			
Card 1 ☐	Card 2 ☐	Card 3 ☐	Card 4 ☐
Top-of-mind brand	Brand familiarity	Advertising	Brand popularity
Clean image	Clean image	Clean image	Feels like home
Prompt service	Prompt service	Reliable service	Error-free service
Previous experience	Friends recommendation	Friends' recommendation	Friends' recommendation

TEST 6			
Card 1 ☐	Card 2 ☐	Card 3 ☐	Card 4 ☐
Brand popularity	Advertising	Brand popularity	Top-of-mind brand
Good value	Good value	Good value	Good value
Courteous service	Reliable service	Courteous service	Error-free service
Friends' recommendation	Friends' recommendation	Previous experience	Previous experience

TEST 7			
Card 1 ☐	Card 2 ☐	Card 3 ☐	Card 4 ☐
Brand popularity	Top-of-mind brand	Brand familiarity	Top-of-mind brand
Clean image	Feels like home	Elegant	Elegant
Prompt service	Reliable service	Courteous service	Prompt service
Frequent customer	Friends' recommendation	Frequent customer	Previous experience

TEST 8			
Card 1 ☐	Card 2 ☐	Card 3 ☐	Card 4 ☐
Brand familiarity	Advertising	Advertising	Advertising
Elegant	Feels like home	Clean image	Feels like home
Courteous service	Prompt service	Reliable service	Prompt service
Friends' recommendation	Friends' recommendation	Friends' recommendation	Previous experience

TEST 9			
Card 1 ☐	Card 2 ☐	Card 3 ☐	Card 4 ☐
Brand popularity	Brand familiarity	Brand familiarity	Top-of-mind brand
Feels like home	Elegant	Elegant	Elegant
Error-free service	Courteous service	Error-free service	Courteous service
Friends' recommendation	Frequent customer	Previous experience	Friends' recommendation

TEST 10			
Card 1 ☐	Card 2 ☐	Card 3 ☐	Card 4 ☐
Top-of-mind brand	Brand familiarity	Advertising	Advertising
Feels like home	Good value	Feels like home	Clean image
Courteous service	Error-free service	Error-free service	Reliable service
Frequent customer	Previous experience	Previous experience	Friends' recommendation

Part VII. Hotel Experience

1) Which category of hotels have you **most frequently stayed** in during the past year? Please pick **one.**

- Budget/Economy (e.g. Motel 6, Quality Inn, La Quinta, etc.) (   )
- Mid-price (e.g., Hampton Inn, Holiday Inn, etc.) (   )
- Upscale (e.g., Hyatt, JW Marriott, Hilton, etc.) (   )
- Luxury (e.g., Four Seasons, Ritz Carlton, etc.) (   )

2) On average how many nights have you stayed at the hotel during the past year?

1-3 days ( ) 4-7 days ( ) 8-15 days ( ) 16+ days (   )

3) What was the purpose of your hotel stay?

Vacation (   ) Business (   ) Visiting Relatives, Friends (   ) other (specify) (   )

4) Please check <u>one box</u> below for the level of agreement regarding <u>how interested</u> you are in hotels (**1 being the *lowest* and 8 being the *highest***).

How often do you use hotels?

Rarely 1☐ 2☐ 3☐ 4☐ 5☐ 6☐ 7☐ 8☐ Often

How much are you involved with hotels in your life?

Low 1☐ 2☐ 3☐ 4☐ 5☐ 6☐ 7☐ 8☐ High

How much are you interested in hotels, relative to others?

Uninterested 1☐ 2☐ 3☐ 4☐ 5☐ 6☐ 7☐ 8☐ Interested

5) Please check <u>one box</u> below for the level of agreement regarding <u>how knowledgeable</u> you are about hotels.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I feel very knowledgeable about hotels	☐	☐	☐	☐	☐
I can give people advice about different brands of hotels	☐	☐	☐	☐	☐
I only need to gather very little information in order to make a wise decision about hotel choices	☐	☐	☐	☐	☐
I feel very confident about my ability to tell the difference in quality between different brands of hotels	☐	☐	☐	☐	☐

Part VIII. Demographic Information

1) Please select your gender.

Male ( ) Female ( )

2) What is your age? ______________

3) What is your marital status?

Single ( ) Married ( ) other (specify) ( )

4) What is your status?

Freshmen ( ) sophomore ( ) Junior ( ) Senior ( ) Graduate Student ( )

other (specify) ( )

5) What is your student enrollment status?

Part time student ( ) Full time student ( ) other (specify) ( )

6) What is your major? ( )

Thank you for your participation!

APPENDIX C

IRB EXEMPTION

Knowledge Enterprise Development

Office of Research Integrity and Assurance

To:	Timothy Tyrrell UCENT
From:	Mark Roosa, Chair Soc Beh IRB
Date:	10/18/2012
Committee Action:	**Exemption Granted**
IRB Action Date:	10/18/2012
IRB Protocol #:	1210008381
Study Title:	An empirical comparison between two multiattribute decision techniques for measuring brand equity

The above-referenced protocol is considered exempt after review by the Institutional Review Board pursuant to Federal regulations, 45 CFR Part 46.101(b)(2) .

This part of the federal regulations requires that the information be recorded by investigators in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects. It is necessary that the information obtained not be such that if disclosed outside the research, it could reasonably place the subjects at risk of criminal or civil liability, or be damaging to the subjects' financial standing, employability, or reputation.

You should retain a copy of this letter for your records.