Customer Value Measurement

MSI Research Award #4-1668

Working Paper

Sara Leroi-Werelds

Sandra Streukens

Sara Leroi-Werelds is a Ph.D. candidate of the Research Foundation Flanders (FWO Vlaanderen) at the Department of Marketing and Strategy at Hasselt University, Belgium. P: Hasselt University, Agoralaan-Building D, BE-3590 Diepenbeek, Belgium. T: + 32 11 26 86 21; E: sara.leroiwerelds@uhasselt.be.

Sandra Streukens is assistant professor at the Department of Marketing and Strategy at Hasselt University, Belgium. P: Hasselt University, Agoralaan-Building D, BE-3590 Diepenbeek, Belgium. T: + 32 11 26 86 28; E: sandra.streukens@uhasselt.be.

The authors thank the members of the MSI Research Review Committee for their constructive comments and insightful advice.

"Making customer value strategies work

begins with an actionable understanding of the concept itself."

- Robert Woodruff -

Report Summary

Although interest regarding customer value has persisted for many years, considerable divergence of opinion exists on how to most adequately conceptualize[1] customer value. The most commonly used value measurement methods include those of Dodds, Monroe, and Grewal (1991), Gale (1994), Holbrook (1999), and Woodruff and Gardial (1996). Though each of these methods intends to capture perceived customer value, substantial differences among them exist in terms of dimensionality (one-dimensional vs. multi-dimensional), nature of costs and benefits (attribute-based vs. consequence-based), and the scope of measurement (relative to competition or not).

Whereas each of the four value conceptualizations has its merits, little is known about which approach is best capable of predicting key marketing variables such as customer satisfaction and loyalty. Furthermore, it remains unclear whether possible performance differences among methods depend on contextual factors such as involvement and type of product. This article addresses these two issues by means of an empirical study using customer data from four different settings (total n=3,360).

[1] In this study, we use the terms conceptualization, measurement method and operationalization interchangeably.

The authors compared the performance of the four measurement methods mentioned above and conclude that customer value should be operationalized in a multi-dimensional, consequence-based way. Thus, customer value is best assessed by means of the methods of Holbrook (1999) or Woodruff and Gardial (1996). Although both methods perform equally well for feel offerings, this is not the case for think offerings. Holbrook's (1999) approach is preferred in case of low involvement think offerings, whereas the method of Woodruff and Gardial (1996) is the best choice for high involvement think offerings. In terms of the three underlying conceptual differences (see above) of the four methods under study, optimal value assessment should rely on a multi-dimensional approach that assesses the costs and benefits at the consequence level. Furthermore, a value measurement approach that explicitly takes into account perceptions of both the focal offering and competing offerings (i.e., Gale, 1994) does not yield additional explanatory power. Although differences in relative performance among the studied approaches exist, involvement and type of offering are not capable of systematically explaining these differences.

This study provides insight in the predictive ability of the four dominant customer value conceptualizations proposed in the academic marketing literature and offers clear directions for choosing the most appropriate value measurement method. For marketing academics and practitioners alike, these findings are of importance when studying relationships in a nomological network that includes perceived customer value. Overall, we hope our work contributes to closing the gap between customer value management theory and practice, signaled by Woodruff (1997), in designing effective marketing strategies.

Introduction

Customer value has been of continuing interest to marketing researchers and practitioners alike. Moreover, it has been recognized as one of the most significant factors in the success of organizations (Butz and Goodstein, 1996; Slater, 1997; Wang, Lo, Chi, and Yang, 2004). In line with Zeithaml's (1988, p. 4) definition that "perceived value is the consumer's overall assessment of the utility of a product based on perceptions of what is received and what is given", there has been a general consensus that customer value involves a trade-off between benefits and costs (e.g., Chen and Dubinsky, 2003; Flint, Woodruff, and Gardial, 2002; Rintamäki, Kuusela, and Mitronen, 2007; Ruiz, Gremler, Washburn, and Carrión, 2008; Slater and Narver, 2000; Ulaga and Chacour, 2001).

Despite the agreement regarding the definition and importance of value, considerable divergence of opinion exists among researchers on how to most adequately conceptualize customer value. This observation is very well illustrated by the great variety of measurement methods forwarded in the literature such as the work of Dodds, Monroe, and Grewal (1991), Gale (1994), Holbrook (1999), and Woodruff and Gardial (1996). Although each measurement method claims to be capable of assessing customer value adequately, no empirical work exists on the relative performance of the different methods in predicting key marketing variables such as customer satisfaction and loyalty, which are leading indicators of a firm's financial performance. Furthermore, it is unknown whether this predictive ability of different value conceptualizations is influenced by contextual factors such as involvement level and type of offering.

Accordingly, the following two research objectives guide our study. First, we aim to assess and compare the performance of the four commonly used customer value measurement methods mentioned above (i.e., Dodds et al., 1991; Gale, 1994; Holbrook, 1999; Woodruff and Gardial, 1996) with regard to their predictive ability of customer satisfaction, repurchase intentions and word-of-mouth in different settings. Second, we examine whether the relative performance of these methods (i.e., the difference between the predictive ability of two methods) systematically varies as a consequence of contextual factors such as type of product (feel versus think products) and level of customer involvement (high versus low involvement).

The importance of our research is illustrated by the fact that "remarkably few firms have the knowledge and capability to actually assess value in practice" (Anderson and Narus, 2004, p. 3). Empirical evidence concerning how to optimally measure perceived customer value represents a necessary condition for realizing the full potential of customer value management. As such, our research offers an attempt to bridge the gap between theory and practice that Woodruff (1997) signals in the area of customer value management.

We organize the rest of this article as follows. First, we present the four commonly used methods for measuring customer value that take central stage in this study and discuss their (dis)similarities. Second, we discuss the data collection procedures. Third, we describe the analytical approaches and empirical results pertaining to our two interrelated research objectives. It should be noted that based on the analytical results pertaining to our first research objective (i.e., the assessment and comparison of the performance of the four commonly used customer value measurement methods with regard to their predictive ability of customer satisfaction,

repurchase intentions and word-of-mouth in different settings), we proceed by proposing and analyzing a series of hypotheses aimed at understanding the differences in predictive ability across the different value measurement methods (i.e., research objective 2). We conclude this paper by summarizing our conclusions, discussing our limitations and making suggestions for further research.

Literature Review

Outcome Variables

Prior research has stated that customer value is an important antecedent of satisfaction and loyalty (Bolton and Drew, 1991; Cronin, Brady, and Hult, 2000; Lai, Griffin, and Babin, 2009; Zeithaml, Berry, and Parasuraman, 1996). In turn, several studies (e.g., Anderson, Fornell, and Lehmann, 1994; Hallowell, 1996; Kamakura, Mittal, de Rosa, and Mazzon, 2002; Loveman, 1998) have indicated that customer satisfaction and customer loyalty are prime determinants of the long-term profitability of the firm.

In line with the literature on the relationship between customer evaluative judgments and financial performance (Anderson et al., 1994; Oliver, 1997), we define customer satisfaction as the cumulative evaluation that is based on all experiences with the supplier's offering over time (Anderson et al., 1994). Loyalty, on the other hand, is approached from a behavioral intentions point-of-view (Cronin et al., 2000; Zeithaml et al., 1996) and includes the intention to repurchase and the willingness to recommend the supplier's offering to others (Lai et al., 2009; Wirtz and Lee, 2003; Zeithaml et al., 1996).

Domains of Difference among Customer Value Measurement Methods

As mentioned before, the value measurement methods of Dodds et al. (1991), Gale (1994), Holbrook (1999), and Woodruff and Gardial (1996) take central stage in our study. To be able to effectively compare and contrast these four value measurement methods, we start with a general outline of how value measurement methods can differ. These so-called *domains of difference* are based on the existing literature about customer value (Woodruff, 1997; Sánchez-Fernández and Iniesta-Bonilllo, 2007) as well as on the thorough evaluation of the four central measurement methods. Below we describe the three domains of difference and after that we will describe the different customer value measurement methods in detail and explain how they relate to these domains of difference.

First of all, we can classify the value measurement methods as one-dimensional or multi-dimensional (Ruiz et al., 2008; Sánchez-Fernández, Iniesta-Bonillo, and Holbrook, 2009). According to the one-dimensional view, customer value is "a single overall concept that can be measured by a self-reported item (or set of items) that evaluates the consumer's perception of value" (Sánchez-Fernández and Iniesta-Bonillo, 2007, p. 430). Although an often mentioned advantage of the one-dimensional measurement method is its simplicity and ease of implementation (Lin, Sher, and Shih, 2005), many researchers (Ruiz et al., 2008; Sweeney and Soutar, 2001) share the notion that the construct of customer value is too complex to be captured by a one-dimensional measurement method. As a response to this critique on the one-dimensional approach, so-called multi-dimensional approaches have been put forward. The basic premise

underlying these multi-dimensional approaches is that customer value consists of several interrelated components or dimensions (Sánchez-Fernández and Iniesta-Bonillo, 2007).

Second, the nature of the benefits and costs included in the model differs across the value conceptualizations. Following Gutman's (1982) means-end chain model, these can be measured at the attribute and/or consequence level. Attributes are concrete characteristics or features of a product or service such as size, shape or on-time delivery. Consequences are more subjective experiences resulting from product use such as a reduction in lead time or a pleasant experience (Gutman, 1982, 1997; Woodruff and Gardial, 1996).

A third and last difference relates to whether or not customer value perceptions are measured relative to the competition.

Dodds, Monroe and Grewal's (1991) Approach

Dodds et al. (1991) focus only on a very narrow aspect of the trade-off underlying customer value as they define perceived value as "a cognitive tradeoff between perceived quality and sacrifice" (Dodds et al., 1991, p. 316). On the basis of this definition, they measure customer value by asking respondents five questions concerning the overall value of the product or service. The approach of Dodds et al. (1991) is considered one-dimensional as the value construct is not divided into different dimensions that tap on specific elements of value. In terms of the second dimension, the nature of the costs and benefits, a distinction between attributes and consequences does not apply as the items of the Dodds et al. (1991) method focus only on customer value at a

very general level. Finally, Dodds et al. (1991) do not measure customer value in relation to the competition.

Empirical studies using the measurement scale of Dodds et al. (1991) include Teas and Agarwal (2000), Agarwal and Teas (2001), Baker, Parasuraman, Grewal, and Voss (2002), Chen and Dubinsky (2003), and Caruana and Fenech (2005).

Gale's (1994) Customer Value Analysis

Compared to the other methods in this study, a distinct feature of Gale's (1994) approach is that it explicitly takes into account the customers' quality and price judgments of an organization relative to those of relevant competitors. The basic premise underlying Gale's (1994) customer value measurement approach, or Customer Value Analysis as he calls it, is that customer value equals the difference between a weighted quality score (termed *market-perceived quality*) and a weighted price score (termed *market-perceived price*).

The basis for constructing a market-perceived quality (price) profile entails asking customers to evaluate relevant quality (price) attributes in terms of performance and importance. These attributes are typically elicited from in-depth or focus group interviews and should cover all relevant aspects related to perceived quality (price). To assess relative customer evaluative judgments, performance evaluations are asked for both the firm's offering and competitors' offerings.

In terms of the nature of the benefits and costs assessed by the measurement method, Gale (1994) stays at the attribute level. Furthermore, as Gale (1994) explicitly distinguishes among various different elements of benefits and costs, his measurement method can be considered multi-dimensional.

Authors following Gale's (1994) Customer Value Analysis include Laitamäki and Korduplesk i (1997), Lam, Shankar, Erramilli, and Murthy (2004) and Setijono and Dahlgaard (2007).

Woodruff and Gardial's (1996) Customer Value Hierarchy

Woodruff and Gardial (1996) presented the Customer Value Hierarchy to understand customer value. The work of Woodruff and Gardial (1996) differed from previous conceptualizations of customer value by suggesting that value creation takes place at the consequence level rather than at the more narrowly defined attribute level. More specifically, they state that value is the result of "the trade-off between the positive and negative consequences of product use as perceived by the customer" (Woodruff and Gardial, 1996, p. 57). According to Woodruff and Gardial (1996), this shift in focus from attributes to consequences will result in value creation that leads to a more pronounced strategic sustainable competitive advantage.

Similar to Gale's (1994) Customer Value Analysis, Woodruff and Gardial (1996) explicitly discern among different elements of the benefits and sacrifices they assess. Consequently, Woodruff and Gardial's (1996) Customer Value Hierarchy can be classified as a multi-dimensional approach. It should be noted that Woodruff and Gardial's (1996) approach does not measure perceived customer value relative to the competition.

Only a few authors follow the approach developed by Woodruff and Gardial (1996), including Flint et al. (2002) and Overby, Gardial, and Woodruff (2004).

Holbrook's (1999) Customer Value Typology

Holbrook (1999) developed a framework, which reflects three underlying dimensions:

- Extrinsic value versus Intrinsic value (an offering appreciated for its functional, utilitarian ability to achieve something vs. an offering appreciated as an end-in-itself)
- Self-oriented value versus Other-oriented value (an offering prized for the effect it has on one self vs. the effect it has on others)
- Active value versus Reactive value (the customer acts on the object vs. the object acts on the customer)

Each of the three dimensions is treated as a dichotomy, though they should be envisioned as a continuum of possibilities running from one extreme to the other with gradations in between (Holbrook, 1999). Using the three dimensions outlined above, Holbrook (1999) developed a matrix representing eight types of customer value: efficiency, excellence, status, esteem, play, aesthetics, ethics, and spirituality. This is also called Holbrook's Typology of Customer Value.

This typology involves the co-existence of different types of customer value. This means that a consumption experience entails many or even all of the value types identified in the typology (Holbrook, 1999). Some of the value types in Holbrook's framework are related in such a way that it is extremely difficult to operationalize them separately. For that reason, some authors

suggest combining these value types in an overarching category. Especially the demarcation between status and esteem can be problematic (Holbrook, 1999) because "the active nature of status and the reactive nature of esteem tend to blur together in ways that render the two hard to distinguish" (Holbrook, 1999, p. 188). Therefore, we follow previous research by combining status and esteem in an overarching category called *social value* (Bourdeau, Chebat, and Couturier, 2002; Gallarza and Saura, 2006; Sánchez-Fernández et al., 2009; Sweeney and Soutar, 2001). Social value arises when one's own consumption behavior serves as a means to influence the responses of others (Holbrook, 2006). Similarly, ethics and spirituality can be combined under the heading of *altruistic value*, as they have in common that "both lie outside the sphere of ordinary marketplace exchanges" (Sánchez-Fernández et al., 2009, p. 101). One can define altruistic value as "a concern for how my own consumption behavior affects others where this experience is viewed as a self-justifying end-in-itself" (Holbrook, 2006, p. 716).

In his earlier work, Holbrook did not consider the cost side of the value construct. Holbrook (1999) admitted that his treatment of the customer value concept implicitly regards value as a cost-free benefit, which means that only the benefit side and not the sacrifice side is included in his approach. One can circumvent this problem by considering the typology as positive outcomes (benefits) and comparing it with negative value inputs (e.g., price, risk, time and effort; Gallarza and Saura, 2006; Oliver, 1997). In a recent study of Holbrook (Sánchez-Fernández et al., 2009) these negative value inputs were considered part of the typology by including monetary cost, time, and effort in efficiency, because efficiency includes the get-versus-give aspects of consumption (Sánchez-Fernández and Iniesta-Bonillo, 2007).

In line with the different dimensions specified by Holbrook (1999), this conceptualization of customer value can be considered a multi-dimensional measurement approach. Regarding the nature of the benefits and sacrifices measured, Holbrook's (1999) method involves both the attribute and the consequence level (Overby et al., 2004; Woodruff, 1997). Finally, in measuring customer value, Holbrook (1999) does not regard performance relative to the competition.

Empirical studies using the Holbrook (1999) framework include Gallarza and Saura (2006) and Sánchez-Fernández et al. (2009).

Comparing and Contrasting the Different Value Measurement Methods

Table 1 summarizes how the customer value measurement methods used in this study relate to the three domains of difference that can be used to classify the various approaches.

[INSERT TABLE 1 ABOUT HERE PLEASE]

Research Design

Settings

In line with our second research objective, it is necessary to take precautions that allow us to empirically assess possible differences in predictive ability of the value measurement methods as a function of contextual factors. Therefore, we collected data across four different settings. The choice of these settings was guided by the Foote, Cone and Belding (FCB) grid (Vaughn, 1980), which classifies customers' purchase decisions on two dimensions: involvement and type of offering. The rationale underlying our choice for the FCB-grid is as follows. Given the different

conceptual perspectives underlying the value measurement methods (see also Table 1 above), we expect that the relative ability of the methods to predict the outcome variables under study depends on the offering's characteristics which correspond to the dimensions of the grid. The products selected as research contexts for our study are soft drinks (low involvement, feel), toothpaste (low involvement, think), day cream (high involvement, feel) and DVD players (high involvement, think). More information regarding these two dimensions as well as their hypothesized impact on the relative performance of the value conceptualizations will be presented in the section where we provide information regarding our second research objective.

Sampling

To enhance the external validity of our research, data were collected using one of the largest marketing research panels in Belgium. Although the respondents were self-selected, they were disqualified if they did not use the product, did not buy the product, or did not pay for the product themselves. Consequently, each respondent was asked to evaluate the soft drink he/she usually drinks, the day cream he/she currently uses, the toothpaste he/she currently uses or the DVD player one currently uses. As explained below, data were obtained from 16 independent samples (i.e., 4 settings * 4 value conceptualizations) each having an effective sample size of 210 respondents.

Questionnaire Design

We opted to construct 16 different questionnaires (i.e., collected from 16 different [sub]samples), so that each questionnaire assesses one value measurement method in one setting. The rationale

behind this choice is threefold. First, we tried to keep the amount of time and effort asked from the respondents as low as possible. Second, we tried to avoid carry-over effects among the different value measurement approaches. Finally, restricting ourselves to between-subject variance allows us to draw statistically valid conclusions among all possible combinations of value measurement approaches.

All questionnaires were identical in terms of the measurement instruments for customer satisfaction, customer loyalty, and the manipulation checks (i.e., measurement of involvement and type of offering). What differed across the questionnaires was the value measurement method employed which, furthermore, needed to be adapted to the particular setting. Starting with the operationalization of the different value conceptualizations, we explain our questionnaire design below. All individual items are listed in Appendices A and B and are evaluated on 9-point Likert scales unless indicated otherwise.

Dodds et al. 's (1991) approach. To assess the performance of the measurement approach suggested by Dodds et al. (1991), we used the five items suggested by the original authors.

Gale's (1994) approach. To generate items for Gale's (1994) Customer Value Analysis, we carried out in-depth interviews using the laddering technique (cf. Woodruff and Gardial, 1996) and listed the attributes people found most important in the four different settings (see Appendix A). In total 28 laddering interviews using respondents that had experience with the product under investigation were conducted (DVD player n=7; day cream n = 6; soft drinks n = 7; toothpaste n = 8). The number of respondents in each setting was determined using the procedure suggested

by Strauss and Corbin (1998), which suggests continuing with laddering interviews until theoretical saturation (i.e., additional interviews do not lead to new information) occurs. Since Gale's method implies a relative approach for measuring customer value, we asked respondents to evaluate the product attributes relative to the competition on a 9-point scale with labels *XYZ is much better* to *XYZ is much worse* (Babakus, Bienstock, and Van Scotter, 2004). In line with Gale's (1994) measurement method, we also needed an importance weight for each attribute. However, because the number of attributes was considerably large, point allocation – as proposed by Gale – was not an option. According to Louviere and Islam (2008), there are two general ways to measure importance: directly or indirectly. These authors compared different ways for measuring importance weights and found high agreement within direct or indirect methods, but less agreement between direct and indirect methods. Since Gale (1994) uses point allocation – and thus directly measures importance weights – a direct approach is required. Therefore, we used the direct rating approach by asking respondent to rate each attribute on a Likert scale anchored at 1 (*very unimportant*) and 9 (*very important*). Furthermore, Bottomley, Doyle and Green (2000) showed that the weights elicited by direct rating are more reliable than those elicited by point allocation in a test-retest situation.

Woodruff's (1996) approach. The generation of items for the measurement method proposed by Woodruff and Gardial (1996) was completely based on the results of the laddering interviews mentioned above.

Holbrook's (1999) approach. For the measurement of Holbrook's (1999) Value Typology, we used existing scales where possible (Excellence: Oliver, 1997 , Efficiency: Ruiz et al., 2008 ,

Social value: Sweeney and Soutar, 2001, Play: Petrick, 2002) and adapted them to the particular settings at hand by means of the laddering interviews described above. An existing scale for aesthetic value was not available, so we used the results of the laddering interviews to generate items. Altruistic value was not mentioned in the interviews, so we did not take this value type into account in our empirical study (Gallarza and Saura, 2006).

Outcome variables. Customer satisfaction was measured using Anderson et al.'s (1994) single-item scale to assess cumulative satisfaction with a market offering. In line with Wirtz and Lee (2003), a 11-point scale was used for this item. Repurchase intentions and word-of-mouth were measured as a proxy for customer loyalty using the scale developed by Zeithaml et al. (1996).

Moderators. Related to our second research objective, it is necessary to formally establish whether the respondents indeed perceive differences regarding the level of involvement and the type of offering. These two variables were measured using the scale developed by Ratchford (1987).

Manipulation Checks

To conduct a manipulation check, we used the average scores of the involvement items and the think/feel items. Regarding the level of involvement, we found significant differences between soft drink and day cream ($M_{\text{soft drink}} = 4.26$, $M_{\text{day cream}} = 4.94$, $p < 0.001$) as well as between toothpaste and DVD player ($M_{\text{toothpaste}} = 4.14$, $M_{\text{DVD player}} = 4.72$, $p < 0.001$). With respect to the type of offering (think vs. feel), significant differences were found between soft drink and

toothpaste ($M_{soft\ drink} = 4.91$, $M_{toothpaste} = 4.39$, $p < 0.001$) as well as between day cream and DVD player ($M_{day\ cream} = 4.76$, $M_{DVD\ player} = 3.99$, $p < 0.001$).

Research Objective 1

Analytical Approach

Analog to multiple regression analysis, predictive ability was evaluated by means of the multiple correlation coefficient R , which is defined as the correlation between the actual value of the dependent variable (y) and the predicted value of the dependent variable ($\hat{y}$). Thus,

$$R = r_{y\hat{y}}$$

Assessing research objective 1 ("to assess and to compare the performance of different commonly used customer value measurement methods with regard to their predictive ability of customer satisfaction, repurchase intentions and word-of-mouth.") involves testing the following set of hypotheses.

$$H_0 : r_{(y\hat{y})D} = r_{(y\hat{y})G} = r_{(y\hat{y})W} = r_{(y\hat{y})H}$$

$$H_A: \textit{at least one } r_{(y\hat{y})} \textit{ is different}$$

The letters D, G, W, H refer to the value measurement methods of Dodds et al. (1991), Gale (1994), Woodruff and Gardial (1996) and Holbrook (1999) respectively. The variable y ($\hat{y}$) represents the actual (predicted) value of satisfaction, repurchase intentions or word-of-mouth.

As each respondent filled out a questionnaire containing only one of the different value measurement methods under study, the four relevant correlation coefficients can be considered independent of one another. Thus, testing the null hypothesis involves testing whether four

independent sample correlation coefficients are statistically equal. For this purpose Zar (1996) proposes the test presented in equation 1.

$$\chi^2_{\alpha,k-1} = \sum_{i=1}^{k}(n_i - 3)z_i^2 - \frac{\left[\sum_{i=1}^{k}(n_i - 3)z_i\right]^2}{\sum_{i=1}^{k}(n_i - 3)} \quad (1)$$

Where

z_i = Fisher z-transformation of correlation coefficient r_i

n_i = sample size on which r_i is based

k = number of independent correlation coefficients

If the null hypothesis of equal independent correlation coefficients is rejected, it is of interest to determine which of the k correlation coefficients are different from which others. Therefore, we used pairwise comparisons based on a Tukey type test. This procedure implies that for each pair of correlation coefficients, r_A and r_B, the following null hypothesis is tested.

$$H_0 : r_A = r_B$$

$$H_1 : r_A \neq r_B$$

To test this null hypothesis we used the following test

$$q = \frac{z_B - z_A}{SE} \quad (2)$$

With

$$SE = \sqrt{\frac{1}{2}\left(\frac{1}{n_A - 3} + \frac{1}{n_B - 3}\right)}$$

The q statistic has a known distribution (see Table B5 of Zar [1996], which lists the critical values of the accompanying q distribution [i.e., $q_{\alpha,n,k}$]).

Parameter Estimation

Partial Least Squares (PLS) path modeling played a prominent role in the assessment of our empirical data. The reasons to opt for PLS path modeling are as follows. First, in line with our objective to evaluate predictive ability of the different value measurement approaches, an estimation approach that ensures optimal prediction accuracy was desirable. Second, PLS path modeling allowed us to estimate measurement models that include both formative and reflective indicators. This is particularly relevant as the literature indicates that value measurement models include both types of measurement (Ruiz et al., 2008). Third, PLS path modeling allowed us to calculate latent variable scores, which are crucial in assessing and comparing the predictive ability of the different value measurement methods under study.

To assess the statistical significance of the parameter estimates, we constructed percentile bootstrap confidence intervals based on 5000 samples (Preacher and Hayes, 2008).

Customer Value Measurement Methods: Measurement Model Structures

Before assessing the predictive ability of the different value measurement methods, it is necessary to correctly specify the accompanying measurement model structures.

In keeping with the existing literature (e.g., Baker et al., 2002), we specified Dodds et al.'s (1991) measurement scale for customer value as a first-order reflective measurement model. It should be

noted that the original scale development process by Dodds et al. (1991) also implies this particular measurement model.

With respect to the Customer Value Analysis suggested by Gale (1994), we started from its basic premise, namely that customer value equals the difference between a weighted quality score (termed market-perceived quality) and a weighted price score (termed market-perceived price). The market-perceived quality (price) score was determined by multiplying the relative performance score (relative price) for each quality (price) attribute by its normalized[2] weight and summing these weighted scores over the relevant quality (price) attributes. Subsequently, following the rationale of Jarvis, MacKenzie, and Podsakoff (2003), we used this market-perceived quality score and market-perceived price score as formative indicators of the customer value construct.

Concerning the customer value measurement approach recommended by Woodruff and Gardial (1996), it is important to distinguish between the first- and second-order constructs. According to research by Ruiz et al. (2008) and Lin et al. (2005), the benefit and sacrifice components (first-order constructs) in this approach should be considered formative components of customer value as customers make an explicit mental trade-off between these components to arrive at an overall value perception (second-order construct). The two first-order constructs – benefits and sacrifices – were modeled according to the guidelines developed by Jarvis et al. (2003): the

[2] In line with Gale's (1994) measurement method, we asked respondents to rate the importance of each attribute. Because the number of attributes was considerably large, a point allocation was not an option, so we used a Likert scale instead and normalized the weights (Bottomley, Doyle, and Green, 2000).

benefits construct consists of diverse positive consequences mentioned during the laddering interviews and, hence, is modeled formatively; the sacrifices construct, on the other hand, is measured by two reflective indicators reflecting the monetary consequences of the product. To model customer value as a second-order construct, we used the two-stage approach (Henseler, Wilson, Gotz, and Hautvast, 2007; Reinartz, Krafft, and Hoyer, 2004, Wilson and Henseler, 2007). In the first stage, the latent variable scores were estimated without the second-order construct (customer value) present, but with all of the first-order constructs (benefits and sacrifices) within the model. In the second stage, the latent variable scores of the first-order factors (benefits and sacrifices) were used as indicators of the second-order construct (customer value) in a separate higher-order PLS-model. It should be noted that consistent with the domain sampling method underlying the development of formative scales, items that yielded non-significant *p*-values in the first stage of the approach were retained in the second stage model (Diamantopoulos and Winklhofer, 2001; Jarvis et al., 2003).

Regarding the Customer Value Typology specified by Holbrook (1999), customer value can be considered a higher-order construct consisting of multiple components (Gallarza and Saura, 2006; Sánchez-Fernández et al. , 2009). Each of Holbrook's (1999) value types can be considered a first-order construct either measured by reflective or formative indicators. Because the different value types are not interchangeable, not necessarily correlated and the direction of causality is from each of the value types to the overall customer value construct, these value types can be considered formative components of customer value (Jarvis et al., 2003). To model customer value as a second-order construct, we again used the two-stage approach (Henseler et al., 2007; Reinartz et al., 2004, Wilson and Henseler, 2007).

Psychometric Properties

We first examined the psychometric properties of all first-order constructs used in our study. In terms of psychometric properties, it is crucial to distinguish between reflective and formative scales (MacKenzie, Podsakoff, and Jarvis, 2005).

Regarding the reflective scales, we assessed unidimensionality, internal consistency reliability, item validity, within-method convergent validity and discriminant validity respectively. Unidimensionality refers to the existence of a single construct underlying a set of items and is assessed following the procedure suggested by Karlis, Saporta, and Spinakis (2003). The test proposed by Jöreskog (1971) was used to gain insight in the internal consistency of the multiple-item constructs. Inspection of the magnitude and significance of the item loadings provide information regarding item validity. Within-method convergent validity and discriminant validity were assessed by means of Fornell and Larcker's (1981) formula of average variance extracted (AVE).

With regard to the formative scales, we only considered item validity and discriminant validity, since the basic principle underlying formative scales requires that every indicator is representative of a different aspect of the construct instead of being a reflection of the underlying construct. Concerning item validity, statistical significance is sufficient to conclude whether a formative indicator is valid or not (Diamantopoulos and Winklhofer, 2001). To obtain evidence for discriminant validity, we assessed whether the latent variable correlations fall within two standard errors of an absolute value of 1 (MacKenzie et al. , 2005).

All relevant psychometric properties of the constructs under study are presented together with the questionnaire in Appendices A and B. Our analyses confirmed favorable psychometric properties with exception of Dodds et al.'s (1991) approach in two settings. In particular, the eigenvalues (see Appendix B) of the construct's inter-item correlation matrix revealed that the scale suggested by Dodds et al. (1991) is not unidimensional for the toothpaste and the DVD player setting. As a result, we did not use the Dodds measurement model in the further analyses of these settings.

Customer Value Measurement Methods: Descriptive Statistics

Tables 2 to 5 present the correlations, the means and the standard deviations for the scores on value (or its dimensions), satisfaction, repurchase intentions, and word-of-mouth per setting for each of the measurement methods.

[INSERT TABLE 2 TO 5 ABOUT HERE PLEASE]

Results and Discussion

Table 6 displays the R-values (i.e., the square root of the coefficient of determination) for each of the settings as well as a pairwise comparison between these R-values. The R^2-values can be found in parenthesis. All R-values (R^2-values) are significantly different from zero, meaning that all four value measurement methods are capable of explaining customer evaluative judgments such as satisfaction, repurchase intentions and word-of-mouth. As a reminder please note that for toothpaste and DVD player the R- and R^2-values are not calculated for the Dodds approach, because the scale did not possess favorable psychometric properties.

[INSERT TABLE 6 ABOUT HERE PLEASE]

The results presented in Table 6 provide the following insights in the performance of the different value measurement methods in predicting customer evaluative judgments.

The methods proposed by Woodruff and Gardial (1996) and Holbrook (1999) are always among the methods with the highest predictive ability. For feel products these two methods perform equally well in predicting all three outcome variables. For think products this is not the case. Regarding low involvement think offerings, the method of Holbrook (1999) is the safest choice as its predictive ability is at least equal to that of Woodruff and Gardial's (1996) approach, whereas for high involvement offerings, the opposite holds. Here, the method of Woodruff and Gardial (1996) is preferred as its performance is at least equal to that of Holbrook's (1999) method.

Although in some instances the methods of Gale (1994) and Dodds et al. (1991) perform equally well as the methods of Woodruff and Gardial (1996) and Holbrook (1999), it is important to note that these first two methods never outperform the latter two methods. Furthermore, for none of the settings or outcome variables we find a significant difference in predictive ability between the methods of Gale (1994) and Dodds et al. (1991). In situations where researchers do prefer to use these suboptimal methods, the choice between Gale (1994) and Dodds et al. (1991) then involves a trade-off between scale length and actionability of the results. It should be noted however that,

given the unsatisfactory psychometric properties, the method by Dodds et al. (1991) is not suitable for think offerings.

Research Objective 2

We can conclude that multi-dimensional, consequence-based value measurement methods such as Woodruff and Gardial (1996) and Holbrook (1999) are in many instances the preferred approaches. However, as can be seen in Table 6, the superiority in predictive ability of these value measurement methods is not consistent across settings. To examine whether a structural pattern can be discerned among these differences in predictive ability (i.e., research objective 2), we will proceed by examining whether the underlying factorial design implied by the use of the FCB matrix (level of involvement * type of offering) moderates the relative performance of the value measurement methods.

As the measurement method of Dodds et al. (1991) did not possess favorable psychometric properties, this method was not included in this part of our study. Hence, in developing the hypotheses below we were interested in comparing the performance in predictive ability between on the one hand the methods of Woodruff and Gardial (1996) and Holbrook (1999), and on the other hand the method of Gale (1994).

Design

Given the different conceptual perspectives underlying the value measurement methods (see also Table 1), we expect that the relative ability of these methods to predict outcome variables, such as satisfaction and loyalty, depends on customer characteristics and product characteristics. To

investigate this, we will use the Foote, Cone and Belding (FCB) grid (Vaughn, 1980). The FCB grid classifies customers' purchase decisions on two dimensions: involvement and type of offering.

Involvement is defined as the attention of a customer to a product or a service because it is somehow important or relevant to him (Ratchford, 1987). High involvement means that the customer has great interest in the product or service at hand, and as a result will be motivated to search for more information. Low involvement, on the other hand, indicates that the customer has little interest in the product or service, and may not bother to pay attention to the same information (Ratchford, 1987; Vaughn, 1980).

Regarding the type of offering, the FCB grid discerns between think and feel offerings. Think offerings are products or services bought to satisfy utilitarian needs, while feel offerings represent products and services bought to satisfy emotional wants. As a result, think offerings involve mainly cognitive information processing, whereas feel offerings involve predominantly affective information processing (Claeys, Swinnen, and Vanden Abeele, 1995).

Hypothesis Development

Involvement. According to Mulvey, Olson, Celsi, and Walker (1994), the level of involvement influences the means-end chains of customers as follows. Customers with a high level of involvement mention more consequences in their laddering interviews compared to customers with a low level of involvement. This means that highly involved customers may have a better understanding of how specific attributes lead to desired consequences. This is consistent with the

study of Celsi and Olson (1988), which states that the customer's level of involvement affects the extent and focus of the comprehension processes by which the customer combines information about product attributes and consequences to form product evaluations and to make brand choices. More specifically, as the customer's level of involvement increases, his comprehension processes become increasingly elaborative and more inferences (i.e., thoughts beyond product information) about the product are made. On the basis of this theoretical foundation, we conjecture that the relative performance of value measurement methods is influenced by the degree of correspondence between the level of involvement associated with the offering and the level of abstraction of the benefits and sacrifices assessed by the value measurement method. This leads to the following hypothesis (H_1).

Hypothesis 1

The difference in performance between value measurement methods that assess benefits and sacrifices at the consequence level (i.e., Woodruff and Gardial; and Holbrook) and value measurement methods that do not assess benefits and sacrifices at the consequence level (i.e., Gale) is larger for high involvement offerings than for low involvement offerings.

Think/feel offerings. Think offerings are mainly bought for utilitarian reasons and involve attention to tangible, objective product features. Feel offerings, on the other hand, are bought for affective reasons. They are considered in terms of their intangible, subjective results and thus, the experience of the customer with the product is of paramount importance (Claeys et al., 1995; Hirschman and Holbrook, 1982; Mittal, 1989; Park and Young, 1983; Ratchford, 1987).

With regard to this distinction between think and feel products, it is interesting to note that the traditional view of a product as a bundle of tangible, objective attributes can be applicable for products whose value is derived from this tangible features (i.e., think products), but this approach is not appropriate for products that are selected because of the intangible and subjective aspects of the consumption experience (i.e., feel products) (Hirschman, 1980; Hirschman and Holbrook, 1982). Furthermore, research conducted by Claeys et al. (1995) has shown that means-end chains underlying think and feel offerings differ in contents. Compared to think offerings, the means-end chain for feel offerings is characterized by a higher level of abstraction. Put differently, the cognitive structure of think offerings contains concrete attributes and functional consequences, whereas the cognitive structure of feel offerings typically involves one abstract attribute and also includes psychosocial consequences. As shown, the customer value measurement methods considered in our study differ regarding the abstraction level at which they tap benefits and sacrifices. Consequently, the relative performance of the different customer value measurement methods is therefore hypothesized to vary for think and feel offerings such that better performance can be expected when there is a match between the type of information processing (think or feel) of the offering and the level of abstraction of the benefits and sacrifices assessed by the value measurement method. Hence, we put forward the following hypothesis (H_2).

Hypothesis 2

The difference in performance between value measurement methods that assess benefits and sacrifices at the consequence level (i.e., Woodruff and Gardial; and Holbrook) and value

measurement methods that do not assess benefits and sacrifices at the consequence level (i.e., Gale) is larger for feel offerings than for think offerings.

Involvement * think/feel offering. Only little research exists on the interaction between the type of offering and the level of involvement. Claeys et al. (1995) infer that under a high level of involvement the difference between think and feel offerings may become more prominent, because under high involvement conditions, the cognitive structure is better organized at the product-knowledge levels (i.e., the attributes) and the self-knowledge levels (i.e., the consequences). This hypothesis finds some support in the literature (Mittal, 1989; Park and Mittal, 1985; Park and Young, 1983). Accordingly, we propose the following hypothesis.

Hypothesis 3

In terms of the relative performance of value measurement methods that assess benefits and sacrifices at the consequence level (i.e., Woodruff and Gardial; and Holbrook) and value measurement methods that do not assess benefits and sacrifices at the consequence level (i.e., Gale), the difference in relative performance for feel and think products will be more pronounced in case of a high level of involvement than in case of a low level of involvement.

Analytical Approach

Relative performance is indicated by the difference in predictive ability of two methods. In general terms the parameter of interest can be expressed as presented in equation 3.

$$\Delta_{pq} = r_p - r_q \qquad (p \neq q) \tag{3}$$

Where r_p and r_q reflect the predictive ability of value measurement methods *p* and q respectively. In the context of the current study, this leads to the following parameters of interest: $\Delta_{WG} = r_W - r_G$ and $\Delta_{HG} = r_H - r_G$, which respectively assess the relative performance of Woodruff and Gardial's (1996) and Holbrook's (1999) method versus Gale's (1994) approach. The relative performance measures Δ_{WG} and Δ_{HG} will be computed for each of the three separate outcome variables: satisfaction, repurchase intentions, and word-of-mouth.

Although we did not put forward hypotheses for differences in relative performance between Holbrook's (1999) and Woodruff and Gardial's (1996) method, these methods are also explicitly compared for reasons of completeness. The need for this additional analysis is further underscored by the findings in Table 6 indicating that these two methods significantly differ in their predictive ability in several instances. The difference in relative performance between Holbrook's (1999) and Woodruff and Gardial's (1996) method is captured by the parameter $\Delta_{HW} = r_H - r_W$.

To examine whether the relative performance of the value measurement methods (i.e., Δ_{WG}, Δ_{HG}, Δ_{HW}) structurally varies as a consequence of the level of involvement and the type of offering, we opted for a factorial structural equation model (FAC-SEM). Originally developed by Iacobucci, Grisaffe, Duhachek, and Marcati (2003), the FAC-SEM approach enables researchers to assess how parameters in a structural model vary as a function of an underlying factorial design. The idea underlying FAC-SEM is analogous to that of n-way ANOVA. Whereas the

parameter of interest in n-way ANOVA is the mean, FAC-SEM focuses on the structural model parameters.

FAC-SEM discerns between interaction effect and main effect hypotheses. For the situation at hand, our hypotheses H1-H3 translate into the FAC-SEM hypotheses presented in Table 7.

[INSERT TABLE 7 ABOUT HERE PLEASE]

FAC-SEM analysis requires combining the data from different settings. For example, to assess the main effect of involvement, we needed to merge the data from the high involvement settings and compare them with the merged data from the low involvement settings. This is challenging as for the methods of Gale, Woodruff and Gardial, and Holbrook different items (i.e., variables) are used across the settings. To overcome this, we proceeded as follows to arrive at a structural model that was identical for all methods and across all settings. We started with estimating twelve (4 settings and 3 methods because Dodds et al. 1991 was not taken into account) models in which $SAT = f(VAL)$, $REP = f(SAT, VAL)$ and $WOM = f(SAT, VAL)$. In these equations the value construct was modeled in line with the suggested model structures outlined above.

For each of the outcome variables we then used the estimation results to construct the predicted values ($\hat{y}$). These predicted values were subsequently correlated to the actual data (i.e., the latent variable scores) on the outcome variables to serve as an estimate for the predictive ability (*R*). Due to the use of the latent variable scores as input, the data structure for each of the twelve

setting-method combinations was equal, which allowed us to construct merged data files needed to examine the different FAC-SEM hypotheses.

Results and Discussion

Below we discuss the results of our FAC-SEM analysis per pair of methods. Similar to n-way ANOVA, we start our interpretation with the highest-order statistically significant interaction (cf. Keppel, 1991). The results of the FAC-SEM analysis are presented in Table 8. The accompanying relative performance statistics as well as the differences in relative performance across the different cells can be derived from Table 6. In the succeeding discussion 'relative performance' refers the difference in performance between two value measurement methods (see also equation 3 above).

[INSERT TABLE 8 ABOUT HERE PLEASE]

Woodruff and Gardial (1996) vs. Gale (1994). Regarding the differences in relative performance of Woodruff and Gardial (1996) and Gale (1994), we find significant interaction effects for satisfaction and word-of-mouth. The aforementioned significant interactions imply that the difference in relative performance for feel and think products depends on the level of involvement. Unraveling the interaction effect for satisfaction, the data show that the difference in relative performance between feel and think products is significantly larger for low involvement offerings than for high involvement offerings. In addition, the interaction effect is disordinal in nature: for low involvement settings, the relative performance is larger for feel offerings than for think offerings, whereas the opposite is true for high involvement settings

(although for high involvement settings this difference between feel and think offerings is not significant). Concerning the interaction effect for word-of-mouth, we find that for high involvement settings the magnitude of the relative performance is different for feel and think offerings, whereas this is not the case for low involvement settings.

Holbrook (1999) vs. Gale (1994). Focusing on the difference in relative performance between the method's of Holbrook (1999) and Gale (1994) in explaining word-of-mouth and repurchase intentions, we also find significant interaction effects. For both outcome variables we find that for high involvement settings the magnitude of the relative performance of Holbrook and Gale is different for feel and think offerings, but this is not the case for low involvement settings.

Holbrook (1999) vs. Woodruff (1996). As mentioned before, we also examine whether the difference in performance between the methods of Holbrook (1999) and Woodruff and Gardial (1996) varies as a function of the underlying factorial design for reasons of completeness. As this additional analysis has a mere descriptive purpose, we continue by addressing the relative performance of the two methods for each of the four cells of our factorial design when the FAC-SEM analysis yields significant effects.

For satisfaction we find a significant disordinal interaction effect. For low involvement settings, Holbrook outperforms Woodruff for think offerings, but both methods perform equally well for feel offerings. For high involvement settings, Woodruff outperforms Holbrook for think offerings, but both methods perform equally well for feel offerings. For repurchase intentions we discern a similar pattern with the exception that both methods perform equally well for high

involvement think products. For word-of-mouth, we find a significant main effect for involvement: the difference in predictive ability between Holbrook's and Woodruff's method is significantly larger for low involvement settings than for high involvement settings.

Our second research objective was based on the expectation that the relative ability of the different value measurement methods to predict outcome variables, such as satisfaction and loyalty, depends on customer characteristics and product characteristics. This expectation was fueled by the different conceptual perspectives underlying the value measurement methods (see Table 1) as well as on the findings of research objective 1 (see Table 6). Although several interaction effects were statistically significant, the findings for our second research objective suggest that no structural pattern can be discerned among the differences in predictive ability.

Conclusion

This study was aimed at assessing and comparing the predictive ability of four commonly used methods (i.e., Dodds, Monroe and Grewal, 1991; Gale, 1994; Holbrook, 1999; Woodruff and Gardial, 1996) for measuring customer value (i.e., research objective 1) as well as at examining the contextual influence on the relative predictive ability of these methods (i.e., research objective 2). In our study we used customer satisfaction, repurchase intentions, and word-of-mouth as criterion variables.

To test the predictive ability of our four measurement methods, we used 16 (i.e., 4 methods * 4 settings) questionnaires and gathered data from 3,360 respondents (i.e., each of the 16

questionnaires was completed by 210 respondents). Partial Least Squares (PLS) path modeling was used to analyze the data.

Our findings provide several insights in the performance of the value measurement methods in predicting customer evaluative judgments.

First, the main results of this study provide support for the view that customer value is too complex to be operationalized as a one-dimensional construct (Petrick, 2002; Ruiz et al., 2008; Sweeney and Soutar, 2001). The one-dimensional measurement approach developed by Dodds et al. (1991) did not perform well in the four research settings. The scale either did not show unidimensionality or performed significantly less well compared to other measurement methods. Thus, our first conclusion is that customer value should be measured in a multi-dimensional way.

Second, it is interesting to note that the best performing methods (i.e., those of Woodruff and Gardial, 1996 and Holbrook, 1999) include benefits and sacrifices at the consequence level, whereas Gale's (1994) approach stays at the attribute level. These findings are in line with the service-dominant logic proposed by Vargo and Lusch (Lusch and Vargo, 2006; Vargo and Lusch, 2004), which states that "there is no value until an offering is used – experience and perception are essential to value determination" (Vargo and Lusch, 2006: 44). Thus, value is fundamentally derived and determined in use rather than in exchange (Vargo, Maglio, and Akaka, 2008), which is consistent with our findings that value should be measured at the consequence level rather than at the attribute level.

Third, the use of a relative value measurement method seems to be of no additional value in terms of predictive ability. The method proposed by Gale (1994) is the only method that assesses relative customer value perceptions and this method never outperforms the methods that only include absolute perceptions (i.e., Woodruff and Gardial, 1996 and Holbrook, 1999). However, it could be interesting to measure customer value in a multi-dimensional, consequence-based, relative way. It could be that such a conceptualization performs even better than the methods of Holbrook (1999) and Woodruff and Gardial (1996), since "in a competitive environment the relative approach seems more consistent with the way consumers make purchase decisions" (Babakus et al., 2004, p. 715).

Building on the findings regarding our first research objective that indicate that the predictive ability of the value measurement methods differs across settings, we assessed whether these differences in performance can be systematically explained by differences in involvement and type of product. Although several interaction effects were statistically significant, the findings for our second research objective suggest that no structural pattern can be discerned among the differences in predictive ability. This implies that our expectations that a particular measurement method performs better when there is a match between the level of abstraction of the benefits and sacrifices (attributes and/or consequences) assessed by the value measurement method, and the type of offering (think/feel) or the level of customer involvement, were not supported.

Nevertheless, our results succeed in providing marketing researchers and organizational managers a better understanding of the conceptualization of customer value. The marketing literature offers quite different conceptualizations of customer value and according to Woodruff

(1997) this fragmentation in conceptual knowledge is (partially) responsible for the lack of good and strong applications of the concept. We compared the performance of four commonly used conceptualizations of customer value and conclude that customer value should be operationalized in a multi-dimensional, consequence-based way.

Limitations and Further Research

Although this study contributes to our knowledge and understanding of customer value and its measurement, several limitations and further research suggestions deserve to be mentioned.

First, other products with more extreme levels of high/low involvement or think/feel could be used. Although the four settings selected for this study differed significantly in terms of involvement (high/low) or type of offering (think/feel), future work could replicate our findings in perhaps more extreme settings. Also, the applicability across different settings could be explored along dimensions other than the high/low involvement and think/feel offering tested in the present study. One dimension for further testing might be the level of product knowledge, which has been shown to affect the means-end associations made by customers (e.g., Graeff, 1997). In addition to addressing these research questions, future work could replicate our findings in service settings as well.

Second, as mentioned in our conclusion, the use of a relative value measurement method seems to be of no additional value in terms of predictive ability. The method proposed by Gale (1994) is the only method that assesses relative customer value perceptions and this method never outperforms the methods that only include absolute perceptions (i.e., Woodruff and Gardial, 1996

and Holbrook, 1999). Therefore, we suggest to investigate whether adjusting the multi-dimensional, consequence-based methods of Holbrook (1996) and Woodruff and Gardial (1996) to include a comparison with the competition provides additional explanatory power.

Third, the present study focused on the relative performance of four commonly used customer value measurement methods in terms of their predictive ability of satisfaction, repurchase intentions and word-of-mouth. We did not consider other measurement issues such as ease of administration, usefulness of results and ease of completion. Future work could explore how customer value measurement methods perform on those facets as well. Furthermore, the selection of a particular measurement method also depends on the objectives of the firm. When a firm is interested in its competitive position with respect to customer value, the methods of Holbrook (1999) and Woodruff and Gardial (1996) provide no clear answer. In this case, a relative approach, such as the one of Gale (1994), is required. As mentioned in the previous paragraph, a multi-dimensional, consequence based, relative approach could provide a solution.

Fourth, in our study we used customer satisfaction, repurchase intentions, and word-of-mouth as criterion variables. Although we deliberately chose to operationalize these outcome variables in a way that is consistent with the majority of existing academic research, we are aware that alternative approaches to measuring the three outcome variables might yield different results.

Finally, measures of actual purchase behavior, rather than behavioral intentions, could enhance the soundness of this study. Unfortunately, such behavioral data are often difficult and expensive to obtain. In addition, it should be noted that, although a significant positive association between

intention and behavior exists, the conversion of (re)purchase intentions into (re)purchase behavior is moderated by various factors, such as type of product, demographics and experience (e.g., Morwitz and Schmittlein, 1992; Seiders, Voss, Grewal, and Godfrey, 2005; Young, DeSarbo, and Morwitz, 1998).

Despite these limitations, this study provides a more comprehensive, in-depth understanding of customer value as well as an important tool for organizational managers since "making customer value strategies work begins with an actionable understanding of the concept itself" (Woodruff 1997, p. 141).

Appendix A

	TOOTHPASTE		SOFT DRINK		DVD PLAYER		DAY CREAM	
Attributes	*Quality attributes (0.98)*		*Quality attributes (1.00)*		*Quality attributes (1.00)*		*Quality attributes (1.00)*	
	Good taste		Good taste		Price-quality relationship		Caring	
	Whitening		Amount of sparkles		Look (e.g., design, color, size)		A well-known brand	
	Against dental caries		Amount of sugar		Quality		Quality	
	User-friendly packaging		Nice feeling in mouth		A well-known brand		Texture (gel, cream)	
	Cleaning		Packaging		User-friendly menu		A nice smell	
	Against dental plaque		A well-known brand		Short start-up time		Price-quality relationship	
	For sensitive teeth		Presence of extra ingredients		User-friendly remote control		Hypoallergenic (= little or no risk at allergic reaction)	
	A well known brand		(caffeine, tea extracts)		Recording possibilities (recorder, hard disk)			
	Quality				Technical possibilities (HDMI,USB port,…)		Working against a specific skin problem (e.g., oily skin, dry skin, redness)	
	Price attribute (-0.24)		*Price attribute (-0.23)*		*Price attribute (-0.31)*		*Price attribute (-0.31)*	
	Price		Price		Price		Price	
Consequences	*Benefits (1.00)*		*Benefits (1.00)*		*Benefits (1.00)*		*Benefits (1.00)*	
	Fresh breath	.56 **	Tastes good	.83 **	Easy to use	.77 **	Makes me feel good	.77 **
	Whiter teeth	.46 **	Thirst-quenching	.65 **	Good picture quality	.76 **	Makes me look good	.77 **
	Helps me to look good	.42 **	Healthier than other soft drinks	.36 **	Good sound quality	.77 **	Enhances my confidence	.65 **
	Enhances my confidence	.29 **	Nice feeling drinking this SD	.64 **	Looks good in my interior	.54 **	Makes my skin feel pleasant	.88 **
	Fresh taste in my mouth	.50 **	Gives me energy	.47 **	Quick start up	.58 **	Helps keeping skin healthy	.92 **
	Less dental caries	.58 **	I won't get fat	.31 **	Allows me to record movies and programs	.20 **	Applying this DC feels nice	.66 **
	Easy to use	.62 **	Bloated feeling (R)	.23 **			Feel clean	.72 **
	Makes brushing enjoyable	.69 **	Refreshing	.73 **	Energy-saving	.51 **	Refreshing	.72 **
	Clean teeth	.78 **	Brand ensures quality	.61 **	Brand ensures quality	.70 **	Brand ensures quality	.77 **
	Less dental plaque	.69 **			Meets my needs	.79 **	Budget-friendly (R)	.96 **
	Helps me feel good	.62 **						
	Healthy teeth	.70 **						
	Less dental pain	.61 **						
	Brand ensures quality	.85 **						
	Sacrifices (-0.32)		*Sacrifices (-0.14)*		*Sacrifices (-0.29)*		*Sacrifices (-0.40)*	
	Budget-friendly (R)	.96 **	Budget-friendly (R)	.97 **	Budget-friendly (R)	.97 **	Budget-friendly (R)	.97 **
	This choice saves me money (R)	.92 **	This choice saves me money (R)	.92 **	This choice saves me money (R)	.92 **	This choice saves me money (R)	.92 **

Note: (R) = reverse scored; Second-order factor loadings in parentheses.

*p < .10 **p < .05

Appendix B

Screening

1. Do you use toothpaste/day cream? yes - no
   Do you drink soft drink? yes - no
   Do you own a DVD player? yes - no
2. Do you sometimes buy toothpaste/day cream/soft drink? yes - no
   Have you bought this DVD player yourself? yes – no
3. Do you pay for toothpaste/day cream/soft drink yourself? yes - no
   Have you paid for this DVD player yourself? yes – no
4. Do you use this DVD player yourself? yes – no

Brand

Which brand of toothpaste/day cream/DVD player are you currently using?

Which brand of soft drink do you drink most of the time?

Note. We used this question for piping in the loyalty, value and satisfaction questions.

Moderators – Manipulation check

Involvement

Adapted from Ratchford (1987)

9-point Likert scale: totally disagree → totally agree

1. The (first) purchase of this particular brand of toothpaste/day cream/soft drink/DVD player is a very important decision.
2. The final choice for this particular brand of toothpaste/day cream/soft drink/DVD player requires a lot of thought.
3. I have a lot to lose when I choose the wrong brand of toothpaste/day cream/soft drink/DVD player.

Think/Feel

Adapted from Ratchford (1987)

9-point Likert scale: totally disagree → totally agree

1. The decision to choose this particular brand of toothpaste/day cream/soft drink/DVD player is mainly based on rational arguments.
2. The decision to choose this particular brand of toothpaste/day cream/soft drink/DVD player is not mainly based on facts.
3. The decision to choose this particular brand of toothpaste/day cream/soft drink/DVD player expresses one's personality.
4. The decision to choose this particular brand of toothpaste/day cream/soft drink/DVD player is based on a lot of feeling.
5. The decision to choose this particular brand of toothpaste/day cream/soft drink/DVD player is mainly based on sensory elements (such as looks, taste, touch or smell).

Value

Dodds, Monroe and Grewal (1991)

9-point Likert scale: totally disagree → totally agree

	TP	SD	DVD	DC
1. This X is a very good value for the money	.80 **	.81 **	.88 **	.82 **
2. At the price shown this X is very economical.	.73 **	.82 **	.69 **	.78 **
3. This is a good buy.	.82 **	.86 **	.89 **	.88 **
4. The price shown for this X is unacceptable. (R)	.42 **	.53 **	.44 **	.65 **
5. This X appears to be a bargain.	.37 **	.68 **	.43 **	.51 **
λ_1	*2.27*	*2.93*	*2.57*	*2.89*
λ_2	*1.14*	*.88*	*1.03*	*.90*
α		*.81*		*.81*
AVE		*.56*		*.55*

Note: (R) = reverse scored; X stands for toothpaste, soft drink, DVD player or day cream.
TP = toothpaste; SD = soft drink; DVD = DVD player; DC = day cream.
*p < .10 **p < .05

Gale (1994)

The items (attributes) are presented in Appendix A.

Importance

9-point scale: very unimportant → very important

Please indicate how important each of the following characteristics of toothpaste/day cream/soft drink/DVD players is to you.

Note. In line with Gale's (1994) measurement method, we asked respondents to rate the importance of each attribute. Because the number of attributes was considerably large, a division of weights was not an option, and we used a Likert scale anchored at 1 (very unimportant) and 9 (very important).

Performance (following Babakus, Bienstock, and Van Scotter, 2004)

9-point scale: much worse → about the same → much better

Please indicate how you evaluate your toothpaste/day cream/soft drink/DVD player relative to the competition.

Woodruff and Gardial (1996)

The items (consequences) are presented in Appendix A.

9-point Likert scale: totally disagree → totally agree

Holbrook (1999)

9-point Likert scale: totally disagree → totally agree

Social value - Adapted from Sweeney and Soutar (2001)	**TP**	**SD**	**DVD**	**DC**
Helps me to feel acceptable.	.94 **	.95 **	.98 **	.85 **
Improves the way I am perceived.	.95 **	.97 **	.99 **	.94 **
Makes a good impression on others.	.91 **	.92 **	.81 **	.95 **
Gives me social approval.	.91 **	.95 **	.95 **	.90 **
λ_1	*3.45*	*3.60*	*3.55*	*3.34*
λ_2	*.23*	*.25*	*.30*	*.32*
α	*.95*	*.96*	*.96*	*.93*
AVE	*.86*	*.90*	*.87*	*.83*
Second-order factor loadings	*0.09*	*0.03*	*-0.14*	*0.21*

Play - Adapted from Petrick (2002)	**TP**	**SD**	**DVD**	**DC**
Makes me feel good.	.82 **	.82 **	.58 **	.80 **
Gives me pleasure.	.91 **	.90 **	.81 **	.93 **
Gives me a sense of joy.	.95 **	.95 **	.90 **	.94 **
Makes me feel delighted.	.91 **	.96 **	.85 **	.94 **
Gives me happiness.	.91 **	.95 **	.82 **	.93 **
λ_1	*4.09*	*4.20*	*3.42*	*4.14*
λ_2	*.56*	*.42*	*.76*	*.52*
α	*.94*	*.95*	*.88*	*.95*

AVE	*.81*		*.84*		*.64*		*.83*	
Second-order factor loadings	*0.39*		*0.47*		*0.35*		*0.56*	
Excellence - Adapted from Oliver (1997)	**TP**		**SD**		**DVD**		**DC**	
The quality is excellent.	.87	**	.92	**	.83	**	.88	**
One of the best regarding quality.	.93	**	.94	**	.91	**	.92	**
High quality product.	.95	**	.94	**	.91	**	.93	**
Superior compared to competing products.	.84	**	.85	**	.81	**	.82	**
λ_1	*3.23*		*3.35*		*3.00*		*3.17*	
λ_2	*.41*		*.36*		*.51*		*.48*	
α	*.92*		*.93*		*.89*		*.91*	
AVE	*.81*		*.84*		*.75*		*.79*	
Second-order factor loadings	*0.99*		*0.98*		*0.91*		*0.96*	
Aesthetic value - Based on laddering interviews	**TP**		**SD**		**DVD**		**DC**	
I think I look good by using this TP/DC/SD.	.59	**	.96	**			.95	**
I think my teeth/skin is beautiful by using this TP/DC.	.93	**					.96	**
I think I have a fresh breath by using this toothpaste.	.88	**						
I think I have a nice figure by drinking this soft drink.			.93	**				
I think this DVD player is beautiful.					.92	**		
This DVD player looks good in my interior.					.92	**		
This DVD player has a beautiful design.					.95	**		
This DVD player has a beautiful color.					.93	**		
λ_1			*1.79*		*3.46*		*1.82*	
λ_2			*.21*		*.22*		*.18*	
α			*.88*		*.95*		*.90*	
AVE			*.89*		*.86*		*.91*	
Second-order factor loadings	*0.65*		*0.21*		*0.55*		*0.79*	
Efficiency - Adapted from Ruiz et al. (2008)	**TP**		**SD**		**DVD**		**DC**	
The price is high (R)	.05		.78		-.15		.05	
The effort I expend to receive X is high (R)	.35	*	-.55		.07		.24	
This TP/DC/DVD is easy to use	.98	**			.86	**	.99	**
Starting up the DVD player requires a lot of time (i.e., the time between turning on the DVD player and the moment the DVD starts). (R)					.48	**		
Second-order factor loadings	*0.42*		*0.00*		*0.68*		*0.47*	

(R) reverse scored; TP = toothpaste; SD = soft drink; DVD = DVD player; DC = day cream.
*p < .10 **p < .05

Satisfaction

Adapted from Anderson, Fornell, and Lehmann (1994)

11-point scale: totally dissatisfied → totally satisfied (cf. Wirtz and Lee, 2003)

Please indicate the extent to which you are satisfied or dissatisfied with your toothpaste/day cream/soft drink/DVD player.

Loyalty

Adapted from Zeithaml, Berry and Parasuraman (1996)

9-point Likert scale: extremely unlikely → extremely likely

Please indicate how likely it is that you would…

1. Say positive things about your toothpaste/day cream/soft drink/DVD player to other people.
2. Recommend your toothpaste/day cream/soft drink/DVD player to someone who seeks your advice.
3. Encourage friends and relatives to buy this toothpaste/day cream/soft drink/DVD player.
4. Consider this toothpaste/day cream/soft drink/DVD player your first choice to buy toothpaste/day cream/soft drink/DVD player.
5. Buy this toothpaste/day cream/soft drink/DVD player again when you need toothpaste/day cream/soft drink/DVD player.
6. Doubt about buying this toothpaste/day cream/soft drink/DVD player again.

Demographics

1. Gender
2. Age
3. Monthly net family income

References

Agarwal, Sanjeev, and R. Kenneth Teas (2001), "Perceived Value: Mediating role of Perceived Risk." *Journal of Marketing Theory and Practice* 9 (Fall), 1-14.

Anderson, Eugene W., Claes Fornell, and Donald R. Lehmann (1994), "Customer Satisfaction, Market Share, and Profitability: Findings from Sweden." *Journal of Marketing* 58 (July), 53-66.

Anderson, James C., and James A. Narus (2004), *Business Market Management: Understanding, Creating, and Delivering Value.* New Jersey: Prentice Hall.

Babakus, Emin, Carol C. Bienstock, and James R. Van Scotter (2004), "Linking Perceived Quality and Customer Satisfaction to Store Traffic and Revenue Growth." *Decision Sciences* 35 (September), 713-737.

Baker, Julie, A. Parasuraman, Dhruv Grewal, and Glenn B. Voss (2002), "The Influence of Multiple Store Environment Cues on Perceived Merchandise Value and Patronage Intentions." *Journal of Marketing* 66 (April), 120-141.

Bolton, Ruth N., and James H. Drew (1991), "A Multistage Model of Customers' Assessments of Service Quality and Value." *Journal of Consumer Research* 17 (March), 375–384.

Bottomley, Paul A., John R. Doyle, and Rodney H. Green (2000), "Testing the Reliability of Weight Elicitation Methods: Direct Rating versus Point Allocation." *Journal of Marketing Research* 37 (November), 508-513.

Bourdeau, Laurent, Jean-Charles Chebat, and Christian Couturier (2002), "Internet Consumer Value of University Students: E-mail-vs.-Web Users." *Journal of Retailing and Consumer Services* 9 (March), 61-69.

Butz Jr., Howard E., and Leonard D. Goodstein (1996), "Measuring Customer Value: Gaining the Strategic Advantage." *Organizational Dynamics* 24 (Winter), 63-77.

Caruana, Albert, and Noel Fenech (2005), "The Effect of Perceived Value and Overall Satisfaction on Loyalty: A Study among Dental Patients." *Journal of Medical Marketing* 5 (July), 245-255.

Celsi, Richard L., and Jerry C. Olson (1988), "The Role of Involvement in Attention and Comprehension Processes." *Journal of Consumer Research* 15 (September), 210-224.

Chen, Zhan, and Alan J. Dubinsky (2003), "A Conceptual Model of Perceived Customer Value in E-commerce: A Preliminary Investigation." *Psychology and Marketing* 20 (April), 323-347.

Claeys, C., A. Swinnen, and P. Vanden Abeele (1995), "Consumer's Means-end Chains for "Think" and "Feel" Products." *International Journal of Research in Marketing* 12 (October), 193-208.

Cronin Jr., J. Joseph, Michael K. Brady, and G. Tomas M. Hult (2000), "Assessing the Effects of Quality, Value, and Customer Satisfaction on Consumer Behavioral Intentions in Service Environments." *Journal of Retailing* 76 (Summer), 193-218.

Diamantopoulos, Adamantios, and Heidi M. Winklhofer (2001), "Index Construction With Formative Indicators: An Alternative to Scale Development." *Journal of Marketing Research* 38 (May), 269-277.

Dodds, William B., Kent B. Monroe, and Dhruv Grewal (1991), "Effects of Price, Brand, and Store Information on Buyers' Product Evaluations." *Journal of Marketing Research* 28 (August), 307-319.

Flint, Daniel J., Robert B. Woodruff, and Sarah Fisher Gardial (2002), "Exploring the Phenomenon of Customers' Desired Value Change in a Business-to-business Context." *Journal of Marketing* 66 (October), 102-117.

Fornell, Claes, and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error." *Journal of Marketing Research* 18 (February), 39-50.

Gale, Bradley T. (1994), *Managing Customer Value: Creating Quality and Service that Customers Can See.* New York: The Free Press.

Gallarza, Martina G., and Irene Gil Saura (2006), "Value Dimensions, Perceived Value, Satisfaction and Loyalty: An Investigation of University Students' Travel Behaviour." *Tourism Management* 27 (June), 437-452.

Graeff, Timothy R. (1997), "Comprehending Product Attributes and Benefits: The Role of Product Knowledge and Means-end Chain Inferences." *Psychology and Marketing* 14 (March), 163-183.

Gutman, Jonathan (1982), "A Means-end Chain Model Based on Consumer Categorization Processes." *Journal of Marketing* 46 (Spring), 60-72.

Gutman, Jonathan (1997). "Means-end Chains as Goal Hierarchies." *Psychology and Marketing* 14 (September), 545-560.

Hallowell, Roger (1996), "The Relationships of Customer Satisfaction, Customer Loyalty, and Profitability: An Empirical Study." *International Journal of Service Industry Management* 7 (4), 27-42.

Henseler, Jörg, Bradley Wilson, Oliver Götz, and Caspar Hautvast (2007), "Investigating the Moderating Role of Fit on Sports Sponsoring and Brand Equity." *International Journal of Sports Marketing and Sponsorship* 8 (July), 321-329.

Hirschman, Elizabeth C. (1980). "Attributes of Attributes and Layers of Meaning." *Advances in Consumer Research* 7 (January), 7-12.

Hirschman, Elizabeth C., and Morris B. Holbrook (1982), "Hedonic Consumption: Emerging Concepts, Methods and Propositions*." Journal of Marketing* 46 (Summer), 92-101.

Holbrook, Morris B. (1999), *Consumer Value: a Framework for Analysis and Research*. London: Routledge.

Holbrook, Morris B. (2006), "Consumption Experience, Customer Value, and Subjective Personal Introspection: An Illustrative Photographic Essay." *Journal of Business Research* 59 (June), 714-725.

Iacobucci, Dawn, Doug Grisaffe, Adam Duhachek, and Alberto Marcati (2003), "FAC-SEM: A Methodology for Modeling Factorial Structural Equations Models, Applied to Cross-Cultural and Cross-Industry Drivers of Customer Evaluations." *Journal of Service Research* 6 (August), 3-23

Jarvis, Cheryl Burke, Scott B. MacKenzie, and Philip M. Podsakoff (2003), "A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research." *Journal of Consumer Research* 30 (September), 199-218.

Jöreskog, Karl Gustav (1971), "Statistical Analysis of Sets of Congeric Tests." *Psychometrika*, 36 (June), 109-133.

Kamakura, Wagner A., Vikas Mittal, Fernando de Rosa, and José Afonso Mazzon (2002), "Assessing the Service-Profit Chain." *Marketing Science* 21 (3), 294-317.

Karlis, Dimitris, Gilbert Saporta, and Antonis Spinakis (2003), "A Simple Rule for the Selection of Principal Components." *Communications in Statistics: Theory and Methods* 32 (3), 643-666.

Keppel, Geoffrey (1991). *Design and Analysis: A Researcher's Handbook.* Upper Saddle River, New Jersey: Prentice-Hall.

Lai, Fujun, Mitch Griffin, and Barry J. Babin (2009), "How Quality, Value, Image, and Satisfaction Create Loyalty at a Chinese Telecom." *Journal of Business Research* 62 (October), 980-986.

Laitamäki, Jukka, and Raymond Kordupleski (1997), "Building and Deploying Profitable Growth Strategies Based on the Waterfall of Customer Value Added." *European Management Journal* 15 (April), 158-166.

Lam, Shun Yin, Venkatesh Shankar, M. Krishna Erramilli, and Bvsan Murthy (2004), "Customer Value, Satisfaction, Loyalty, and Switching Costs: An Illustration from a Business-to-Business Service Context." *Journal of the Academy of Marketing Science* 32 (July), 293-311.

Lin, Chien-Hsin, Peter J. Sher, and Hsin-Yu Shih (2005), "Past Progress and Future Directions in Conceptualizing Customer Perceived Value." *International Journal of Service Industry Management* 16 (4), 318-336.

Loveman, Gary W. (1998), "Employee Satisfaction, Customer Loyalty, and Financial Performance: An Empirical Examination of the Service Profit Chain in Retail Banking." *Journal of Service Research* 1 (August), 18-31.

Louviere, Jordan J., and Towhidul Islam (2008). "A Comparison of Importance Weights and Willingness-to-Pay Measures Derived from Choice-Based Conjoint, Constant Sum Scales and Best–Worst Scaling." *Journal of Business Research*, 61(September), 903-911.

Lusch, Robert F., and Stephen L. Vargo (2006), *The Service-Dominant Logic of Marketing: Dialog, Debate, and Directions.* New York: M.E. Sharpe Inc.

MacKenzie, Scott B., Philip M. Podsakoff, and Cheryl Burke Jarvis (2005), "The Problem of Measurement Model Misspecification in Behavioral and Organizational Research and Some Recommended Solutions."*Journal of Applied Psychology* 90 (July), 710-730

Mittal, Banwari (1989), "Must Consumer Involvement Always Imply More Information Rearch." *Advances in Consumer Research* 16 (January), 167-172.

Morwitz, Vicki G., and David Schmittlein (1992), "Using Segmentation to Improve Sales Forecasts Based on Purchase Intent: Which "Intenders" Actually Buy?" *Journal of Marketing Research* 29 (November), 391-405.

Mulvey, Michael S., Jerry C. Olson, Richard L. Celsi, and Beth A. Walker (1994), "Exploring the Relationships Between Means-End Chains and Involvement." *Advances in Consumer Research* 21 (1) 51-57.

Oliver, Richard L. (1997), *Satisfaction: A Behavioral Perspective on the Consumer.* New York: McGraw-Hill.

Overby, Jeffrey W., Sarah Fisher Gardial, and Robert B. Woodruff (2004), "French versus American Consumers' Attachment of Value to a Product in a Common Consumption Context: A Cross-National Comparison." *Journal of the Academy of Marketing Science* 32 (September), 437-460.

Park, C. Whan, and Banwari Mittal (1985), "A Theory of Involvement in Consumer Behavior: Problems and Issues." In *Research in Consumer Behavior*, ed. Jagdish N. Sheth, 201-231. Greenwich, CT: JAI Press.

Park, C. Whan, and S. Mark Young (1983), "Types and Levels of Involvement and Brand Attitude Formation." *Advances in Consumer Research* 10 (January), 320-324.

Petrick, James F. (2002), "Development of a Multi-Dimensional Scale for Measuring the Perceived Value of a Service." *Journal of Leisure Research* 34 (2), 119-134.

Preacher, Kristopher, and Andrew Hayes (2008), "Asymptotic and Resampling Strategies for Assessing and Comparing Indirect Effects in Multiple Mediator Models." *Behavior Research Methods* 40 (3), 879-891.

Ratchford, Brian T. (1987), "New Insights About the FCB Grid." *Journal of Advertising Research* 27 (August/September), 24–38.

Reinartz, Werner, Manfred Krafft, and Wayne D. Hoyer (2004), "The Customer Relationship Management Process: Its Measurement and Impact on Performance." *Journal of Marketing Research* 41 (August), 293-305.

Rintamäki, Timo, Hannu Kuusela, and Lasse Mitronen (2007), "Identifying Competitive Customer Value Propositions in Retailing." *Managing Service Quality* 17 (6), 621-634.

Ruiz, David Martin, Dwayne D. Gremler, Judith H. Washburn, and Gabriel Cepeda Carrión (2008), "Service Value Revisited: Specifying a Higher-Order, Formative Measure." *Journal of Business Research* 61 (December), 1278-1291.

Sánchez-Fernández, Raquel, and M. Ángeles Iniesta-Bonillo (2007), "The Concept of Perceived Value: A Systematic Review of the Research." *Marketing Theory* 7 (December), 427-451.

Sánchez-Fernández, Raquel, M. Ángeles Iniesta-Bonillo, and Morris B. Holbrook (2009), "The Conceptualisation and Measurement of Consumer Value in Services." *International Journal of Market Research* 51 (January), 93-113.

Seiders, Kathleen, Glenn B. Voss, Dhruv Grewal, and Andrea L. Godfrey (2005), "Do Satisfied Customer Buy More? Examining Moderating Influences in a Retailing Context." *Journal of Marketing* 69 (October), 26-43.

Setijono, Djoko, and Jens J. Dahlgaard (2007), "Customer Value as a Key Performance Indicator (KPI) and a Key Improvement Indicator (KII)." *Measuring Business Excellence* 11 (June), 44-61.

Slater, Stanley F. (1997), "Developing a Customer Value-Based Theory of the Firm." *Journal of The Academy of Marketing Science* 25 (March), 162-167.

Slater, Stanley F., and John C. Narver (2000), "Intelligence Generation and Superior Customer Value." *Journal of the Academy of Marketing Science* 28 (January), 120-127.

Strauss, Anselm L., and Juliet M. Corbin (1998), *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. London: Sage Publications.

Sweeney, Jillian C., and Geoffrey N. Soutar (2001), "Consumer Perceived Value: The Development of a Multiple Item Scale." *Journal of Retailing* 77 (Summer), 203-220.

Teas, R. Kenneth, and Sanjeev Agarwal (2000), "The Effects of Extrinsic Product Cues on Consumers' Perceptions of Quality, Sacrifice, and Value." *Journal of the Academy of Marketing Science* 28 (March), 278-290.

Ulaga, Wolfgang, and Samir Chacour (2001), "Measuring Customer-Perceived Value in Business Markets." *Industrial Marketing Management* 30 (August), 525-540.

Vargo, Stephen L., and Robert F. Lusch (2004), "Evolving to a New Dominant Logic for Marketing." *Journal of Marketing* 68 (January), 1-17.

Vargo, Stephen L., Paul P. Maglio, and Melissa Archpru Akaka (2008). "On Value and Value Co-Creation: A Service Systems and Service Logic Perspective." *European Management Journal* 26 (June), 145-152.

Vaughn, Richard (1980), "How Advertising Works: A Planning Model." *Journal of Advertising Research* 20 (October), 27-33.

Wang, Yonggui, Hing Po Lo, Renyong Chi, and Yongheng Yang (2004), "An Integrated Framework for Customer Value and Customer-Relationship-Management Performance: a Customer-Based Perspective from China." *Managing Service Quality* 14 (2/3), 169-182.

Wilson, Bradley, and Jörg Henseler (2007). *Modeling Reflective Higher-Order Constructs using Three Approaches with PLS Path Modeling: A Monte Carlo Comparison.* Paper presented at the Australian and New Zealand Marketing Academy Conference. Otago, Australia, December 3-5.

Wirtz, Jochen, and Meng Chung Lee (2003), "An Examination of The Quality and Context-Specific Applicability of Commonly Used Customer Satisfaction Measures." *Journal Of Service Research* 5 (May), 345-355.

Woodruff, Robert B. (1997), "Customer Value: The Next Source for Competitive Advantage*." Journal of The Academy Of Marketing Science* 25 (March), 139-153.

Woodruff, Robert B., and Sarah Fisher Gardial (1996), *Know Your Customer: New Approaches To Understanding Customer Value and Satisfaction.* Cambridge, MA: Blackwell Publications.

Young, Martin R., Wayne S. DeSarbo, and Vicki G. Morwitz (1998), "The Stochastic Modeling of Purchase Intentions and Behavior." *Marketing Science,* 44 (2), 188-202.

Zar, Jerrold H. (1996) *Biostatistical Analysis.* New Jersey: Prentice-Hall.

Zeithaml, Valerie A. (1988), "Consumer Perceptions of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence." *Journal of Marketing* 52 (July) 2-22.

Zeithaml, Valerie A., Leonard L. Berry, and A. Parasuraman (1996), "The Behavioral Consequences of Service Quality." *Journal of Marketing* 60 (April) 31-46.

Table 1

Differences between Measurement Methods

	Dodds et al. (1991)	**Gale (1994)**	**Woodruff and Gardial (1996)**	**Holbrook (1999)**
1. Approach	One-dimensional	Multi-dimensional	Multi-dimensional	Multi-dimensional
2. Nature of costs/benefits	n.a.	Attributes	Consequences	Attributes and consequences
3. Competition	No	Yes	No	No

Table 2

Summary of Correlations, Means and Standard Deviations for the Dodds Method

	Think							Feel						
	VAL	SAT	REP	WOM	*M*	*SD*		VAL	SAT	REP	WOM	*M*	*SD*
VAL	—	.48**	.47**	.45**	6.65	1.28	VAL	—	.32**	.27**	.35**	6.24	1.46
SAT	.34**	—	.52**	.56**	7.78	1.64	SAT	.41**	—	.64**	.50**	8.26	1.23
REP	.33**	.43**	—	.73**	6.41	1.65	REP	.33**	.55**	—	.58**	7.29	1.53
WOM	.42**	.38**	.52**	—	6.26	1.75	WOM	.34**	.57**	.55**	—	6.84	1.48
M	5.89	7.91	7.14	6.07			*M*	5.56	8.38	7.60	6.44		
SD	1.06	1.42	1.56	1.81			*SD*	1.59	1.24	1.38	1.95		

Note. Correlations for the high involvement offerings are presented above the diagonal, and correlations for the low involvement offerings are presented below the diagonal. Means and standard deviations for the high involvement offerings are presented in the vertical columns, and means and standard deviations of the low involvement offerings are presented in the horizontal rows.

VAL = value; SAT = Satisfaction; REP = Repurchase Intentions; WOM = Word-of-Mouth.

*p < .05 **p < .01

Table 3

Summary of Correlations, Means and Standard Deviations for the Gale Method

Think								Feel							
	MPQ	MPP	SAT	REP	WOM	*M*	*SD*		MPQ	MPP	SAT	REP	WOM	*M*	*SD*
MPQ	—	-.35**	.43**	.51**	.58**	6.15	1.12	MPQ	—	-.36**	.45**	.45**	.46**	6.78	1.26
MPP	-.44**	—	-.13	-.14*	-.18**	3.88	1.51	MPP	-.15*	—	-.15*	-.14	-.12	4.08	1.69
SAT	.46**	-.17*	—	.59**	.69**	7.80	1.71	SAT	.37**	-.18*	—	.65**	.55**	8.46	1.28
REP	.37**	.00	.59**	—	.62**	6.30	1.55	REP	.35**	-.05	.46**	—	.57**	7.44	1.43
WOM	.49**	-.15*	.54**	.61**	—	6.25	1.90	WOM	.49**	-.07	.47**	.50**	—	6.90	1.57
M	6.28	4.21	8.31	7.24	6.30			*M*	6.61	4.33	8.69	7.79	6.78		
SD	1.14	1.32	1.27	1.65	1.96			*SD*	1.06	1.76	1.00	1.23	1.62		

Note. Correlations for the high involvement offerings are presented above the diagonal, and correlations for the low involvement offerings are presented below the diagonal. Means and standard deviations for the high involvement offerings are presented in the vertical columns, and means and standard deviations of the low involvement offerings are presented in the horizontal rows. MPQ = Market-Perceived Quality; MPP = Market-Perceived Price; SAT = Satisfaction; REP = Repurchase Intentions; WOM = Word-of-Mouth.

*p < .05 **p < .01

Table 4

Summary of Correlations, Means and Standard Deviations for the Woodruff Method

	Think								Feel						
	BEN	SAC	SAT	REP	WOM	*M*	*SD*		BEN	SAC	SAT	REP	WOM	*M*	*SD*
BEN	—	-.22**	.65**	.48**	.70**	6.58	1.08	BEN	—	-.34**	.59**	.50**	.70**	7.16	1.25
SAC	-.33**	—	-.14*	-.17*	-.15*	4.00	1.74	SAC	-.27**	—	-.24**	-.17*	-.32**	4.19	2.33
SAT	.50**	-.34**	—	.55**	.68**	7.70	1.69	SAT	.61**	-.02	—	.53**	.54**	8.35	1.26
REP	.43**	-.01	.51**	—	.65**	6.28	1.66	REP	.52**	-.06	.60**	—	.59**	7.40	1.44
WOM	.51**	-.18**	.46**	.55**	—	6.43	1.81	WOM	.50**	-.19**	.56**	.53**	—	7.03	1.40
M	6.28	4.73	7.96	7.07	5.98			*M*	6.19	5.27	8.16	7.50	6.33		
SD	1.26	1.87	1.30	1.61	1.81			*SD*	1.10	2.19	1.22	1.35	1.79		

Note. Correlations for the high involvement offerings are presented above the diagonal, and correlations for the low involvement offerings are presented below the diagonal. Means and standard deviations for the high involvement offerings are presented in the vertical columns, and means and standard deviations of the low involvement offerings are presented in the horizontal rows. BEN = benefits; SAC = Sacrifices; SAT = Satisfaction; REP = Repurchase Intentions; WOM = Word-of-Mouth.

*p < .05 **p < .01

Table 5

Summary of Correlations, Means and Standard Deviations for the Holbrook Method

	Think									
	AEST	EFF	EXC	PLAY	SOC	SAT	REP	WOM	*M*	*SD*
AEST	—	.02	.47**	.55**	.25**	.34**	.23**	.33**	5.36	1.90
EFF	.10	—	.08	-.12	-.33**	.24**	.22**	.11	6.90	1.31
EXC	.51**	.18**	—	.35**	.09	.54**	.47**	.53**	5.96	1.46
PLAY	.69**	-.05	.30**	—	.44**	.14*	.02	.21**	4.46	1.69
SOC	.49**	-.27**	.04	.57**	—	-.08	-.13	-.01	2.06	1.59
SAT	.41**	.29**	.70**	.23**	.03	—	.54**	.56**	7.94	1.34
REP	.36**	.23**	.73**	.16*	-.09	.67**	—	.63**	6.15	1.49
WOM	.52**	.10	.69**	.42**	.23**	.59**	.59**	—	6.19	1.69
M	5.61	6.96	6.67	4.30	2.71	8.01	7.18	6.10		
SD	1.79	1.25	1.47	2.14	1.90	1.73	1.70	1.82		
	Feel									
	AEST	EFF	EXC	PLAY	SOC	SAT	REP	WOM	*M*	*SD*
AEST	—	.16*	.60**	.56**	.26**	.49**	.46**	.52**	6.52	1.67
EFF	-.19**	—	.05	-.06	-.35**	.20**	.21**	.06	6.97	1.36
EXC	.22**	-.12	—	.40**	.23**	.67**	.57**	.52**	6.72	1.36
PLAY	.52**	-.17*	.35**	—	.48**	.31**	.30**	.43**	5.60	1.92
SOC	.59**	-.33**	.09	.40**	—	.07	.04	.26**	3.27	1.97
SAT	.13	.02	.66**	.25**	-.01	—	.76**	.58**	8.21	1.21
REP	-.04	-.07	.53**	.16*	-.11	.55**	—	.67**	7.39	1.50
WOM	.24**	-.03	.51**	.42**	.15*	.59**	.61**	—	6.81	1.56
M	2.59	6.39	6.82	4.29	2.14	8.38	7.51	6.08		
SD	1.93	1.45	1.65	2.13	1.72	1.36	1.58	2.29		

Note. Correlations for the high involvement offerings are presented above the diagonal, and correlations for the low involvement offerings are presented below the diagonal. Means and standard deviations for the high involvement offerings are presented in the vertical columns, and means and standard deviations of the low involvement offerings are presented in the horizontal rows. AEST = Aesthetics; EFF = Efficiency; EXC = Excellence; PLAY = Play; SOC = Social Value; SAT = Satisfaction; REP = Repurchase Intentions; WOM = Word-of-Mouth.

*p < .05 **p < .01

Table 6

Comparison between the Coefficients of Determination

		Satisfaction					Word-of-Mouth					Repurchase Intentions			
		D	**G**	**W**	**H**		**D**	**G**	**W**	**H**		**D**	**G**	**W**	**H**
toothpaste	**D**					**D**					**D**				
Think - Low involv	**G**		.46(.21)		**	**G**		.61(.37)		*	**G**		.62(.38)		**
	W			.56(.31)	**	**W**			.63(.40)		**W**			.62(.38)	**
	H		**	**	.71(.50)	**H**		*		.72(.52)	**H**		**	**	.78(.61)
		D	**G**	**W**	**H**		**D**	**G**	**W**	**H**		**D**	**G**	**W**	**H**
soft drink	**D**	.47(.22)		**	**	**D**	.60(.36)				**D**	.63(.39)			
Feel - Low involv	**G**		.38(.14)	**	**	**G**		.58(.33)			**G**		.55(.31)		
	W	**	**	.74(.55)		**W**			.59(.35)		**W**			.67(.45)	
	H	**	**		.67(.45)	**H**				.62(.39)	**H**				.64(.40)
		D	**G**	**W**	**H**		**D**	**G**	**W**	**H**		**D**	**G**	**W**	**H**
DVD player	**D**					**D**					**D**				
Think - High involv	**G**		.43(.19)	**	**	**G**		.76(.58)		**	**G**		.69(.48)		
	W		**	.73(.54)	*	**W**			.76(.58)	**	**W**			.61(.38)	
	H		**	*	.62(.38)	**H**		**	**	.62(.38)	**H**				.61(.37)
		D	**G**	**W**	**H**		**D**	**G**	**W**	**H**		**D**	**G**	**W**	**H**
day cream	**D**	.42(.18)		**	**	**D**	.56(.32)		**		**D**	.65(.43)			*
Feel - High involv	**G**		.45(.20)	*	**	**G**		.60(.36)	*		**G**		.73(.53)		
	W	**	*	.62(.38)		**W**	**	*	.73(.54)		**W**			.67(.45)	
	H	**	**		.68(.47)	**H**				.64(.41)	**H**	*			.77(.60)

Note: This table displays the *R*-values with the R^2-values in parenthesis. D = Dodds; G = Gale; W = Woodruff and Gardial; H = Holbrook.
*p < .10 **p < .05

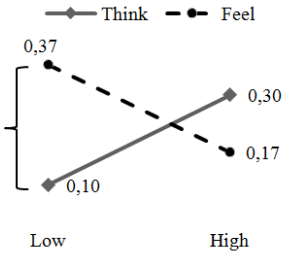
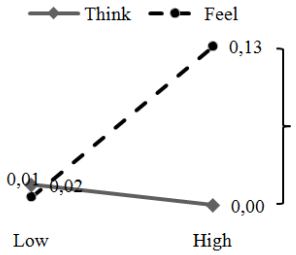
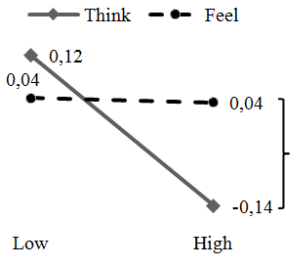
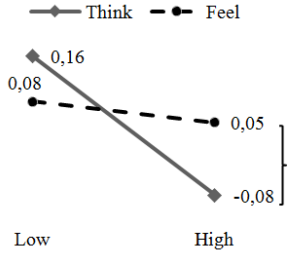
Table 7

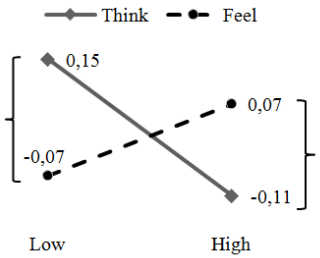
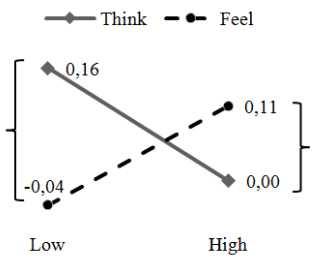
FAC-SEM Hypotheses

Woodruff and Gardial vs. Gale	
Main effect involvement (H1)	$H_0 : \Delta_{WG(High)} \leq \Delta_{WG(Low)}$
	$H_A : \Delta_{WG(High)} > \Delta_{WG(Low)}$
Main effect think/feel (H2)	$H_0 : \Delta_{WG(Feel)} \leq \Delta_{WG(Think)}$
	$H_A : \Delta_{WG(Feel)} > \Delta_{WG(Think)}$
Interaction effect (H3)	$H_0 : (\Delta_{WG(Feel)} - \Delta_{WG(Think)})_{High} \leq (\Delta_{WG(Feel)} - \Delta_{WG(Think)})_{Low}$
	$H_A : (\Delta_{WG(Feel)} - \Delta_{WG(Think)})_{High} > (\Delta_{WG(Feel)} - \Delta_{WG(Think)})_{Low}$
Holbrook vs. Gale	
Main effect involvement (H1)	$H_0 : \Delta_{HG(High)} \leq \Delta_{HG(Low)}$
	$H_A : \Delta_{HG(High)} > \Delta_{HG(Low)}$
Main effect think/feel (H2)	$H_0 : \Delta_{HG(Feel)} \leq \Delta_{HG(Think)}$
	$H_A : \Delta_{HG(Feel)} > \Delta_{HG(Think)}$
Interaction effect (H3)	$H_0 : (\Delta_{HG(Feel)} - \Delta_{HG(Think)})_{High} \leq (\Delta_{HG(Feel)} - \Delta_{HG(Think)})_{Low}$
	$H_A : (\Delta_{HG(Feel)} - \Delta_{HG(Think)})_{High} > (\Delta_{HG(Feel)} - \Delta_{HG(Think)})_{Low}$
Holbrook vs. Woodruff and Gardial	
Main effect involvement (H1)	$H_0 : \Delta_{HW(High)} = \Delta_{HW(Low)}$
	$H_A : \Delta_{HW(High)} \neq \Delta_{HW(Low)}$
Main effect think/feel (H2)	$H_0 : \Delta_{HW(Feel)} = \Delta_{HW(Think)}$
	$H_A : \Delta_{HW(Feel)} \neq \Delta_{HW(Think)}$
Interaction effect (H3)	$H_0 : (\Delta_{HW(Feel)} - \Delta_{HW(Think)})_{High} = (\Delta_{HW(Feel)} - \Delta_{HW(Think)})_{Low}$
	$H_A : (\Delta_{HW(Feel)} - \Delta_{HW(Think)})_{High} \neq (\Delta_{HW(Feel)} - \Delta_{HW(Think)})_{Low}$

Table 8

FAC-SEM Results

Woodruff and Gardial vs. Gale	**Satisfaction**	**Word-of-Mouth**	**Repurchase Intentions**
Interaction effect involvement * think/feel	[-.59; -.23]	[.02; .25]	ns
			
Main effect involvement	ns	ns	ns
Main effect think/feel	ns	ns	ns
Summary hypotheses tests	H3 supported in opposite direction (disordinal interaction)	H3 supported (disordinal interaction)	H3, H2, H1 not supported
Holbrook vs. Gale	**Satisfaction**	**Word-of-Mouth**	**Repurchase Intentions**
Interaction effect involvement * think/feel	ns	[.11; .38]	[.07; .33]
			
Main effect involvement	ns	ns	ns
Main effect think/feel	ns	ns	ns
Summary hypotheses tests	H3,H2,H1 not supported	H3 supported (disordinal interaction)	H3 supported (disordinal interaction)

Additional analysis ***Holbrook vs. Woodruff and Gardial***	***Satisfaction***	***Word-of-Mouth***	***Repurchase Intentions***
*Interaction effect involvement * think/feel*	*[.24; .59]* 	*ns*	*[.17; .43]* 
Main effect involvement	*ns*	*[-.31; -.06]*	*ns*
Main effect think/feel	*ns*	*ns*	*ns*

Note: The differences indicated with an accolade are significant at the .05-level.