A multiple-item scale for measuring customer loyalty development

Rosalind McMullan

Department of Nutrition and Food Science, Auburn University, Auburn, Alabama, USA

Abstract

Purpose – This paper seeks to explore the complex inter-relationships between the attitudinal and behavioural dimensions of customer loyalty development, by examining the dynamic processes by which customer loyalty is initiated and sustained using a mixed methods approach. In doing so, the paper highlights the absence of valid and reliable measures of customer loyalty development and discusses the use of the multi-phase model of customer loyalty development.

Design/methodology/approach – This model is the basis for the construction of a multi-item scale to measure customer loyalty development. A mixed methods design is specified and stages in the construction of the scale are discussed including measures of validity and reliability.

Findings – The findings of the research demonstrate the validity and reliability of the loyalty scale and highlight the sustaining and mediating effects associated with different levels of loyalty development.

Research limitations/implications – The study is set within the passenger ferry sector. Future research will seek to make empirical generalisations in relation to the application of the loyalty scale.

Practical implications – The main implications of this research are to emphasise the importance of sustaining and developing customer loyalty based on a differentiated approach to rewarding customers who have different levels of loyalty development. The findings highlighted the need to acknowledge the importance of reciprocity in terms of which aspects of service customers value within different levels of loyalty.

Originality/value – The main contributions of this paper are the presentation of the loyalty scale and the confirmation of the plateau of customer loyalty development.

Keywords Customer loyalty, Customer service management, Consumer behaviour, Behaviourally-anchored rating scales

Paper type Research paper

An executive summary for managers and executive readers can be found at the end of this article.

Introduction

The development of customer loyalty has become an important focus for marketing strategy in recent years due to the benefits associated with retaining existing customers (Gwinner et al., 1998; Hagen-Danbury and Matthews, 2001). Despite this, the concept of customer loyalty remains relatively unexplored (Hart et al., 1999). Whilst numerous studies have distinguished between the attitudinal and behavioural dimensions of loyalty (e.g. Jacoby and Kyner, 1973; Dick and Basu, 1994; Knox and Walker, 2001), these have not adequately explored the complex inter-relationships between the two dimensions, and the dynamic processes by which loyalty is initiated and sustained. Finding an accurate measure of customer loyalty is extremely important due to its link with profitability (Reichheld, 2003). The underpinning purpose of this paper is to contribute to the knowledge and understanding in measuring customer loyalty development.

This paper begins by reviewing progress made within the literature relating to frameworks for understanding customer loyalty and its measurement. The paper discusses existing approaches to understanding and measuring customer loyalty development and presents Oliver’s (1999) model as the basis for developing a multi-item scale. The scale’s development, pilot, validity and reliability tests are detailed with conclusions stating implications of the loyalty scale for researchers and practitioners.

In reviewing the literature in relation to customer loyalty it is important to note differences in terminology including brand loyalty (e.g Jacoby and Chesnut, 1978), customer loyalty (e.g Oliver, 1997) and service loyalty (Gremler and Brown, 1999). A detailed review of such terms may be read in Knox and Walker’s (2001) paper. These differences are sometimes semantic, but in general the term used tends to frame the focus of the research. This paper is concerned with customer loyalty to a brand, product or service and as such is customer orientated.

Customer loyalty

There is recognition of a need for greater knowledge and understanding in relation to customer loyalty (Knox and
Walker, 2001). This results from uncertainty that exists over the meaning and measurement of the construct and the absence of academic literature in this area (Oliver, 1997; 1999; Hart et al., 1999). Most analyses of loyalty have been from a behavioural perspective, excluding attitudinal type data and concentrating on a deterministic perspective using stochastic models (Tellis, 1988; Ehrenberg, 1988; Ehrenberg and Goodhardt, 2000). A problem associated with this type of analysis, is that loyalty is about much more than just repeat purchase; someone who keeps buying may be doing so out of inertia, indifference or exit barriers rather than loyalty (Reichheld, 2003). Recent studies have concentrated on the relationship between customer loyalty and quality, satisfaction (Selnes, 1993; Mittal and Lasser, 1998; Oliver, 1999; Martensen et al., 2000; McDougall and Levesque, 2000) profitability (Hallowell, 1996) or lack of profitability (Reinartz and Kumar, 2000) and frequency programme effectiveness (Dowling and Uncles, 1997; O’Malley, 1998; Shoemaker and Lewis, 1999). Thus, despite all the interest in the general concept and the universal belief in the benefits of loyalty, progress in measuring and clearly defining it has been very limited (Knox and Walker, 2001). Table I summarises the main contributions of studies within the literature, which have sought to understand customer loyalty.

The studies presented in Table I collectively enhance knowledge and understanding of customer loyalty. Some of the studies highlighted have contributed to defining the construct whilst others have approached its measurement. Progress has also been made in identifying and understanding antecedents of customer loyalty through the use of multi-item measurement scales. In reviewing these approaches it is clear that there is an absence of an instrument capable of measuring customer loyalty development whilst identifying what is important for sustaining and developing loyalty or rendering it vulnerable. One aim of this paper is to overcome this absence. There are numerous benefits associated with being able to identify different groups of customers. For example, identifying loyal customers allows this group to be harnessed as promoters of the business through word of mouth marketing; secondly, by identifying different groups it is possible to ascertain the level of profitability each generates (Reichheld, 2003).

Theoretical framework for the development of the multi-item loyalty scale

Oliver (1999) hypothesised that there are four phases or plateau in the development of customer loyalty. This research will refer to these as phases. Each phase has a number of characteristics or dimensions, which act as either sustainers (attracting the customer to stay) or vulnerabilities (pulling the customer towards a substitute). The first three phases and their characteristics are based on existing validated research, however the fourth remains untested (Fishbein and Ajzen, 1972; Jacoby and Chesnut, 1978; Dick and Basu, 1994; Oliver, 1999). One aim of this research is to test the fourth phase of the model.

Figure 1 shows that in addition to the four phases and their characteristics of customer loyalty development, there are two mediating factors, sustaining and vulnerability elements. The mediating factors allow modelling of the continued influence of competitors, advertising, service failure and other external influences that sustain or make an existing customer’s loyalty development vulnerable.

As customers progress through the phases of loyalty development, the sustainers and vulnerability elements change to reflect the degree of involvement. The theory is that once a customer has found a product or service that he or she enjoys (meeting with expectations of cost, quality and benefits), and continues to use, he or she becomes less concerned with seeking alternatives and does not respond to advertising or competitive threats (Oliver, 1999). One way to test Oliver’s theory and the four-phase model of customer loyalty development is through a multi-item scale. The loyalty scale was constructed, to include the four phases, their characteristics and mediating factors in the development of a customer’s loyalty. The procedures followed in the development of the loyalty scale are now discussed.

Developing multi-item scales

Numerous advantages have been highlighted in the use of scaling techniques including the meaningful comparison of two results at a specific stage in time and the subsequent measure over time to check stability (Rajecki, 1990). One of the main values of a scale is its ability to measure a concept by using multiple indicators rather than one, which facilitate tapping the complexity of concepts (De Vaus, 1996). A single observation may be misleading and lacking in context thus multi-item measurement scales can help overcome these distortions. Scales also allow for greater precision, specifically in relation to ranking or classifying groups and identifying subsequent differences or similarities (Green et al., 1988).

Lastly, by summarising the information presented by a number of questions into one variable (in this case customer loyalty development) the analysis is simplified. However, problems such as interpretation and wording of the question may affect the validity of multi-item measurement scales (Oskamp, 1991). The main problem however, is the way in which response sets can invalidate questionnaire answers. Several types of response sets exist including carelessness, social desirability, extremity of response and acquiescence (Edwards, 1969; Rotter, 1966; Bradburn and Sudman, 1979; De Vaus, 1996). Numerous methods were employed in this research to partially control or overcome response sets bias (Williams, 1992; Knox and Walker, 2001). Five stages, drawn from the literature (Bearden et al., 1993; De Vaus, 1996), were taken to develop the loyalty scale, as illustrated in Figure 2.

Stage 1. Outline and delineate the construct's domain

The first stage related to the theoretical definition with the construct’s domain being thoroughly outlined and delineated (Bearden et al., 1993). This was derived from a thorough review of the literature and an expert opinion. Based on the literature review customer loyalty was operationally defined for this study to have six characteristics. The first characteristic is based on the deterministic philosophy of purchasing being more than a random event, that purchases are “biased” or preferred in favour of one alternative over another. The second characteristic related to a behavioural response or a purchase. It is insufficient to study attitudes in isolation of purchase behaviours within a marketing context. The third characteristic related to purchase behaviours being expressed over a period of time. Expression of intention of
Stage 2. Develop a set of questions to measure the concept
A set of questions (items) was developed to measure customer loyalty development (De Vaus, 1996). The questions consisted of a mixture of favourable and unfavourable statements to which respondents would be asked to rate their point of agreement or disagreement. The statements were selected to reflect orientation to the attitude of interest. This helped to distinguish between different groups of people and their responses. The responses ranged from strongly agree to strongly disagree.

Stage 1. Define the concept
A set of questions (items) was developed to measure customer loyalty development (De Vaus, 1996). The questions consisted of a mixture of favourable and unfavourable statements to which respondents would be asked to rate their point of agreement or disagreement. The statements were selected to reflect orientation to the attitude of interest. This helped to distinguish between different groups of people and their responses. The responses ranged from strongly agree to strongly disagree.

Secondary research is recommended for developing a set of validated and reliable questions for use in a scale (Bearden et al., 1993; Green et al., 1988; De Vaus, 1996; Oliver, 1997). There are two complementary approaches to this, one conceptual the other empirical. The first approach was used to examine the conceptual content of the items. The second approach was used after piloting the scale to obtain a correlation matrix of the items. Items will normally have...
modest correlations (0.3 or above) with each other item in the scale (De Vaus, 1996).

Stage 3. Trim and refine pool of items
A number of existing scales were reviewed and a pool of 122 items generated. The scales related directly or indirectly to the antecedents, sustainers and vulnerabilities of customer loyalty development. These scales were examined using criteria for validity and reliability (Bearden et al., 1993). The criteria included the number of items per scale, the Cronbach’s alpha or reliability level of each scale and best practice. A panel of experts was formed to validate, trim and refine the initial items. The panel consisted of five experts; three academics who specialised in service quality, customer loyalty and services marketing; and two marketing practitioners, one of whom is responsible for managing a customer loyalty programme. The panel’s brief was to evaluate each item based on criteria that examined the theoretical definition, the construct’s domain and the operational definition (Bearden et al., 1993). In other words, the scale items needed to be consistent with the literature.

The optimum length of scale is debated within the literature with suggestions ranging from 20 to no longer than 33 items (Raju, 1980; Bearden et al., 1993; Pritchard et al., 1999). The panel sought to reduce the number of items from 122, whilst ensuring that each of the four phases of customer loyalty development was represented. The pilot multi-item scale consisted of six items tapping the cognitive phase (C), seven tapping the affective phase (A), nine tapping the conative phase (CO) and six tapping the action phase (AC). The multi-item scale also included items relating to attraction and vulnerability elements. Avoiding duplication of items optimised clarity. The items were arranged into statements within a questionnaire format and Likert scoring developed from 1-7 to allow an extensive range of scoring. The multi-item scale consisted of 28 items and was administered to a sample of customers who broadly represented characteristics of those chosen for the survey proper.

Stage 4. Pilot items and refine
The validity of the pilot multi-item scale was tested using Factor Analysis SPSS Version 9 and based on this analysis minor revisions were made. The scale was piloted amongst a sample of restaurant diners who belonged to a University training restaurant dining club during November 1999 (Beggs and Gilmore, 2001; McMullan and Gilmore, 2003). Restaurant customers were considered to be an appropriate market segment due to the individual’s freedom of choice of where to dine, in terms of price, service quality, and range of cuisine on offer and atmosphere. In other words, the purchasing decision was based on customers’ prior knowledge of eating out within an area (cognitive), what type of food and service he or she preferred (cognitive), where he or she had eaten recently and whether this was favourable or unfavourable (affective) and where he or she, based on these preceding factors, intended to eat out next (conative).
A multiple-item scale for measuring customer loyalty development

Rosalind McMullan

The passenger ferry sector used in the main study broadly shared these characteristics. Both sectors are within services industries and share common characteristics such as freedom of choice, prior knowledge of service, preferences and intentions. Changes included changing phraseology to make statements clearer, changing US English to UK English, ordering the questions to reduce respondent fatigue from similar phase questions, altering the service context from restaurant to passenger ferry sector. In addition, a statement relating to individual attention was removed and an additional switching price related statement inserted. The Likert rating of 1-7 was reduced to 1-5, in order to ease respondents’ understanding and interpretation (Churchill, 1979; Bearden et al., 1993). All changes were made in consultation with the expert panel.

The main quantitative study involved a postal survey, which included the 28 multi-item loyalty scale (see Appendix). This was administered during July 2001 to customers of a leading passenger ferry company operating within the UK. The survey was administered to passengers who had previously sailed with the company on a particular route. A random sample of the company’s database, which was made up of a population of 60,000 existing customers across the United Kingdom (UK), identified 3,000 names and addresses spread evenly across regions. This represented 5 per cent of the company’s population and met with criteria allowing the findings to be generalized (De Vaus, 1996).

Numerous steps were taken to increase the response rate including Dillman’s total design method (Dillman, 1978). Incentives in the form of a 10 per cent voucher off the next sailing were offered to all respondents who completed and returned the questionnaire within a three-week time frame in order to optimise the response rate. There are numerous reasons to support the use of incentives, despite the response set bias that may occur as a result. Research studies on postal surveys identify five factors, of which incentives are one, that are effective in increasing the response rates in public opinion surveys (Paxson, 1995). Incentives compensate the respondent for his or her time (Dillman, 1978) whilst acknowledging the norm of reciprocity (Gouldner, 1960; Gendall et al., 1998). Incentives also provide cost benefits to the research. A study by Brennan et al. (1993) found that a prepaid incentive of $1 and one reminder produces approximately the same response as an equivalent survey with no incentive but two reminders. The study found similar results when replicated within the UK using 20 pence (34 cents) as an incentive (Jobber and O’Reilly, 1996). No reminders were used in this study. Incentives are also advocated for methodological purposes where a large number of responses is required in order to apply statistical tests such as factor analysis (Turley, 1999).

Before analysis was carried out the data were coded and organised. The questionnaires were scanned using an optical mark reader (OMR) and the data were imported into SPSS Version 9. Advantages of using an OMR are efficiency and an absence of human error associated with manual data input. The data were screened for errors and missing data were coded.

Stage 5. Plot development scores for individuals and add up individual scores
Stage five in the construction of the multi-item scale related to scoring respondents’ responses. A multi-item scale score is a summary of an individual’s responses to a number of questions. An unweighted factor based scale was used due to ease of use and interpretation (Green et al., 1988; Bryman and Cramer, 1997). This approach allowed the identification of the development of customer loyalty through the four phases. The rationale for this is best illustrated by considering two respondents with the same score, whose opinions may have differed. Furthermore, scale scores must be interpreted in relative terms, as they are not absolute (e.g. an individual can not be 75 per cent loyal, rather they can have a high comparative score). Thus, it was necessary to plot scores within a distribution to identify high, moderate and low scores. In order to overcome the problem of upper and lower limits, minimum and maximum values were specified (Tull and Hawkins, 1990). Reichheld (2003) supports this approach arguing that customer surveys should be kept simple for ease of interpretation and criticised the interpretation of scores based on complex weighting algorithm. Consequently, it was considered that the term level would be a more appropriate description of the numeric score derived from the loyalty scale than phase. The term level is used within the findings.

One of the aims of this research was to establish a method to classify, compare and measure differing groups of customers, rather than employ ranking methods. As such, each statement on the loyalty scale is viewed as equal, for example a cognitive statement is of equal value to an affective item; therefore weighting the statements was inappropriate. This approach is supported by within the literature (Green et al., 1988). Furthermore, the loyalty scale was derived from Oliver’s (1993) model, which detailed phases or plateaux of loyalty development. None of the issues within his model was given greater weight. However, further research could examine the validity of categorising the items by type for example, price, facilities, service level and status.

Findings
The data were considered to be at ordinal level (Cohen and Holliday, 1982). Empirical evidence exists to support the treatment of ordinal variables as if they conform to interval scales in order to have the widest choice of tests (Freeman, 1965; Labovitz, 1967, 1970). The results of the unspecified factor analysis are shown in Table II. A component matrix was generated to ensure that the analysed variables had reasonable correlations with other variables (Norusis, 1985). Unrotated and rotated component matrices were inspected and variables that did not or correlated weakly with others were excluded (correlations less than or equal to 0.3) (De Vaus, 1996). All but one variable correlated well on the three components. The result of KMO of sampling adequacy was 0.906 and Barlett’s test was 8648.984, which is considered a high Chi-square, significant at 0.00. The results of these tests rendered the data very factorable and consequently the factor analysis was generated.

The un-specified factor analysis points to six factors, having an eigenvalue of over 1, the first three accounting for the greatest amount of variance (Table II). Table II shows each factor and the extent to which variance or eignevalues can be explained by each factor. Three tests were applied to this six-factor solution in order to confirm validity before reliability analysis (De Vaus, 1996). These tests were Kaiser’s criterion, a scree test, and to overcome weaknesses within the former
Kaiser’s criterion is used to select those factors, which have an eigenvalue greater than one. Kaiser’s criterion is recommended for data where the number of variables is less than 30, in this case there were 28, and where the average communality is greater than or equal to 0.70 or when the number of subjects is greater than 250, in this case there were 950 subjects (Bryman and Cramer, 1997). There were 950 cases when missing data were excluded from the analysis. This data set met two of the assumptions but failed in the other, as the mean communality was 0.543.

The second test was the scree test (Cattell, 1966). The scree test showed a break between the steep slope of the initial factors and a gentle one for the remainder, implying that the latter were less important. The greatest degree of variance was explained by factors 1-3 with the factors levelling between 5-7. The factors to be retained were those which came before the point at which the eigenvalues levelled.

The third test, RanEigen (random eigen), was carried out to ensure the appropriate number of factors was retained. A weakness of Kaiser’s criterion and the scree test is that often too many components are extracted, and it is not always clear where to draw the line that discriminates “significant” from “random” (Enzmann, 1997). The results of the RanEigen identified three factors with a potential weak fourth factor, which is consistent with the scree test.

Based on the results of these tests, it was decided to exclude the weak fourth factor and specify the conditions of the factor analysis to an optimum three factors solution. The extraction method was principal component analysis with varimax rotation. The factors were rotated to increase their interpretability and identify more clearly what they represent. The rotated matrix compared more favorably with the unrotated matrix in this respect. Varimax rotation, a method of orthogonal rotation, was specified in order to increase the interpretability of factors. Varimax rotation was chosen over oblimin rotation as examination of the correlation matrix showed that factors were reasonably uncorrelated. Varimax rotation assumes that the factors are unrelated. Factors are rotated to maximise the loadings of the items. The items are used to identify the conceptual meaning of the factors (Bryman and Cramer, 1997).

Table III shows the item number and the extent to which it correlates or loads under each factor. The highest loading per item and factor is taken in all cases. For example item q1_17_q1 (item 1 or Question 1) loads highest on Factor 1 and is excluded from Factors 2 and 3. There is no absolute rule in relation to how high a co-efficient should be before it is said to load on a factor, however it would be unusual to include co-efficients below 0.3 (De Vaus, 1996; Bryman and Cramer, 1997). Figure 3 highlights the conceptual analysis of the factors identifying three themes. The three themes consist of items that sustain a customer’s loyalty (Factor 1) and those

### Table II Six-factor solution and with corresponding items

<table>
<thead>
<tr>
<th>Component</th>
<th>Total</th>
<th>Initial Eigenvalues</th>
<th>Cumulative %</th>
<th>Total</th>
<th>Extraction sums of squared loadings</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>% of variance</td>
<td>Cumulative %</td>
<td></td>
<td>% of variance</td>
<td>Cumulative %</td>
</tr>
<tr>
<td>2</td>
<td>2.321</td>
<td>8.290</td>
<td>34.696</td>
<td>2.321</td>
<td>8.290</td>
<td>34.696</td>
</tr>
<tr>
<td>3</td>
<td>2.014</td>
<td>7.194</td>
<td>41.890</td>
<td>2.014</td>
<td>7.194</td>
<td>41.890</td>
</tr>
<tr>
<td>4</td>
<td>1.278</td>
<td>4.566</td>
<td>46.456</td>
<td>1.278</td>
<td>4.566</td>
<td>46.456</td>
</tr>
<tr>
<td>5</td>
<td>1.162</td>
<td>4.150</td>
<td>50.606</td>
<td>1.162</td>
<td>4.150</td>
<td>50.606</td>
</tr>
<tr>
<td>6</td>
<td>1.041</td>
<td>3.717</td>
<td>54.323</td>
<td>1.041</td>
<td>3.717</td>
<td>54.323</td>
</tr>
<tr>
<td>7</td>
<td>0.992</td>
<td>3.543</td>
<td>57.867</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.895</td>
<td>3.195</td>
<td>61.062</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.844</td>
<td>3.016</td>
<td>64.077</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.828</td>
<td>2.958</td>
<td>67.035</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.778</td>
<td>2.779</td>
<td>69.814</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.758</td>
<td>2.707</td>
<td>72.521</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.689</td>
<td>2.459</td>
<td>74.980</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.664</td>
<td>2.371</td>
<td>77.351</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.636</td>
<td>2.272</td>
<td>79.623</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.582</td>
<td>2.080</td>
<td>81.703</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.558</td>
<td>1.992</td>
<td>83.695</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.527</td>
<td>1.882</td>
<td>85.577</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>0.503</td>
<td>1.797</td>
<td>87.374</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.483</td>
<td>1.726</td>
<td>89.100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>0.451</td>
<td>1.611</td>
<td>90.711</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>0.446</td>
<td>1.592</td>
<td>92.303</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>0.406</td>
<td>1.450</td>
<td>93.753</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>0.386</td>
<td>1.380</td>
<td>95.133</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>0.363</td>
<td>1.296</td>
<td>96.429</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>0.360</td>
<td>1.286</td>
<td>97.715</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>0.330</td>
<td>1.180</td>
<td>98.895</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>0.309</td>
<td>1.105</td>
<td>100.000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
items that present vulnerability, which could be considered as the “deal breakers” in relation to price (Factor 2) and service (Factor 3). Factor 1, the factor with the greatest number of items, includes cognitive items such as choice, punctuality, reservation, information, facilities and affective items including preference, enjoyment, loyalty and recommendation. Factor 1 has been labelled “Loyalty Sustainers” as conceptually it consists of those issues, which sustain and develop customer’s loyalty further. In contrast to the sustaining and mediating effect discussed by Oliver (1999), many of the items that sustain a customer’s loyalty are internal. It includes some weaker items, which relate to choosing the right ferry operator, punctuality, promotional offers and inertia. These items could be dropped as their co-efficients are below 0.5 but above 0.3, rendering them weak. The issues were duplicated to some extent by other items, thus the lower co-efficient provides for choosing the best item and creating a more parsimonious scale. The co-efficient of item 20 (Q20) loaded marginally higher on Factor 1 than Factor 2. This is interesting to note, as it seems to challenge the notion of inertia.

Factor 2 is labelled “Loyalty Vulnerabilities: Price” and is characterised by price-related items such as bargain hunting, value for money, and switching for £10 or £20. This demonstrates the key areas of price that cause potential vulnerabilities. Two of the items in this factor were weak (below 0.5 but above 0.3). Based on this, items 13 (Q13) and 28 (Q28) could be excluded; however, location is an important element within services (Q13) and the extent to which preference exists is also an important means of discriminating (Q28). Factor 3 is solely concerned with service vulnerabilities and is labelled “Loyalty Vulnerabilities: Service”. One of the items in Factor 3 is weak (below 0.5 but above 0.3). Item 19 (Q19) relates to the challenge posed by a new service. In the context of this study, the introduction of the low-cost airline operators presented an important area of vulnerability and as such this item adds value. The three factors appear to mirror the

<table>
<thead>
<tr>
<th>Question or item number</th>
<th>Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>q1_17_q1</td>
<td>0.424</td>
</tr>
<tr>
<td>q1_17_q2</td>
<td>0.559</td>
</tr>
<tr>
<td>q1_17_q3</td>
<td>0.497</td>
</tr>
<tr>
<td>q1_17_q4</td>
<td>0.579</td>
</tr>
<tr>
<td>q1_17_q5</td>
<td>0.520</td>
</tr>
<tr>
<td>q1_17_q6</td>
<td>0.494</td>
</tr>
<tr>
<td>q1_17_q7</td>
<td>2.083E-02</td>
</tr>
<tr>
<td>q1_17_q8</td>
<td>-0.154</td>
</tr>
<tr>
<td>q1_17_q9</td>
<td>0.673</td>
</tr>
<tr>
<td>q1_17_q10</td>
<td>0.737</td>
</tr>
<tr>
<td>q1_17_q11</td>
<td>0.617</td>
</tr>
<tr>
<td>q1_17_q12</td>
<td>0.642</td>
</tr>
<tr>
<td>q1_17_q13</td>
<td>-5.045E-02</td>
</tr>
<tr>
<td>q1_17_q14</td>
<td>0.617</td>
</tr>
<tr>
<td>q1_17_q15</td>
<td>0.180</td>
</tr>
<tr>
<td>q1_17_q16</td>
<td>-3.594E-02</td>
</tr>
<tr>
<td>q1_17_q17</td>
<td>-1.799E-02</td>
</tr>
<tr>
<td>q18_28_q18</td>
<td>0.154</td>
</tr>
<tr>
<td>q18_28_q19</td>
<td>7.208E-02</td>
</tr>
<tr>
<td>q18_28_q20</td>
<td>0.496</td>
</tr>
<tr>
<td>q18_28_q21</td>
<td>0.715</td>
</tr>
<tr>
<td>q18_28_q22</td>
<td>5.514E-02</td>
</tr>
<tr>
<td>q18_28_q23</td>
<td>0.548</td>
</tr>
<tr>
<td>q18_28_q24</td>
<td>0.436</td>
</tr>
<tr>
<td>q18_28_q25</td>
<td>0.429</td>
</tr>
<tr>
<td>q18_28_q26</td>
<td>0.334</td>
</tr>
<tr>
<td>q18_28_q27</td>
<td>0.658</td>
</tr>
<tr>
<td>q18_28_q28</td>
<td>0.401</td>
</tr>
</tbody>
</table>

Notes: extraction method: principal component analysis; rotation method: varimax with Kaiser normalization; rotation converged in five iterations

Figure 3 Loyalty scale items loading on factors

Factor 1 Loyalty Sustainers (16 items)
Q1 When travelling by ferry operator, it is important that I choose the right ferry operator
Q2 I know ______ well enough to evaluate it against other ferry operators
Q3 ______ is always punctual
Q4 Making a reservation with ______ is always straightforward
Q5 Information is always easily obtained from ______
Q6 I regularly take advantage of ______ promotional offers
Q7 ______ facilities are appealing
Q8 In general ______ is exactly what I need from a ferry operator
Q9 ______ as a choice of ferry operator hasn’t worked out as well as I thought it would
Q10 I could do it over again, I’d choose an alternative ferry operator to
Q11 I have a preference for ______ as a ferry operator
Q12 I usually travel with the same ferry operator
Q13 I enjoy crossing with ______
Q14 I consider myself loyal to ______
Q15 I would get tired of travelling by ______ every time I travel
Q16 I have recommended ______ to my friends

Factor 2 Loyalty Vulnerabilities: Price (8 items)
Q7 When deciding on a ferry operator, I am not interested in bargain hunting
Q8 When choosing a ferry operator, I compare prices of different companies to be sure I get the best value for money
Q13 When choosing a ferry operator, the port is the most important factor
Q15 I would try an alternative ferry operator if it were £10 less expensive than ______
Q22 I would try an alternative ferry operator if it were £20 less expensive than ______
Q25 When I travel, I choose the ferry operator that I know best
Q26 If I like a ferry operator I rarely switch from it just to try something different
Q28 If ______ is not available it makes little difference to me and I travel with an alternative ferry operator

Factor 3 Loyalty Vulnerability Service (4 items)
Q16 I would try an alternative ferry operator if the alternative offered increased facilities over ______
Q17 I would try an alternative ferry operator if the alternative offered increased status over ______
Q18 I would change ferry operator if the alternative’s staff was friendlier
Q19 When I see a new ferry service different from the usual, I will try it
A multiple-item scale for measuring customer loyalty development

Rosalind McMullan

plateaux that Oliver proposes. The factors are not conceptually distinct in terms of Oliver’s phases, but overlap. However, the items clearly represent those issues which sustain or render vulnerable customer loyalty development. For example cognitive and affective items collectively make up Factor 1. Sustainers dominate Factor 1, whilst Factors 2 and 3 are characterised by vulnerabilities.

Reliability analysis was carried out to ensure the factors were reliable (Bearden et al., 1993). The results are based on 1,017 cases. It is important to note that the factor analysis was based on 950 cases. The procedure for factor analysis provides an opportunity to exclude missing cases, which was applied to make the data more factorable. The same facility is not available under reliability analysis. Scale mean, variance, correlation and alpha if item was to be deleted are presented. The results of the reliability analysis of Factor 1, which included 16 items, shows a Cronbach’s alpha of 0.8762 (standardised item alpha of 0.8825) indicating reliability (Table IV). The reliability analysis of Factor 2 indicated a low reliability score with a standardised item alpha of 0.6834.

Factor 3 had a standardised item alpha of 0.4940. Whilst items could be excluded to increase the reliability scores of these two factors, their conceptual make up is stronger with the retention of the weaker items. An examination of the intraclass correlation coefficients and the interrater reliability estimates, served as a check of the analysis to ensure that no items needed to be excluded from Factor 1 of the loyalty scale, to improve the reliability (Bearden et al., 1993).

The reliability analysis of Factor 1 compares favourably with other scales used within marketing. For example the reliability of Oliver’s (1997) scale to measure “Satisfaction” achieved 0.82, “SERVQUAL’s” reliability ranged between 0.87-0.90 (Parasuraman et al., 1988) and Slama and Tashchian’s (1987) scale “Purchasing Involvement”, had a Cronbach’s alpha of 0.86. Therefore, the loyalty scale has a comparable level of reliability at the upper limit in relation to the aforementioned scales.

The internal reliability of the loyalty scale was examined by asking participants, face to face, to determine if the scale had correctly categorised their phase of loyalty development, which was then compared to their individual scores. Four focus groups took place nine months after the loyalty scale was administered to gauge if the respondent’s level of loyalty development had changed. Each focus group was composed of between six to nine respondents. As the duration of each focus group was approximately 60 minutes, a similar incentive to that used within the survey was employed to attract participants and to reward them for their time and effort. The authors of this research conducted the focus groups. Each focus group covered six main discussion points, including opinions about travelling by ferry, choosing a ferry operator, preferred service dimension, comparisons with other forms of transport, loyalty towards the ferry operator and awareness of promotional offers. The discussion points were based on findings, which emerged from analysis of the scale’s findings. The analysis was structured by the sequence of the discussion point and by scores determined by the loyalty scale. This approach to focus group analysis is advocated within the literature (Coffey and Atkinson, 1996; Shaw, 1999; Krueger and Casey, 2000).

In general, analysis of the focus groups found the loyalty scale to be reliable, with the majority of participants in each score band, displaying antecedents, sustaining and vulnerable elements associated with the appropriate level of loyalty development. Whilst most respondents remained in the same level of loyalty development areas of vulnerability had emerged. During the nine months interval there had been slippage by a few participants to a lower level of loyalty development due to unresolved dissatisfaction with some elements of the company’s service and persuasion and trial of low-cost alternatives (such as low-cost airlines). This stage of

Table IV Results of reliability analysis

<table>
<thead>
<tr>
<th>Scale mean if item deleted</th>
<th>Scale variance if item deleted</th>
<th>Item – total statistics</th>
<th>Corrected item – total correlation</th>
<th>Squared multiple correlation</th>
<th>Alpha if item deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q2M1</td>
<td>58.2763</td>
<td>50.9757</td>
<td>0.4219</td>
<td>0.2406</td>
<td>0.8731</td>
</tr>
<tr>
<td>Q2M2</td>
<td>58.2104</td>
<td>51.1230</td>
<td>0.4992</td>
<td>0.2955</td>
<td>0.8700</td>
</tr>
<tr>
<td>Q2M3</td>
<td>58.6441</td>
<td>50.4027</td>
<td>0.4384</td>
<td>0.2163</td>
<td>0.8727</td>
</tr>
<tr>
<td>Q2M4</td>
<td>58.2094</td>
<td>50.2524</td>
<td>0.5329</td>
<td>0.4369</td>
<td>0.8684</td>
</tr>
<tr>
<td>Q2M5</td>
<td>58.3786</td>
<td>50.6823</td>
<td>0.4677</td>
<td>0.3846</td>
<td>0.8711</td>
</tr>
<tr>
<td>Q2M6</td>
<td>58.9125</td>
<td>49.7394</td>
<td>0.3793</td>
<td>0.1795</td>
<td>0.8781</td>
</tr>
<tr>
<td>Q2M9</td>
<td>58.2606</td>
<td>50.6043</td>
<td>0.5543</td>
<td>0.4471</td>
<td>0.8679</td>
</tr>
<tr>
<td>Q2M10</td>
<td>58.3569</td>
<td>49.4916</td>
<td>0.6535</td>
<td>0.5434</td>
<td>0.8640</td>
</tr>
<tr>
<td>Q2M11</td>
<td>58.3835</td>
<td>49.9867</td>
<td>0.5339</td>
<td>0.3982</td>
<td>0.8683</td>
</tr>
<tr>
<td>Q2M12</td>
<td>58.3176</td>
<td>49.9788</td>
<td>0.5684</td>
<td>0.4118</td>
<td>0.8670</td>
</tr>
<tr>
<td>Q2M14</td>
<td>58.6155</td>
<td>48.3865</td>
<td>0.6309</td>
<td>0.4546</td>
<td>0.8639</td>
</tr>
<tr>
<td>Q3M3</td>
<td>58.7080</td>
<td>48.2680</td>
<td>0.5553</td>
<td>0.4019</td>
<td>0.8675</td>
</tr>
<tr>
<td>Q3M4</td>
<td>58.1701</td>
<td>50.8697</td>
<td>0.6324</td>
<td>0.4725</td>
<td>0.8663</td>
</tr>
<tr>
<td>Q3M6</td>
<td>58.9489</td>
<td>48.0682</td>
<td>0.6124</td>
<td>0.4729</td>
<td>0.8646</td>
</tr>
<tr>
<td>Q3M7</td>
<td>58.5162</td>
<td>51.9764</td>
<td>0.3432</td>
<td>0.1482</td>
<td>0.8763</td>
</tr>
<tr>
<td>Q3M10</td>
<td>58.4307</td>
<td>49.6549</td>
<td>0.6180</td>
<td>0.4311</td>
<td>0.8652</td>
</tr>
</tbody>
</table>

Notes: reliability coefficients 16 items; alpha = 0.8762; standardized item alpha = 0.8825
the research served as a useful method for testing the reliability of the loyalty scale.

Conclusions of the findings

The research findings provide conclusions in relation to Oliver’s (1999) model. Oliver’s (1999) action phase had not been tested empirically until this study. Whilst this research concludes that the action phase antecedents exist in the development of customer loyalty, very few participants exhibited its antecedents. This is evidenced by the lack of inertia, due to situational and mediating effects, which either sustain or render vulnerable the level of customer loyalty development. Therefore, conclusions identify that customer loyalty development is a composite mix of antecedents, sustaining and vulnerability elements. Thus it is the conclusion of this research that loyalty is present only when there is evidence of each of the phases. This may be measured by the loyalty scale, which provides a reliable and valid measure of the level of customer loyalty development based on Oliver’s (1999) hypothetical model. The researchers also confirm that measuring the level of loyalty development is as suggested by Oliver (1999). Levels are a composite mix of phases, which supports Oliver’s hypothesis in relation to whether these three phases may be in synchrony rather than linearly related. In practical terms therefore, the loyalty scale allows managers to identify the most important aspects of their service in relation to the development of their customers’ loyalty. The lack of inertia demonstrated by customers is also an important indicator of their proactive approach. This has implications for managers as it highlights that customers may have a preference, but if an alternative becomes available and customers feel that the preferred company could be doing more to secure their loyalty, the possibility of switching becomes greater. Therefore, many service providers could create a greater level of affective switching costs, which would help to combat the vulnerabilities posed by a new entrant.

An important contribution of the loyalty scale is that it successfully models situational and mediating effects and may be used to identify the most influential sustaining and vulnerability elements affecting each level of customer loyalty development. Knowledge of the situational and mediating effects allows managers to prioritise issues for action within each category of loyalty development. For example, the findings from the focus groups in relation to promotional offers showed how hit and miss these appeared. Use of results from the loyalty scale may include finding out more about customer perceptions of promotional offers, by level of loyalty development.

A further conclusion relates to the analytical perspective of customer loyalty development. Oliver’s model examines customer loyalty development from the perspectives of academics and organisations. Future use of the loyalty scale should consider this bias. This was overcome within this research through the use of focus groups, which provided an analysis of customer loyalty development from a customer’s perspective.

Managerial implications

The research highlights a number of implications for service managers. The first issue specifically relates to the passenger ferry sector. The respondents were well educated in relation to the market, services on offer and competition. Respondents kept up to date on the provision of services and evaluated all aspects of the services. This would suggest that this is a mature, highly competitive market, which points to a need to differentiate customers’ perceptions of the company to a greater extent in contrast to some earlier findings. The loyalty scale may be used to differentiate in conjunction with existing demographic, behavioural or financial data to produce for example correlations matrices, adding value to the existing information held by organisations for operational management.

This research underlined the importance for practitioners of using a combination of research methods in customer research. For example, by identifying customers by level of loyalty development information may be generated on trends within the levels, and followed up with qualitative research such as focus groups to probe and explain trends to gain greater levels of understanding from the perspective of the customer. Focus groups in this study were run nine months after the survey administration to gauge the respondent’s level of loyalty development. The findings highlighted that the majority of respondents remained at the same level of loyalty development. However, vulnerabilities and opportunities to sustain or develop their loyalty also existed at each level. The main area of vulnerability to all levels of participants’ loyalty development was the threat of new competition in terms of the no-frills airline operators. During the nine months interval between the loyalty scale’s administration and the focus group discussions, there was slippage by some participants to a lower level of loyalty development due to dissatisfaction with some elements of the company’s service and persuasion and trial of low-cost alternatives. This group needs to be appropriately managed to reduce the level of defection and poor word of mouth reports. Countering a damaged reputation requires a company to create very appealing and often costly incentives to induce dissatisfied customers back.

The main implication of this finding is to emphasise the importance of sustaining and developing customer loyalty based on a differentiated approach to rewarding customers who have different levels of loyalty development. The findings highlighted the company’s need to acknowledge the importance of reciprocity in terms of which aspects of service customers valued within different levels of loyalty. Supplementing the loyalty scale with focus groups also allows management to be aware of issues, which are being evangelised or recommended by loyal customers, and also the opportunity to ascertain what issues could be improved to promote this further. It is important to remember that customers benchmark not just from what similar service companies are doing, but what the best service providers in general are doing. In this research, participants referred to providers of ferries, airlines, retailers and cruise liners. Most of the items within Factor 1 may be internally controlled, which is good news for managers. Factors 2 and 3 are externally influenced which highlights the importance of managing internal factors well.

The main implication of this research to managers is that the loyalty scale provides an easy to use instrument through which the development of customer loyalty may be measured, in addition to identifying situational and mediating effects. The valid and reliable loyalty scale may also be used within the context of complex services. The research has also added to the services loyalty literature providing a greater level of understanding on how loyalty develops and the importance of comprehending situational and mediating effects.
References


industry findings from Denmark”, Total Quality Management, Vol. 11 Nos 4-6, pp. 544-53.


Appendix

Figure A1 Multi-item loyalty scale

Please answer the following statements by circling the number in the scale which best reflects the strength of your opinion in relation to that statement.

Key

1. When travelling by ferry it is important that I choose the right ferry operator.
2. I know ______ well enough to evaluate it against other ferry operators.
3. ______ is always punctual.
4. Making a reservation with ______ is always straightforward.
5. Information is easily obtained from ______.
6. I regularly take advantage of ______ promotional offers.
7. When deciding on a ferry operator, I am not interested in bargain hunting.
8. When deciding on a ferry operator I compare prices of different companies to be sure I get the best value for money.
9. ______ facilities are usually appealing.
10. In general, ______ is exactly what I need from a ferry operator.
11. ______ as a choice of ferry operator hasn’t worked out as well as I thought it would.
12. If I could do it over again, I’d choose another ferry operator to ______.
13. When choosing a ferry operator, the port is the most important factor.
14. I have a preference for ______ as a ferry operator.
15. I would try an alternative ferry operator if it was £10 less expensive than ______.
16. I would try an alternative ferry operator if the alternative offered increased facilities over ______.
17. I would try an alternative ferry operator if the alternative offered increased status.
18. I would change ferry operator if the alternative offered better facilities.
19. When I see a new ferry service I would try it.
20. I usually travel with the same ferry operator.
21. I enjoy crossing with ______.
22. I would try an alternative ferry operator if ______.
23. I consider myself to be loyal to ______.
24. I would get tired of travelling with ______ every time I travel by ferry.
25. When I travel, I choose the ferry operator that I know best.
26. If I like a ferry operator, I rarely switch from ______ to try something different.
27. I have recommended ______ to my friends.
28. If ______ is not available, it makes little difference to me and I travel with an alternative ferry operator.
Executive summary and implications for managers and executives

This summary has been provided to allow managers and executives a rapid appreciation of the content of this article. Those with a particular interest in the topic covered may then read the article in toto to take advantage of the more comprehensive description of the research undertaken and its results to get the full benefits of the material present.

Constructing a scale to measure customer loyalty development

Finding an accurate measure of customer loyalty is important because it is closely linked with profitability. While there has been much research into the relationship between customer loyalty and quality, satisfaction, profitability and the effectiveness of frequency programmes, there is no instrument capable of measuring customer loyalty development while identifying what is important for sustaining and developing loyalty or rendering it vulnerable. McMullan uses Oliver’s (1999) four-phase model of customer loyalty development as the basis for constructing a scale to measure customer loyalty development:

- **Outline and delineate the construct’s domain.** McMullan advances the view that: purchases are biased or preferred in favour of one alternative over another; it is insufficient to study attitudes in isolation of purchase behaviours within a marketing context; expression of intention of purchases over a period of time will give a temporal indication of the customer’s loyalty to the supplier; the research must focus on a decision-making unit, in this case individual customers; a customer’s loyalty may or may not develop in a sequential way through four phases; and the decision to purchase is a function of an evaluative psychological decision-making process.

- **Develop a set of questions to measure the concept.** The questions consist of a mixture of favourable and unfavourable statements to which respondents are asked to rate their point of agreement or disagreement. The statements are selected to reflect orientation to the attitude of interest. This helps to distinguish between different groups of people and their responses.

- **Trim and refine the pool of items.** This can be done using a panel of experts.

Measuring customer loyalty development in the passenger ferry sector

Three themes emerged when McMullan used the scale for measuring customer loyalty development in the passenger ferry sector:

- **Loyalty sustainers.** These include cognitive items such as choice, punctuality, reservation information and facilities, and affective items such as enjoyment, loyalty and recommendation.

- **Loyalty vulnerabilities: price.** These include price-related items such as bargain hunting and value for money.

- **Loyalty vulnerabilities: service.** One aspect of this is the challenge posed by a new service, such as the arrival of low-cost airlines.

The internal reliability of the loyalty scale was examined by asking participants, face to face, to decide if the scale had correctly categorised their phase of loyalty development, which was then compared to their individual scores. Focus groups took place nine months after the loyalty scale was administered to gauge if the respondent’s level of loyalty development had changed. Analysis of the focus groups found the loyalty scale to be generally reliable. While most respondents remained in the same level of loyalty development, areas of vulnerability had emerged. A few respondents had slipped to a lower level of loyalty development during the nine months’ interval, because of unresolved dissatisfaction with some elements of the company’s service and persuasion to try low cost alternatives. These respondents obviously need to be appropriately managed to reduce the level of defection and poor word-of-mouth reports.

(A précis of the article “A multiple-item scale for measuring customer loyalty development”. Supplied by Marketing Consultants for Emerald.)
This article has been cited by:


3. Arun Kumar Agariya, Deepali Singh. 2014. CRM Index Development and Validation in Indian Banking Sector. *International Journal of Customer Relationship Marketing and Management* **3**:2, 10-32. [CrossRef]

4. Armando Maria Corsi, Sarah Rann Overton, Leonardo Casini. 2014. The impact of the new wine common market organization (CMO) on behavioural loyalty towards product attributes: A case from Italy. *Journal of Consumer Behaviour* n/a-n/a. [CrossRef]


