E-tail Brand Equity: Scale Development and Validation

George Christodoulides*
University of Birmingham, UK

Professor Leslie de Chernatony
University of Birmingham, UK

Dr Eric Shiu
University of Birmingham, UK

Dr Olivier Furrer
University of Nijmegen, The Netherlands

Dr Temi Abimbola
University of Central England, UK

February 2004

School Working Paper Series

The authors would like to acknowledge funding for this research from the CIM

* Corresponding author:
Birmingham Business School, Winterbourne, 58 Edgbaston Park Road,
Birmingham B16 2RT, tel.: +44 121 414 2725, fax: +44 121 414 7791 email:
gxc123@bham.ac.uk

© Christodoulides, de Chernatony, Shiu, Furrer and Abimbola 2004
Abstract
This paper presents the procedures followed in developing and validating a scale to measure e-tail brand equity (eBREQ). Following a comprehensive literature review, an exploratory research study involving depth interviews with experts and consumers was undertaken to identify the facets of brand equity in an e-tail environment. An initial pool of 82 items was generated to reflect these facets. The results from a subsequent web based survey showed e-tail brand equity to be a second-order construct with three correlated but distinct dimensions: emotional connection, online experience, and responsive service nature. A series of psychometric tests demonstrated that the resulting 8-item eBREQ scale is both valid and reliable. The implications of this research for marketing researchers and practitioners are discussed.

Keywords
Brand equity, e-tailing, scale development, validity, reliability

Introduction
Brand equity has been the subject of increasing interest and scholarly investigation for over a decade. Despite writings preaching the demise of brands in the new economy (e.g. Sinha 2000, Dussart 2001), branding takes centre stage in e-tailing, and strategies to build brand equity remain a business priority for marketers as brands continue to grow in importance (Simms 2003, p.18). What renders this intangible so alluring is that it reflects a “storehouse of future profits that result from past marketing activities” (Ambler 2000, p.55, original emphasis).

No attempt to measure brand equity to date has taken into account the internet and its related technologies. Experience with this interactive medium has shown that simply replicating offline marketing effort online is inadequate (Meyers and Gerstman 2001). As is well documented, the internet’s unique characteristics (cf. Hoffman and Novak 1996) bear implications for developing and managing brands (de Chernatony 2001). This does not imply that the principles which have informed brand management for half a century are not relevant today (Porter 2001). A ‘brand’ is a universal concept regardless of setting. What changes online is the enactment of the brand (de Chernatony and Christodoulides 2004). It is therefore postulated that the ways in which brand equity is created online are different from traditional contexts.

Extant brand equity measures have been developed and validated in classical contexts. As yet brand equity measurement has not been re-examined in light of the internet. The purpose of this study is to systematically develop and validate a psychometrically rigorous instrument to measure e-tailers’ brand equity (i.e. the equity of retailers who capitalise on the unique capabilities of the internet to trade online).

The paper opens with a literature review on brand equity. It then describes how exploratory research was conducted in parallel with the literature search
to identify the dimensions of e-tail brand equity. The paper goes on to explain the procedures followed to refine the initial pool of 82 items into the proposed 8-item eBREQ scale and presents a series of tests performed to assess its reliability, validity and unidimensionality of the three dimensions of the scale. The last section highlights the usefulness of the eBREQ scale for practicing marketers and marketing researchers, and makes recommendations for future research.

**Literature Review: Brand Equity**

In spite of the attention brand equity received over the past decade, our understanding of the construct has been impeded by a lack of consensus about its conceptualisation, resulting in a plethora of diverse methodological approaches to its measurement. As Berthon et al. (2001, p.1) note, “perhaps the only thing that has not been reached with regard to brand equity is a conclusion.”

Brand equity can be analysed on two levels, depending on the beneficiary of value (firm/consumer). Marketing research has largely concentrated on consumer based brand equity (hereafter CBBE) as opposed to firm based brand equity. This is because the consumer based approach offers insights into consumer behaviour which can be easily converted into actionable brand strategies (Keller 1993). Conversely, the firm based approach centres around financial valuation, and provides little usable information for managers. As a result, significant advances have been made in terms of CBBE’s conceptualisation (e.g. Keller 1993, Erdem and Swait 1998), measurement (e.g. Park and Srinivasan 1994, Yoo and Donthu 2001a, Vazquez et al. 2002, Netemeyer et al. 2004), and validation of measuring instruments (Agarwal and Rao 1996, Mackay 2001, Washburn and Plank 2002).

Current knowledge of CBBE has evolved from two paradigms: cognitive psychology and signalling theory in information economics (Czellar and Dennis 2002). The dominant stream of research has been grounded in cognitive psychology, focusing on memory structure (Aaker 1991, Keller 1993). Aligning with the psycho-cognitive framework, Keller (1993, p.2) defined CBBE as “the differential effect of brand knowledge on consumer response to the marketing of the brand.” According to this conceptualisation, a brand is positively valued when the customer reacts more favourably to the marketing of a product with a known brand name compared to an identical but unbranded product. Brand knowledge is conceptualised as an “associative network memory model” consisting of two dimensions: brand awareness and brand associations in consumer memory. Positive CBBE is yielded when the consumer is aware of the brand and also has strong, unique and favourable brand associations.

In parallel, brand equity research rooted in information economics takes into consideration the imperfect and asymmetrical nature of markets. In light of this, economic agents are required to transmit information about their specific characteristics by means of signals. According to Erdem and Swait (1998), brand names act as signals to consumers. From this perspective, a brand signal becomes the sum of that brand’s past and present marketing activities.
Imperfect and asymmetrical market information produces uncertainty in consumers’ minds about available products and services. A credible brand signal generates customer value by: (1) reducing perceived risk, (2) reducing information search costs, and (3) creating favourable attribute perceptions (ibid.). Under this paradigm, CBBE is consequently defined as “the value of a brand signal to consumers” (ibid, p.140).

Hitherto little research was directed towards understanding brand equity in an e-business context. In their small scale consumer study, Michel and Vergne (2002) noted that brand equity is based on different dimensions depending on the type of the brand’s site (informational versus transactional). In another study, Page and Lepkowska-White (2002) drew on Keller’s (1993) framework to identify four categories of factors which drive ‘web equity’ via awareness and image (i.e. marketer and non-marketer communications, site design, vendor characteristics, and product/service characteristics). These factors are product of the literature alone and need to be subject to empirical examination. No study to date has explored brand equity measurement in an e-business context.

Scale Development
As no research has produced a valid, reliable and parsimonious scale to measure e-tail brand equity, the current research sets out to fill this gap by developing and testing the eBREQ scale. The development of the scale followed the iterative procedures suggested by Churchill (1979), Gerbing and Anderson (1988), and Netemeyer et al. (2003) in conjunction with Aaker’s (1996) guidelines on how to devise a brand equity measurement system for a particular context.

The process began with a literature review on brand equity. From this, we adopted the widely quoted MSI definition of brand equity, i.e. “a set of associations and behaviours on the part of a brand’s consumers, channel members and parent corporation that enables a brand to earn greater volume or greater margins than it could without the brand name and, in addition, provides a strong, sustainable and differential advantage” (Srivastava and Shocker 1991, p.5). This definition aligns with our views on the ontology of brands which fall within the scope of the ‘holistic’ approach as opposed to the ‘product plus’ approach to branding (Styles and Ambler 1995). Under the ‘product plus’ approach, the brand is considered to be an addition to the product (ibid). In contrast, proponents of the ‘holistic view’ contend that brands supervene on products without, though, being confined to the product space (Grassl 1999). Brands reflect the totality of all marketing mix elements (Ambler and Styles 1997).

An exploratory stage was undertaken to identify the facets of brand equity in e-tail space. Aaker’s (1996) Brand Equity Ten served as a starting point as it represents a comprehensive set of equity manifestations in classical contexts. To supplement this list with equity manifestations appropriate for an e-tail environment, 16 depth interviews with brand experts were conducted, revealing 10 additional facets, namely brand experience, interactivity, customisation, relevance, design, customer service, order fulfilment, quality of
brand relationships, communities, and site logs (Christodoulides and de Chernatony 2004). Following the interviews with experts, two focus groups were carried out to explore the ensuing facets of e-tail brand equity through consumers’ own descriptions.

Out of the 20 candidate facets of e-tail brand equity, market share, market price, distribution coverage, and site logs represent indirect manifestations of CBBE in the sense that information cannot be collected directly from consumers. As individual scores on these facets would be meaningless for a consumer based exercise, a decision was made to exclude them from the instrument. The literature was again consulted and items from validated scales were used, when available, to form an initial pool of 82 items. All items were evaluated with 7-point Likert scales anchored at 1=“strongly disagree” and 7=“strongly agree.” Scale items were randomly ordered to avoid any systematic order effect or cluster answering effect.

In line with the nature of this research, a web based questionnaire was developed to facilitate data collection. A personalised email was sent to a sample of 5,000 UK internet shoppers inviting them to participate in the survey. All target subjects were registered members of ipoints™, a UK online reward scheme. As an incentive for completing the survey, respondents were offered: a donation of 20p to their preferred charity, 25 ipoints™, and a summary of the results. Elimination of incomplete data resulted in 375 usable surveys (7.5% response rate).

There are three possible explanations for this low rate, which nonetheless is not atypical in web surveys; response rates as low as 4% have been reported in the literature (cf. Grandcolas et al. 2003). First, as the novelty effect of the internet has started wearing off, internet users have become apathetic towards online research, leading to declining response rates (Wilson and Laskey 2003, McDonald and Adam 2003). Second, evidence suggests that internet users have become victims of email overload, forcing them to delete up to 70% of their messages based solely on the subject line (Burkeman 2001). Third, users’ increasing use of sophisticated software tools such as automatic filtering of unwanted email messages may have prevented the invitation from reaching all inboxes. To test for the possibility of non-response bias, Armstrong and Overton’s (1977) recommended procedure was adopted. This assumes that late respondents in a sample are similar to theoretical non-respondents. Using a series of $t$- and $\chi^2$-tests to compare early and late respondents on their demographic characteristics and internet habits, no significant differences appeared between the two groups.

Evaluation procedures and analyses
Prior to any analysis, the sample of 375 respondents was split in half using the random sample function of SPSS. The first half ($n_1=188$) was used to develop the scale, while the second half ($n_2=187$) was used to validate the results (Churchill 1979).
**Item analysis**

Following Churchill’s (1979) paradigm, the first step in purifying the instrument was to calculate coefficient alpha and item-to-total correlations to delete “garbage items.” The sub-scales of leadership, personality and communities failed to achieve acceptable levels of internal consistency (<0.7 recommended by Nunnally 1978) and their items were deleted. Next, an item was deleted from a sub-scale if its item-to-total correlation was below 0.4 (Wang et al. 2001). A further 6 items were removed. Reliabilities of the remaining sub-scales ranged from 0.71 for loyalty to 0.90 for design.

**Exploratory Factor Analysis (EFA)**

A series of exploratory factor analyses was subsequently performed on the remaining 66 items to identify the factor structure of e-tail brand equity. The suitability of the data set for EFA was examined using the KMO statistic and Barlett’s test of sphericity. The KMO statistic which measures the overall sampling adequacy of the factor analysis produced a value of .92 which falls into the range of being ‘superb’ (Field 2000, p.455). Barlett’s test yielded an $\chi^2$ value of 9710.6 (p<0.001) which suggests that the intercorrelation matrix contains sufficient common variance to render factor analysis appropriate.

The sample data was then examined using principal components analysis as the extraction technique and varimax as the orthogonal rotation method. To improve unidimensionality and discriminant validity through EFA, four commonly employed decision rules were initially applied to identify the factors underlying e-tailers’ brand equity: (1) using the latent root criterion as a cut-off value for extraction; (2) deleting items with insignificant factor loadings at the 0.05 level ($\leq 0.45$ for our sample size) (cf. Hair et al. 1998, p.112); (3) deleting items with significant factor loadings ($\geq 0.45$) on two or more factors; and (4) excluding single item factors from the standpoint of parsimony.

The outcome of EFA suggested an 11-factor solution, accounting for 68.6% of the variance. As emerging factors comprised as many as 16 items, EFA was repeated to reduce the items to a more tractable number. A more stringent criterion specifying that items with loadings less than 0.67 (the square root of 0.45) on a given factor be deleted was introduced in lieu of decision rule 2 from above (Shimp and Sharma 1987). Out of 66 items, 22 survived this process, loading on 5 distinct factors. Based on the shared meaning among the items of each factor, the 5 factors were labelled as emotional connection, online experience, responsive service nature, order fulfilment and trust; reliability coefficients were 0.91, 0.91, 0.78, 0.82, and 0.86 respectively.

**Confirmatory Factor Analysis (CFA)**

Following Gerbing and Anderson’s (1988) advice, the remaining 22 items were examined via CFA to establish the unidimensionality of each emerging factor. A measurement model was thus specified to have the retained 5 factors (latent variables) with the 22 selected items (manifest variables), and each item was prescribed to load on only one factor according to the structure indicated in EFA. The factor analytic model was estimated using the LISREL 8.30 maximum likelihood method (Jöreskog and Sörbom 1996).
The measurement model turned out to be a poor representation of the data, with fit indices failing to meet acceptable levels (cf. Hu and Bentler 1999). The $\chi^2$ test was 462.44 ($p<0.001$) with 199 degrees of freedom. Goodness of fit index (GFI) and adjusted goodness of fit index (AGFI) were 0.82 and 0.77 respectively. The comparative goodness of fit indexes measured by the comparative fit index (CFI), incremental fit index (IFI), and non-normed fit index (NNFI) were 0.90, 0.90 and 0.89 respectively. The root mean square error of approximation (RMSEA) was 0.084 (the 90% confidence interval for the RMSEA was 0.074-0.094).

**Model re-specification**

In order to detect misfitting parameters and achieve a clear factor structure with unidimensional factors only, CFA was initially used in an exploratory manner (cf. Netemeyer et al. 1996, Lastovicka et al. 1999, Yoo and Donthu 2001b). Scale items exhibiting high modification indices or residuals were subsequently removed from the variable list. These may *inter alia* stem from model misspecification, non-normally distributed data or nonlinear relationships among some variables and adversely affect the overall model fit (Jöreskog and Sörbom 1996). Significant cross-loading items were dropped for three reasons. First, the objective of this research was to develop a valid, reliable and parsimonious eBREQ scale of multiple independent dimensions, and this required a clean factor structure. Second, removing cross-loadings helps factor interpretation. Third, when not underpinned by theory, cross-loadings may be attributed to a statistical artefact (Yoo and Donthu 2001b).

Following a series of iterative procedures, the 8-item eBREQ measurement model was supported by values of fit. The Satorra-Bentler $\chi^2$ statistic was 15.84 ($p=0.54$) with 17 degrees of freedom. GFI and AGFI were 0.96 and 0.92 respectively. CFI, IFI and NNFI were 0.99, 0.99 and 0.98 respectively. Item loadings on their corresponding dimensions ranged from 0.64 to 0.94 as shown in Table 1. The smallest $t$-value of the loadings was 9.20, which indicates highly significant loadings.

**Scale Validity and Reliability**

A CFA was performed on the second sample ($n=187$) to test the underlying structure of the items from the previous analysis. The 3-dimensional eBREQ measurement model was re-evaluated through LISREL 8.30, and showed satisfactory fit. The model’s Satorra-Bentler $\chi^2$ was 16.02 ($p=0.52$) with 17 degrees of freedom, and its relative $\chi^2$ was 0.94, which is <2 as recommended by Carmines and McIver (1981). RMSEA and SRMR were 0.00 and 0.033 respectively. CFI and IFI were 0.99 each and NNFI was 0.98. GFI and AGFI were 0.96 and 0.92 respectively. These fit indices exhibit an excellent level of fit with the model in the new data (Hu and Bentler 1999). The item loadings to their constructs ranged from 0.59 ($t$-value=6.43) to 0.93 (21.50) as shown in Table 1.
Figure 1: eBREQ Conceptual Model

Composite reliability estimates which are measures of internal consistency, analogous to Cronbach’s alpha, were next computed from LISREL results (Fornell and Larcker 1981). These were 0.89 for emotional connection, 0.85 for online experience and 0.69 for responsive service nature, which exceeded the minimum recommended level of 0.60 (Bagozzi and Yi 1988).

Having established the internal psychometric properties of the eBREQ scale, the next step involved an assessment of its validity. Ping (2004) recommends that validity be determined on the basis of the following minimum criteria: content validity, criterion validity, and construct validity (including discriminant and convergent validity).

Table 1: eBREQ Items: A LISREL Completely Standardised Solution
As care was taken with the procedures followed (cf. Churchill 1996) involving exploratory research with both experts and consumers, the eBREQ scale is considered content valid.

An aggregate score of e-tail brand equity was then required to assess the criterion and construct validity of the eBREQ scale. Simply summating the raw scores of the 8 items would not be appropriate as the 3 dimensions would likely not contribute equally to e-tail brand equity. For this purpose, a higher-order eBREQ model was fitted whose 3 dimensions are related to a higher-order factor, labelled higher-order e-tail brand equity. Causal paths of this higher-order model to the 3 dimensions were used as weights to create an index for e-tail brand equity (Yoo and Donthu 2001a). The eBREQ model and the higher-order model were statistically equivalent and their fit indices were thus identical. Path coefficients were 0.66 for emotional connection, 0.54 for online experience and 1.08 for responsive service nature. The weight of a dimension was calculated as the fraction of the path estimate of that dimension over the sum of the three path estimates. For example, the weight of emotional connection is 0.29, which derives from $0.66/(0.66+0.54+1.08)$. As a result, the eBREQ index is equal to 0.29 (the mean of emotional connection) + 0.24 (the mean of online experience) + 0.47 (the mean of responsive service nature).

The scale’s criterion validity was assessed by correlating the eBREQ index to an independent 3-item overall brand equity measure (OBE) adapted from Yoo and Donthu (2001a). The internal reliability of OBE was 0.90. The two were correlated at 0.56 ($p<0.01$) indicating criterion validity. Convergent validity was gauged through the average variance extracted (AVE) that measures the percentage of total variance of the data accounted for by each construct. AVE for the three factors was 0.74 for emotional connection, 0.66 for online experience, and 0.53 for responsive service nature as indicated in Table 2. All of them exceeded the 0.50 threshold advocated by Fornell and Larcker (1981).

Evidence of discriminant validity among the dimensions of e-tail brand equity was provided by three separate tests. The first test compares the AVE for each factor with the squared pairwise correlation between factors. In order to establish discriminant validity, the AVE for a given factor should be higher than all $\phi^2$ estimates involving that factor (Fornell and Larcker 1981). The results provide good support of discriminant validity as shown in Table 2.

<table>
<thead>
<tr>
<th>Standardised correlation between factors</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional connection and Online experience</td>
<td>0.35 (0.12)$^a$</td>
</tr>
<tr>
<td>Online experience and Responsive service nature</td>
<td>0.58 (0.34)</td>
</tr>
<tr>
<td>Emotional connection and Responsive service nature</td>
<td>0.71 (0.50)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Average variance extracted for each factor</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional connection</td>
<td>0.74</td>
</tr>
<tr>
<td>Online experience</td>
<td>0.66</td>
</tr>
<tr>
<td>Responsive service nature</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Note: $^a$Numbers in brackets represent squared pairwise correlations

Table 2: Standardised Correlation Matrix and AVE
The second test calls the scale developer to examine each correlation between all pairs of factors within the eBREQ measurement model (Anderson and Gerbing 1988). In order for two factors to be independent, their pairwise correlation should be smaller than 1. This was tested by constructing a 95% confidence interval around each correlation. Confidence intervals were (0.17, 0.53) for emotional connection and online experience; (0.41, 0.75) for online experience and responsive service nature; and (0.58, 0.83) for emotional connection and responsive service nature. None of the confidence intervals subsumed the value 1, supporting the eBREQ scale’s discriminant validity.

The third test compares the $\chi^2$ statistic among the 3-factor eBREQ model and alternative measurement models with fewer factors (see Table 3). Evidence of discriminant validity exists if the $\chi^2$ of each unconstrained model (with more factors) is significantly lower than the $\chi^2$ of each constrained model (model with fewer factors). The smallest $\chi^2$ difference between the 1-factor and the 2-factor models was 18.40 (p<0.001). The improvement in $\chi^2$ from each of the 2-factor models to the 3-factor eBREQ model was 182.70, 56.05 and 28.34 respectively. This suggests that treating the individual dimensions as distinct factors is superior to aggregating the dimensions.

Confidence in a scale increases if it behaves as expected in relation to other constructs. The construct validity of eBREQ was further assessed vis-à-vis three independent constructs, namely attitude towards the service provider (e-tailer), consistent image, and purchase intention. The attitude towards the e-tailer was captured by means of 3-bipolar items adopted from Bruner et al. (2001). A 3-item modified measure from Loiacono et al. (2002) was used to quantify consistent image. The two scales achieved reliabilities of 0.96 and 0.89 respectively. Purchase intention was measured through a single item Graeff (1997). The correlation of the eBREQ index with the constructs was significant at the 0.01 level: 0.59 with attitudes towards the e-tailer, 0.34 with purchase intention and 0.42 with consistent image. In addition, the correlation between individual eBREQ dimensions and the relevant constructs was consistent and significant. This is good confirmation of the scale’s construct validity.

<table>
<thead>
<tr>
<th>Competing Models</th>
<th>$\chi^2$</th>
<th>df</th>
<th>$\chi^2$/df</th>
<th>$\Delta\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>One factor model</td>
<td>217.12***</td>
<td>20</td>
<td>10.86</td>
<td>NA</td>
</tr>
<tr>
<td>Two factor correlated model ($\phi_{EC:OE}$ = 1)</td>
<td>198.72***</td>
<td>19</td>
<td>10.46</td>
<td>18.40***</td>
</tr>
<tr>
<td>Two factor correlated model ($\phi_{OE:RSN}$ = 1)</td>
<td>72.07***</td>
<td>19</td>
<td>3.79</td>
<td>145.05***</td>
</tr>
<tr>
<td>Two factor correlated model ($\phi_{EC:RSN}$ = 1)</td>
<td>44.36***</td>
<td>19</td>
<td>2.33</td>
<td>172.76***</td>
</tr>
<tr>
<td>Three factor correlated model</td>
<td>16.02</td>
<td>17</td>
<td>0.94</td>
<td>182.70***a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>56.05***b</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28.34***c</td>
</tr>
</tbody>
</table>

Note: EC= Emotional connection; OE= Online experience; RSN= Responsive service nature

*a Chi-square difference over the two-factor model (EC & OE combined)

*b Chi-square difference over the two-factor model (OE & RSN combined)

*c Chi-square difference over the two-factor model (EC & RSN combined)

*** p<0.001

Table 3: Comparison of Competing Models
Discussion
The results of the various tests demonstrate that the eBREQ scale possesses strong psychometric properties, and confirm that emotional connection, online experience and responsive service nature constitute 3 independent yet correlated dimensions of e-tail brand equity. Consisting of only 8 items (cf. Miller 1956), the scale is parsimonious and is of value to practitioners and researchers.

The eBREQ scale can serve as a diagnostic for e-tail marketers to track the equity of their brands. By assessing individual dimensions of e-tail brand equity, marketers are able to identify areas of strength and weakness. By extending this consumer based exercise to include competing brands, marketers can make concomitant adjustments to their positioning strategies in order to enhance their e-tail brand’s competitive stance. E-tail marketers can also use the eBREQ scale to monitor the long-term impact of various marketing activities on their brand equity to optimise the effectiveness of budget allocation. Furthermore, the eBREQ scale can help marketers gain insights into the relationship between their e-tail brand performance and bottom line measures such as profit and sales.

In addition to its managerial usefulness, the eBREQ scale enables researchers to advance marketing theory in the areas of online branding and retailing. More specifically, through using the eBREQ scale researchers can evaluate how different marketing strategies impact on e-tailers’ brand value. With the aid of the eBREQ scale, researchers can also demonstrate how marketing expenditure for e-tail brands is an investment that results in quantifiable outcomes. One limitation of the current study is that it was conducted exclusively in the UK. Other researchers are encouraged to undertake replication studies in different countries to assess the scale’s equivalence across nations. The present study focused on business-to-consumer shoppers and researchers may wish to explore whether the eBREQ scale holds for business-to-business customers.
References:


