

Open Research Online

The Open University's repository of research publications and other research outputs

An assessment of the validity and dimensionality of E-S-QUAL as a measure of internet retail service quality

Conference or Workshop Item

How to cite:

Rafiq, Mohammed; Fulford, Heather; Lu, Xiaoming and Ellis-Chadwick, Fiona (2008). An assessment of the validity and dimensionality of E-S-QUAL as a measure of internet retail service quality. In: Marketing: Shifting Focus from Mainstream to Offbeat, ANZMAC Conference., 1-3 Dec 2008, Sydney, Australia.

For guidance on citations see [FAQs](#).

© 2008 The Authors

Version: Accepted Manuscript

Link(s) to article on publisher's website:
<http://www.anzmac2008.org/>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's [data policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

An Assessment of the Validity and Dimensionality of E-S-QUAL as a measure of Internet Retail Service Quality

by

Mohammed Rafiq, Heather Fulford, Xiaoming Lu and Fiona Ellis-Chadwick

Abstract

Despite its acknowledged importance, there are few rigorous empirical studies examining internet retail service quality. An exception is the development of the E-S-QUAL scale by PZM (2005). Whilst E-S-QUAL demonstrates excellent psychometric properties, the scale is in need of external validation. We present a reassessment and validation of the E-S-QUAL in the internet grocery sector. Our results show that there are potential discriminant validity problems with the Efficiency and System Availability dimensions. Our analysis also shows problems of high correlation between Privacy and Fulfilment dimensions. This may be because privacy is less important for frequent users (such as grocery shoppers). Consistent with Parasuraman et al, our results suggests that website related factors are the most important in the overall evaluation of internet retail service quality.

Introduction

With the increasing importance of Internet retailing, service quality in the online environment has been recognised as an important factor in determining the success or failure of e-commerce ventures (Santos, 2003; Yang, 2001). A number of existing studies on e-service quality have attempted to identify the elements that define customers' perception of service quality, and to build models that outline the differences between customers' expectations and the real service experience (Janda et al., 2002; and Zeithaml et al., 2002). Much of the early empirical research on internet retail service quality, the focus of this study, concentrated on developing measures for the evaluation of Web sites, or process quality. However, Wolfinbarger and Gilly (2003) argue that measuring e-service quality should go beyond the website interface. This is because a customer's online buying experience consists of everything from information search, product evaluation, decision making, making the transaction, delivery, returns and customer service. It is apparent that measures for evaluating just websites may not be sufficient for measuring service quality across various stages of the service delivery online. This is also in line with Parasuraman et al's, (2005, henceforth PZM) study who state that the purpose of developing scales for e-service quality is to measure the whole experience of customers regarding the service received rather than to generate information for website designers.

There have been relatively few rigorous empirical studies examining internet retail service quality to date. Yoo and Donthu (2001) using a sample of students, developed the SITEQUAL scale to measure the perceived quality of an Internet shopping site. This resulted in a nine-item scale with four dimensions namely, ease of use, aesthetic design, processing speed, and security. Students were asked to visit evaluate each site and therefore did not include all aspects of online buying particularly fulfilment and therefore did not include all aspects of e-service quality. Loiacono et al (2002) developed the WEBQUAL scale with 12 dimensions namely, informational fit-to-task, interactivity, trust, response time, ease of understanding, intuitive operations, visual appeal, innovativeness, flow-emotional appeal, consistent image, online completeness, and better than alternative channels.. PZM (2005) question the dimensionality of WebQual, in addition, it does not include a fulfilment dimension. Nor had the student sample necessarily made a purchase off the website they were

evaluating. Barnes and Vidgen (2002,) developed and refined a scale, also called WebQual. The scale provides an index of a web site's quality based on customer perceptions weighted by importance of perceived characteristics and has five factors namely usability, design, information, trust, and empathy. The scale does not provide a comprehensive evaluation e-service quality as it does not require the respondents to have completed the purchasing process, they are asked instead to visit and to collect information on a book of their choice on one of three online book stores and to rate the site. Its specificity to book buying and the fact that it is based on a convenience sample of university students and staff limits its applicability. Janda, Trocchia and Gewinner (2002) develop an internet service quality scale consisting of five dimensions, performance (expectations regarding order fulfilment), access (the retailer's ability to provide a variety of products from across the world), security (financial security and privacy), sensation (website features) and information. Reliability of their sensation dimension (with a Cronbach's alpha score of 0.61) is questionable and a couple of the items of access dimension (namely availability of world wide products and the support of multiple languages and currencies) are not likely to be applicable to all internet retailers.

Based on an extensive and systematic scale development process, including online and offline focus groups, and an online survey of a customer panel, Wolfinbarger and Gilly (2003), developed the eTailQ scale. Their analysis suggests eTailQ has four dimensions namely-website design, fulfilment/reliability, privacy/security, and customer service. However, whilst recognising the rigour of the development procedures, PZM question the consistency and distinctness of the privacy/security, and customer service dimensions. For instance the customer service dimension includes, items relating to the company's willingness to respond to customer needs, the company's interest in solving problems, as well as the promptness with which inquiries are answered.

Based on their preliminary work (Zeithaml et al. 2000) arguing that in any assessment of internet service quality, the focus ought to be on all cues and encounters that occur before, during, and after the transaction, PZM developed a 22-item scale with four dimensions namely, Efficiency (8 items), System availability (4 items), Fulfilment (7 items), and Privacy (3 items). Whilst PZM. meticulously followed procedures for developing new scales and the scale demonstrates excellent psychometric properties, the scale is need of external validation. To date we are aware of only one attempt at rigorous external validation of E-S-QUAL by Boshoff (2006). Boshoff concludes that E-S-QUAL captures the essence of e-service quality, its dimensionality, however, it needs further examination. This is because in Boshoff's study a six factor scale provided a better fit to the data than the original four factor E-S-QUAL scale. However, because the respondents in Boshoff' survey were purchasers of books, DVDs, CDs, and other gifts , it is possible that Boshoff's results are limited to infrequently/ occasionally purchased relatively low value items.

Given the importance of reliable and valid research measures, we present a reassessment and validation of E-S-QUAL in the internet grocery sector. Online grocery shopping is an interesting area for testing E-S-QUAL because shopping for groceries is an activity that is repeated at regular time intervals, therefore online service quality is likely to be even more important because of the frequency of the transactions and the amount that customers spend on groceries. We also examine whether E-S-Qual is second order construct which PZM hint at but do not provide full details for in their paper. Furthermore, there is emerging evidence of the importance of service quality as an antecedent of customer satisfaction and customer and loyalty in online retailing (see for instance, Shankar et al 2003; Anderson and Srinivasan, 2003; and Gefen 2002), the study therefore also tests the nomological validity of E_SQUAL in this context.

Research Method

Given the problems with the service recovery part of their scale, E-RecS-Qual, and that PZM regarded their findings on E-RecS-Qual as preliminary, this paper focuses on validating E-S-QUAL. We therefore used PZM's, (2005) 22-item four dimensional E-S-QUAL scale in our study. However, instead of a 5 point Likert scale used in the PZM study we adopted a seven point Likert scale. A seven point scale extends the range and variability of responses and potentially increases the reliability of the results. To measure customer satisfaction in the online environment we adopted the measure from Jones and Suh's (2000) study. Loyalty was measured using Zeithaml, Berry, and Parasuraman (1996) 5 item scale. Both these scales also employed seven point Likert scales for the reasons mentioned above.

The data presented in this research were collected via a web-based survey using self-administered questionnaires. The questionnaires were distributed to online grocery shoppers by the use of an Internet panel administered by a market research company. A total of 519 responses were received within a week and 491 questionnaires remained for further analysis after data screening. The relatively large sample size provides some confidence that the sample is representative of online grocery shoppers. In this survey 49.3 per cent of the respondents were males and 50.7% females. There were 45 (9.2%) respondents who were less than 25 years; 186 (38.2%) were between 25-40; 190 (39 %) were between 41-55; 40 (8.2%) were between 56-60; and 26 (5.3%) were over 60 years old. Almost 50% of e-grocery shoppers' annual salary is over £30,000 (compared with £23,244 UK average for 2006 (Office of National Statistics)), 15% between £25,000-29,999 and 11.3% between £20,000-24,999. The above average income is representative of internet shoppers in general and grocery shoppers in particular (see for example Mintel 2007). In the sample, 69.7 per cent shopped online with Tesco, 14.5 per cent with ASDA, 10.6 per cent with Sainsbury's, 0.6 per cent with Waitrose and 4.7 per cent with a number of smaller operators such as Ocado and Foodferry. We believe this to be an accurate reflection of the market share of the online food retailers in the summer of 2006. On average, 64 per cent of the respondents had shopped with the offline store before trying the online store. That is, around 64 per cent of the respondents were transferring their loyalty from their offline store to the online store. The most loyal were Tesco shoppers, of whom 93 per cent had shopped with the offline store before shopping with the online store.

Results

As recommended by Garver and Mentzer (1999), we began the analysis by looking at each of the dimensions separately to assess whether each of the items loaded on the dimension that it was supposed to, and to assess the unidimensionality of the constructs. All the items loaded quite well on the appropriate dimensions. The standardised loadings ranged from 0.79 to 0.88 for the Efficiency Dimension, 0.78 to 0.85 for System Availability, 0.63 to 0.85 for Fulfilment and 0.76 to 0.85. However, two items of the Fulfilment dimensions had correlated errors; namely item FUL 1: It delivers orders when promised; and item FUL 7: It makes accurate promises about delivery of products. On reflection these items do appear to be very close and therefore item FUL 7 was arbitrarily deleted as both items had the same standardised loadings. The high standardised loadings suggested that the constructs are unidimensional. The coefficient alphas ranged from 0.83 for the Privacy dimension to 0.96 for the Efficiency dimension and are similar to the PZM. As all the coefficient alphas exceeded the minimum value of 0.7, this suggests that the measures are reliable. Given this, we proceeded to the next stage of conducting CFA by looking at all the dimensions together. The standardised loadings were all above 0.7 except for item FUL 5 which had a loading of 0.63. Overall, the standardised loadings show a similar pattern to those obtained by PZM. The goodness-of-fit statistics are also similarly good. In our data, however, RMSEA is significantly lower at 0.08

compared with 0.11 in PZM data. The inter factor correlations ranged from 0.54 to 0.88 (see Table 1)). In the PZM study the inter-correlations ranged from 0.62 to 0.77. The high standardised loadings together with the high coefficient alphas provide support for E-S_QUAL's convergent validity. To assess discriminant validity we followed the same procedure as PZM and fixed each inter-factor correlation, one at a time, and redid the CFA and examined the difference in the chi-square statistic for the original and the constrained models. Each of the CFAs produced a significant $\Delta\chi^2$ values (with 1 df) except for when the correlation between Efficiency and System Availability was constrained to 1. This resulted in chi-square difference value of 1.6 which is not statistically significant for 1 df, suggesting that there is problem with discriminant validity for the Efficiency and System Availability measures. This is not surprising given that the high correlation between the two constructs of .89 (see Table 2). However, the finding does contrast with that of PZM who found that discriminant validity was supported for all four of their factors.

TABLE 1: Inter- factor correlations between E_S_QUAL Dimensions

	Efficiency	Systems Availability	Fulfilment	Privacy
Efficiency	1.00	.77	.68	.62
Systems Availability	.878	1.00	.68	.64
Fulfilment	0.633	.588	1.00	.65
Privacy	0.585	.535	.813	1.00

Note: The inter-factor correlations for the UK study are shown below the diagonal and above the diagonal inter-correlations for Wal-Mart in the PZM study

The high inter-factor correlations in both studies suggest that a second-order factor structure underlies E-S-QUAL. PZM do not report the findings of the second order E-S-QUAL model except to say that the factor loadings and fit statistics were similar to the first order model. Notwithstanding the problem mentioned above, we estimated a second order measurement model for E-S-QUAL. Although the second-order factor loadings were high (Efficiency had a factor loading of .94, System Availability .91, Fulfilment.7 and privacy .65), the initial estimated model resulted in a poor level of fit. The observed χ^2 for this model is 922.60 and the χ^2/df ratio at 4.99 exceeded 3 as recommended by Bagozzi and Yi (1988). The goodness-of-fit index (GFI) and adjusted GFI (AGFI) at 0.84 and 0.80, were much lower than the recommended level of 0.90. In an effort to address the problems we examined the modification indices (MI). Firstly, two of the Efficiency items (EFF 5 and EFF7) had errors that were correlated with the residual of the Systems Availability dimension, and as correlations with residuals are substantively uninterpretable, the model was re-estimated with the error covariance of EFF5 and EFF6 specified as free parameters. Examination of the MI also showed evidence of covariance between the residuals of Fulfilment and Privacy dimensions and their regression weights also showed evidence of cross loading. Such misspecification means that the Fulfilment items could measure Privacy or vice versa. In Wolfinbarger and Gilly's (2003) study privacy not a significant factor in predicting e-service quality. This may be the case in this study due to the high correlation between Fulfilment and Privacy. Also, Efficiency, Systems Availability and Fulfilment could be regarded as core elements of the online grocery service, whilst Privacy is an augmented part of the service. Therefore, we re-estimated the model with Privacy specified as free parameter.

The re-specified full measurement model yields an overall χ^2 value of 373.01 with 101 degrees of freedom and $\chi^2/df=3.69$). The GFI and AGFI are at 0.91 and 0.88. Although the

value of AGFI is still below the recommended level of 0.90, Other goodness- of-fit indices are IFI (0.96), CFI (0.96), TLI (0.95) and RMSEA (0.08) appears to be adequate. All parameter estimates are statistically significant this time.. The substantial improvement in the model fit between the initial 4 dimension model and the final 3 dimensional model suggests that, Efficiency, Systems Availability and Fulfilment are most appropriate for estimating E-S-QUAL for our data. All the model loadings were statistically significant and the second order loadings of the three dimensions are high. Efficiency had a regression weight of .95, System Availability .88, and Fulfilment .67. These findings together suggest that E-S-QUAL is a second order construct.

We tested the nomological validity of the resulting second order model by modelling E-S-QUAL as antecedent of customer satisfaction and customer satisfaction as an antecedent of customer loyalty in a structural model. E-S-QUAL was modelled as second order latent construct with three dimensions. As previously mentioned, we measured customer satisfaction adoption a measure of online satisfaction employed by Jones and Suh (2000) and loyalty was measured using Zeithaml, Berry, and Parasuraman (1996) 5 item scale. The model showed very good fit. The observed χ^2 for this model is 740.01 and the χ^2/df ratio at 2.99 just below 3 as recommended by Bagozzi and Yi (1988). The CFI at .952, NFI at .90, RFI at .92 and TLI at .95 are all above the recommended level of 0.90. RMSEA at 0.06 also suggests evidence of good fit. Furthermore, the standardised regression weight between E-S-QUAL and customer satisfaction was 0.56 and that between customer satisfaction and customer loyalty was 0.69. This provides strong support for the nomological validity of the three dimensional, second order formulation of E-S-QUAL.

Discussion

E-S-QUAL provides a useful starting point for assessing e-service quality. However, our study has shown that E-S-QUAL needs further refining. Our results show that there are potential discriminant validity problems with the Efficiency and System Availability Dimension. The high correlation between the dimensions suggests that they are too closely correlated and therefore need more refinement to improve their discriminant validity. Our analysis also showed that potential problem of high correlation with the Privacy and Fulfilment dimensions. This may be because , as Wolfinbarger and Gilly(2003) suggest that, privacy is less important for frequent users (such as grocery shoppers) as experience may allay concerns regarding website security and privacy issues. In our study ,as grocery shopping is a frequent activity and 72 per cent of the respondents shopped at least once or more times a month online for groceries, and 69 per cent of the respondents had been shopping with their e-grocer for six months or more. Interesting the Efficiency dimension is the most service quality dimension both in the PZM and our study. This suggests that website related factor are more important in the overall service quality outcome. The slightly lower ranking of fulfilment may suggest that overall customers are relatively happy with this dimension rather than its lack of importance. Together this suggests that e-tailers should be putting more of their efforts in the website related factors in the E-S-Qual model. The findings of this study are limited to the e-grocery market and therefore E-S-Qual's psychometric properties need to be tested and validated in other retail contexts in order to arrive at a more comprehensive evaluation of the validity of E-S-QUAL. Although there were a small number of respondents in the sample from pure internet retailers Ocado and Foodferry, the majority of the respondents were from grocery retailers with both bricks and clicks stores. Therefore, a comparative study between pure internet and bricks and clicks store would be interesting. The study is also limited to just one country; a useful extension would be to expand the research into a number of other countries with different competitive, consumer and business environments.

REFERENCES

- Anderson, R. E. and Srinivasan, S. S. (2003) E-satisfaction and E-loyalty: A Contingency Framework, *Psychology & Marketing*, 20 (2), pp. 123-138.
- Bagozzi, R. P. and Yi, Y. (1988) On the evaluation of structural equation model, *Journal of Academy of Marketing Science*, 16 (1), pp.74-94.
- Barnes, Stuart J. and Richard T. Vidgen (2002), "An Integrative Approach to the Assessment of E-Commerce Quality," *Journal of Electronic Commerce Research*, 3 (3), 114-27.
- Boshoff, C (2006) "A Psychometric Assessment of E-S-QUAL: A Scale To Measure Electronic Service Quality", *Journal of Electronic Commerce Research*. Vol. 8, No.. 1; p. 101-114
- Garver, Michael S. and John T. Mentzer (1999), "Logistics Research Methods: Employing Structural Equation Modeling to Test for Construct Validity," *Journal of Business Logistics*, Vol. 20, No. 1, pp. 33-57.
- Gefen, D. (2002) Customer Loyalty in E-Commerce, *Journal of the Association for Information Systems*, 3, pp. 27-51.
- Janda, S., Trocchia, P. J. and Gwinner, K. P. (2002) Consumer perception of Internet retail service quality, *International Journal of Service Industry Management*, 13 (5), pp. 412-431.
- Jones, M. A. and Suh, J. (2000) Transaction-specific satisfaction and overall satisfaction: an empirical analysis, *The Journal of Services Marketing*, 14 (2), pp.147-159.
- Loiacono, E. T., Watson, R. T. and Goodhue, D. L. (2002) WebQual: a measure of Web site quality, American Marketing Association, Chicago.
- Mintel Report (2007) Food Shopping-UK-November2007, Mintel International Group Limited, London.
- Oliver, R. L. (1980) A cognitive model of the antecedents and consequences of satisfaction decisions, *Journal of Marketing Research*, 17, pp. 460-469.
- Parasuraman, A., Zeithaml, V. A. and Malhotra, A. (2005) E-S-QUAL: A Multiple-Item Scale for Assessing Electronic Service Quality, *Journal of Service Research*, 7 (3), pp. 213-233.
- Santos, J. (2003) E-service quality: a model of virtual service quality dimensions, *Managing Service Quality*, 13 (3), pp. 233-246.
- Shankar, V., Smith, A. K. and Rangaswamy, A. (2003) Customer satisfaction and loyalty in online and offline environments, *International Journal of Research in Marketing*, 20 (2), pp.153-175.
- Spreng, R. A. and Mackoy, R. D. (1996) An empirical examination of a model of perceived service quality and satisfaction, *Journal of Retailing*, 72 (2), pp. 201-
- Wolfinbarger, M. and Gilly, M. C. (2003) e-TailQ: dimensionalizing, measuring and predicting etail quality, *Journal of Retailing*, 79 (3), pp. 183-198.
- Yang, Z. (2001) Consumer perceptions of service quality in Internet-based electronic commerce, *Proceedings of the 30th EMAC Conference*, Bergen, pp. 8-11.
- Yoo, B. and Donthu, N. (2001) Developing a scale to measure the perceived quality of an Internet shopping site (SITEQUAL), *Quarterly Journal of Electronic Commerce*, 2 (1), pp. 31-45.
- Zeithaml, V. A., Berry, L. L. and Parasuraman, A. (1996) The behavioural consequences of service quality, *Journal of Marketing*, 60 (2), pp. 31-47.
- Zeithaml, V. A., Parasuraman, A. and Arvind, M. (2002) Service Quality Delivery Through Web Sites: A Critical Review of Extant Knowledge, *Journal of Academy of Marketing Science*, 30 (4), pp. 362-375.
- Zeithaml, V. A., Parasuraman, A. and Malhotra, A. (2000) A conceptual Framework for Understanding e-Service Quality: Implications for Future Research and Managerial Practice, Marketing Science Institute, Cambridge, MA, Working Paper, Report No. 00-115.