CONJOINT ANALYSIS OF E-TAILER QUALITY: RELIABILITY, PREDICTIVE VALIDITY AND OTHER CONSIDERATIONS

Rose Sebastianelli, rose.sebastianelli@scranton.edu, (570) 941-4287
Nabil Tamimi, nabil.tamimi@scranton.edu, (570) 941-4288
Department of Operations and Information Management, Kania School of Management,
University of Scranton, 320 Madison Avenue, Scranton, Pennsylvania, 18510

ABSTRACT

Conjoint analysis is used to determine the relative importance of five factors (reputation, site usability, security, delivery, and customer support) on electronic retailer (e-tailer) quality. Results on the reliability and predictive validity of the estimated individual conjoint models are presented. Caveats in using conjoint for this purpose are considered.

Keywords: E-Tailing, Quality, Conjoint

INTRODUCTION

The growth in electronic commerce (e-commerce) has been extraordinary, with many of today's business transactions being conducted via the Internet. Although the projected annual growth in online retail sales is increasing at a decreasing rate, the latest quarterly e-commerce report published by the U.S. Census Bureau estimates total e-commerce sales for 2011 at $194.3 billion, an increase of 16.1% from 2010. Moreover, e-commerce sales account for an increasing percentage of total retail sales, 4.6% in 2011 compared with 4.3% in 2010 (http://www.census.gov). Consequently, determining what creates a quality shopping experience for online consumers is paramount for electronic retailers (e-tailers) attempting to meet and/or exceed customer expectations and stay competitive in an ever more crowded virtual marketplace.

Examining quality in the e-commerce arena (e-quality) has been an area of research interest for more than a decade. While early studies generally emphasized conceptualizing e-quality using multi-dimensional frameworks, a number of researchers have approached the issue empirically by developing multi-item scales to measure e-quality constructs. The common theme among this body of work has been the identification and evaluation of factors affecting online consumers' perceptions of e-quality and/or satisfaction with e-commerce. Our study, while within the context of this line of inquiry, focuses on determining the effects of selected electronic retailing (e-tailing) dimensions on consumers' judgments of e-tailer quality. To do this we use conjoint analysis, a method long used in marketing research studies (e.g., Green and Rao, 1971) but relatively new to e-commerce (e.g., Schaupp and Belanger, 2005). Conjoint analysis provides the means to decompose overall judgments into "part-worths" that show the relative contributions of each attribute (factor) to the response variable (i.e., perceived quality of the e-tailer website). A documented advantage of this approach is that it forces respondents to "trade-off" among factors in developing their overall evaluations. In this way, the relative importance of each factor can be discerned in a more realistic decision making scenario (when asked to rate the importance of factors directly, respondents have a tendency to rate them all as important).

In this paper we present the results of an experimental study in which participants provide overall judgments (rankings) of e-tailer quality described in terms of five attributes (reputation of retailer, site usability, security, delivery and customer support) shown to be salient in previous e-tailing quality research. These data serve as the basis for estimating individual level conjoint models from which the relative importance of the factors is derived. We evaluate the reliability and predictive validity of the estimated conjoint models as well as address some important caveats in the use of this methodology for e-commerce research.

BACKGROUND

Early studies examining quality in the e-commerce arena (e-quality) used dimensions related to service quality as a starting point. For example, Cox and Dale (2001) argued that while the lack of human interaction during an online experience makes some service quality factors irrelevant for virtual operations, some do apply. These include accessibility, communication, credibility, understanding and appearance. Van Iwaarden *et al.* (2003) found that all previously defined service quality dimensions are applicable to websites. Their survey results on the quality aspects perceived to be important in website design and website use correspond directly to the well known scale for measuring service quality, SERVQUAL (Berry and Parasuraman, 1991).

In a more comprehensive framework, Madu and Madu (2002) borrowed dimensions from both service and product quality to define e-quality. They identified the following 15 e-quality dimensions: performance, features, structure, aesthetics, reliability, storage capability, serviceability, security and system integrity, trust, responsiveness, product/service differentiation and customization, web store policies, reputation, assurance, and empathy. Some of these dimensions were redefined for e-commerce; for example, reliability, a dimension of both product quality and service quality, here refers to "consistency of web site performance." Others are unique to virtual operations. For instance, trust, which includes attributes that affect the willingness of users to disclose personal information over the Internet, or security/system integrity, which refers to the ability to safeguard and protect online transactions.

Researchers taking an empirical rather than conceptual approach to quality in e-commerce have developed scales for measuring latent e-quality constructs. SITEQUAL, developed by Yoo and Donthu (2001), identified the following four e-quality dimensions: (1) ease of use, (2) aesthetic design, (3) processing speed and (4) security of personal and financial information. These four dimensions were based on collecting data from convenience samples of students on nine items. Using information gathered from website designers and consumers, as well as undergraduate students' ratings of e-commerce sites, Loiacono, Watson and Goodhue (2002) developed WebQual. They identified the following 12 dimensions: 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. Both SITEQUAL and WebQual focused on the website interface.

Given that a consumer's experience with electronic retailers (e-tailers) goes beyond the website interface, others have developed scales that consider all aspects of purchasing via the Internet. For example, eTailQ (Wolfinbarger and Gilly, 2003) consists of 40 attributes that were reduced

to four underlying e-tailing dimensions: (1) fulfillment / reliability, (2) website design, (3) customer service and (4) security / privacy. Equally comprehensive are the E-S-QUAL and E-RecS-QUAL scales developed by Parasuraman, Zeithaml and Malhotra (2005). Based on data collected from a random sample of Internet users, E-S-QUAL yielded the following four e-tailing quality dimensions: efficiency (ease of using the site), fulfillment (extent to which the site's promises are fulfilled), system availability (correct technical functioning) and privacy (degree of protection) while the E-RecS-QUAL revealed three e-customer service dimensions: responsiveness (effective handling of problems), compensation (degree consumers are compensated for problems) and contact (availability of assistance).

STUDY DESIGN AND DATA COLLECTION

In an effort to balance the need to represent salient e-tailing quality dimensions cited in the literature while keeping the number of overall judgments required by study participants in the conjoint task manageable, we chose to use five factors to describe various e-tailers. The first factor we use is *reputation of the retailer*. Vendor characteristics (including store reputation) have not only been included in a number of studies on online shopping attitudes and behavior (see the comprehensive review of research by Li and Zhang, 2002), but we contend that this factor is related to a number of dimensions in SERVQUAL that apply to e-tailing quality (i.e., reliability, responsiveness) as well as to the frequently cited issue of trust (e.g., Lee and Lin, 2005). The second and third factors we include are *site usability* and *security*, both of which have been repeatedly included on various e-quality scales (i.e., SITEQUAL, WebQual, eTailQ, E-S-QUAL). In order to capture aspects of online shopping that go beyond the website interface, we also include *delivery* (to represent the fulfillment dimension in eTailQ and E-S-QUAL) and *customer support* (to represent e-customer service dimensions from E-RecS-QUAL) as the fourth and fifth factors. Each of these five factors is conceptualized at two levels as shown in Table 1.

Table 1: Factors (Attributes) and Levels for E-Tailer Descriptions

Attribute	Levels
Reputation	1. Lesser known small specialty retailer. 2. Widely recognized, established retailer.
Site Usability	1. Web interface is not user friendly. 2. User friendly web interface.
Security	1. No visible assurance of secure transactions. 2. Encryption technologies to assure secure transactions.
Delivery	1. Order tracking number provided. 2. Excellent record for on-time delivery.
Customer Support	1. Toll free number is available for customer service. 2. Live person chat support available for customer service.

A complete factorial design of the five attributes results in a total of 32 (2×2×2×2×2) different e-tailer descriptions. The full profile approach was used to collect the data. Because this method uses the complete set of attributes, it yields 32 e-tailer profiles for participants to judge. In order to alleviate respondent fatigue during the conjoint task, we used a ½ fractional factorial design to

implement the full profile approach. This reduced the number of evaluative judgments required by each participant from 32 to 16.

Research participants were recruited from a university setting to take part in an experimental study about "online shopping." Flyers were used to publicize the study and provide details regarding compensation and session times. The conjoint task was one of several tasks to be completed during the session. A total of 121 research participants took part in the study, which was carried out over multiple sessions. For the conjoint task, participants were presented with 16 multiple cue stimulus cards, each a verbal description of an e-tailer in terms of the five attributes. All groups of cards were randomized prior to presentation. Participants were instructed as follows: *You have 16 cards (designated A through P) and each describes a particular online retailer. Arrange the cards from best online retailer to worst online retailer based on the description provided. Place the rank (1 =best to 16= worst) on the line next to the retailer designation.* A subset of participants (n = 34) completed a second conjoint task. The set of cards used in the second conjoint task were created according to a different ½ fractional factorial design than those used in the original conjoint task. These participants, therefore, evaluated two sets of e-tailers, each set being described by the same attributes and their levels, but neither set containing duplications of the other. This procedure is referred to as the alternate forms method with spaced testing and is used to assess the reliability and predictive validity of the estimated conjoint models.

DATA ANALYSIS AND RESULTS

Participants in the study included undergraduate students, graduate students and professional staff. The age of respondents ranges from 19 to 66 years, with an average of 27 years. About 30% are older than 22 years of age, suggesting that the majority consists of typical undergraduate students. They are predominately female (64.5%) and most are employed at least part-time (26.5% full-time and 38% part-time). Additional information regarding online shopping behaviors, such as browsing and purchasing from e-tailers, is provided in Tables 1, 2 and 3. Over 60% have made between 1 and 6 online purchases in the last 3 months; almost 80% make at least 6 purchases from online retailers, on average, per year. The vast majority of respondents (almost 75%) report frequent browsing of e-tailers, on a daily or weekly basis.

Table 1: Online Purchasing Behavior of Study Participants

Number of online purchases made in the last 3 months.	Number	Percent (%)
None	9	7.4 %
1 to 3	39	32.2 %
4 to 6	37	30.6 %
7 to 9	14	11.6 %
10 or more	22	18.2 %

Table 2: Online Purchasing Behavior of Study Participants

Frequency of purchasing from online retailers (per year).	Number	Percent (%)
0 to 5	26	21.5 %
6 to 12	37	30.6 %
13 to 20	33	27.3 %
21 to 36	16	13.2 %
36 or more	9	7.4 %

Table 3: E-tailer Browsing Behavior of Study Participants

Frequency of browsing online retailers.	Number	Percent (%)
Daily	37	30.6 %
Weekly	53	43.8 %
Monthly	26	21.5 %
Only a few times per year.	5	4.1 %

The rankings of e-tailer descriptions were analyzed using monotone analysis of variance (MONANOVA), a method that is especially suited for factorial designs and appropriate for estimating part-worths for attribute levels from ordinal judgments. MONANOVA selects the best monotone transformation of the data, over all possible monotone transformations, such that the greatest percentage of variance can be accounted for by the main effects. Conjoint models were estimated at the individual level for all 121 participants. Summary results for the estimated part-worths of the attribute levels across all participants are presented in Table 4.

Table 4: Summary Results: Estimated Part-Worth Values for Attribute Levels

Model	Minimum	Maximum	Mean	Std. deviation
Intercept	8.063	9.125	8.502	0.084
Lesser known small specialty retailer.	-4.000	2.000	-1.492	1.134
Widely recognized, established retailer.	-2.000	4.000	1.492	1.134
Web interface is not user friendly.	-4.000	0.750	-1.626	0.925
User friendly web interface.	-0.750	4.000	1.626	0.925
No visible assurance of secure transactions.	-4.375	0.125	-3.194	1.111
Encryption technologies to assure secure transactions.	-0.125	4.375	3.194	1.111
Order tracking number provided.	-2.000	2.000	-0.029	0.675
Excellent record for on-time delivery.	-2.000	2.000	0.029	0.675
Toll free number is available for customer service.	-1.625	1.750	-0.042	0.547
Live person chat support available for customer service.	-1.750	1.625	0.042	0.547

The range in the estimated part-worth values between the levels of a given attribute is an indication of the attribute's importance; the greater the range the more important the attribute's contribution to the overall judgment of e-tailer quality. Table 5 presents a summary of the

relative importance of the five attributes. The relative importance is expressed as the mean percentage contribution (across all participants) to the overall judgment (the standard deviation in percentages indicates the level of consistency among participants). On average, the factor most important to perceptions of e-tailer quality is *security*, followed by *site usability* and *reputation of retailer*. *Delivery* and *customer support* are found to be substantially less important.

Table 5: Average Relative Importance of Attributes

Attribute	Relative Importance (Mean %)	Standard Deviation (%)
Security	43.546 %	15.174
Site Usability	22.412 %	12.868
Reputation	21.087 %	14.097
Delivery	7.326 %	5.514
Customer Support	5.628 %	5.281

In order to assess the reliability and predictive validity of the estimated individual conjoint models, we used the rankings for the alternate set of 16 e-tailer profiles provided by the subset of 34 participants who took part in a second conjoint task. First, these rankings were analyzed using MONANOVA to obtain a second set of estimated individual conjoint models. Pearson correlations between the part-worths of the first and second models were then computed for each of the 34 individuals. The individual level models demonstrate a high level of reliability; the correlations between model part-worths range from 0.871 to 1.000, with a mean of 0.972 and a median of 0.981. To assess predictive validity, we used the first set of estimated individual conjoint models to predict each participant's rankings for the alternate set of 16 e-tailer profiles. The predicted rankings for the second set of e-tailer profiles were then correlated (by computing Spearman rank correlations) to the actual rankings. Results indicate a high level of predictive validity; the Spearman rank correlations range from 0.626 to 0.997, with a mean of 0.875 and a median of 0.928. Table 6 shows the results at the individual level.

Table 6: Reliability and Predictive Validity Results

Individual	Reliability (Pearson Correlations)	Predictive Validity (Spearman Rank Correlations)
1	0.980	0.830
2	0.979	0.944
3	0.992	0.971
4	0.970	0.874
5	0.994	0.962
6	0.987	0.938
7	0.993	0.968
8	0.992	0.972
9	0.963	0.774
10	1.000	0.997
11	0.926	0.806
12	0.871	0.671
13	0.996	0.982
14	0.980	0.807
15	0.923	0.626
16	0.971	0.875
17	0.980	0.918
18	0.999	0.997
19	0.995	0.962
20	0.991	0.960
21	0.990	0.711
22	0.987	0.782
23	0.996	0.952
24	0.982	0.953
25	0.941	0.632
26	0.996	0.986
27	0.956	0.798
28	0.980	0.947
29	0.937	0.724
30	0.996	0.958
31	0.999	0.997
32	0.935	0.829
33	0.952	0.876
34	0.921	0.771

DISCUSSION AND CONSIDERATIONS

The results from our conjoint study indicate that *security* is the most important factor, on average, affecting perceptions of e-tailing quality, followed by *site usability* and *reputation of the retailer*. The major advantage of using conjoint methodology, as already noted, is the ability to force individuals to "trade-off" among factors when judging e-tailer quality thereby providing a more realistic way to determine the relative importance of various factors. Furthermore, the

individual level conjoint models we estimated exhibit high levels of reliability and predictive validity. It appears that conjoint methodology has potential for e-commerce research that focuses on understanding how various attributes affect users' preferences, opinions, and perceptions about e-quality.

Although conjoint analysis has been a popular marketing research tool for decades, its use in e-commerce research has been limited. One study that is similar to ours is that of Schaupp and Belanger (2005) who used conjoint methodology to evaluate online consumer satisfaction. They considered factors related to technology (security, usability and site design, privacy), shopping factors (convenience, trustworthiness, delivery), and product factors (merchandising, product value, product customization). Each factor was conceptualized at three levels; for example, delivery was defined as (1) provide a tracking number, (2) minimization of delivery time and (3) customer made aware of delays. They used the full profile approach with written paragraph descriptions, but had participants rate their level of satisfaction with each e-tailer resulting in metric (rather than ranked) responses. Consequently, they used ordinary least squares to estimate the conjoint models. Based on data collected from a sample of undergraduate students, they found that the most important factor was *privacy* (46.9%) followed by *merchandising* (15.1%); none of the remaining factors had mean percentages above 10%. Interestingly, in their study *security* ranked last in terms of relative importance. Even though the studies are not exact replications, the results are so different as to warrant further examination.

As in most research, a number of decisions are required in a conjoint study. These include selecting the type of model, data collection method, the factors and levels, measurement scale for the response, and estimation method. Since conjoint analysis allows us to determine the relative importance of attributes by decomposing an overall judgment, it is important to realize the potential for affecting an attribute's relative importance by how we conceptualize its levels. Consider how *security* is conceptualized in our study versus that of Schaupp and Belanger (2005). We use two levels that are quite extreme: *no visible assurance of secure transactions* versus *encryption technologies to assure secure transactions*. Schaupp and Belanger (2005), on the other hand, used three levels, none of which indicate a complete lack of security (*confirmation screen*, *encryption* and *password/ID protected accounts*). Can this be a plausible explanation for why security is the most important attribute in our study and the least important attribute in theirs? At the very least it raises an important consideration, or perhaps a caveat, in the use of conjoint analysis in e-commerce research.

One possible way to reduce the risk of this potential problem (of an attribute's relative importance being affected by how its levels are defined) is to use conjoint methodology with real rather than fictitious e-tailing descriptions. At a minimum we need to be aware that the risk of this undesirable effect exists. As with any research methodology, conjoint analysis is not without its limitations. Nonetheless, its advantage of providing a more realistic task makes it a promising methodology for better understanding the factors affecting consumers' perceptions of online shopping and other e-commerce transactions and online activities.

REFERENCES

Berry, L.L. and Parasuraman, A. (1991). *Marketing Services – competing through quality*. New York, NY: The Free Press.

Cox J. and Dale, B.G. (2001). "Service quality and e-commerce: an exploratory analysis." *Managing Service Quality*, 11(2), 121-31.

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

Lee, G.G. and Lin, H.F. (2005). "Customer perceptions of e-service quality in online shopping." *International Journal of Retail & Distribution Management*, 33 (2/3), 161-176.

Li, N. and Zhang, P. (2002). "Consumer online shopping attitudes and behavior: an assessment of research." *Proceedings of Eighth American Conference on Information Systems*, 508-517.

Loiacono, E., Watson, R. and Goodhue, D. (2002). "WEBQUAL: a measure of web site quality." *Marketing Educators' Conference: Marketing Theory and Applications*, 13, 432-437.

Madu, C.N. and Madu, A.A. (2002). "Dimensions of e-quality." *International Journal of Quality & Reliability Management*, 19(3), 246-58.

Parasuraman, A., Zeithaml, V.A. and Malhotra, A. (2005). "E-S-QUAL: a multi-item scale for assessing electronic service quality." *Journal of Service Research*, 7(3), 213-33.

Schaupp, L. and Belanger, F. (2005). "A conjoint analysis of online consumer satisfaction." *Journal of Electronic Commerce Research*, 6(2), 95-111.

Van Iwaarden, J., van der Wiele, T., Ball, L. and Millen, R. (2003). "Perceptions about the quality of websites: a survey amongst students at Northeastern University and Erasmus University." *Information & Management*, 41, 947-959.

Wolfinbarger, M. and Gilly, M.C. (2003). "eTailQ: dimensionalizing, measuring and predicting e-tail quality." *Journal of Retailing*, 79(3), 183-98.

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), 31-46.