

# Developing a Scale to Measure the Perceived Quality of An Internet Shopping Site (SITEQUAL)

Boonghee Yoo and Naveen Donthu

Yoo, Boonghee and Naveen Donthu (2001), "Developing a Scale to Measure the Perceived Quality of an Internet Shopping Site (SITEQUAL)," *Quarterly Journal of Electronic Commerce*, 2 (1), 31-47

An earlier version of this paper was presented at the 2000 Academy of Marketing Science Conference, Montreal, Canada, in May 2000. The authors thank Foundation Committee of St. Cloud State University, and Marketing Roundtable and Electronic Commerce Institute of Georgia State University for partial funding of this project.

## ABSTRACT

As Internet shopping gradually moves from a novelty to a routine way of shopping, the quality of the Internet sites will play an important role in differentiating sites. Internet shopping sites must be of high quality to attract consumers and influence their shopping decisions. The purpose of this study is to develop and validate a psychometrically rigorous instrument to measure the perceived quality of an Internet shopping site (i.e., SITEQUAL). Candidate items were generated based mainly on consumers' own descriptions. Multiple methods and samples produced a 9-item scale of SITEQUAL, which consists of four dimensions. This scale can be used to evaluate the quality of Internet shopping sites and examine how site quality affects visitors' online behavior, such as search patterns, site patronization, and buying decisions.

*Keywords:* Internet shopping site, online shopping, online shopping mall, e-tailing, site quality, perceived quality, quality criteria, scale development, SITEQUAL, and web design.

## INTRODUCTION

The value of Internet-based transactions is growing quickly and is estimated to grow 20-fold over the next five years, reaching \$11 trillion by 2010 (Mullaney 2000). In 1999, United States' businesses spent about \$4.3 billion on Internet ads and are likely to spend around \$28 billion in 2005 (Kranhold 2000). This trend shows that the Internet is fast becoming a new way to market to consumers. Consumers are rapidly adopting Internet shopping (Donthu and Garcia 1999), and even baby boomers' and seniors' online shopping grew by 18.4% in 1999, making them the Internet's fastest-growing demographic groups (Montgomery 2000). U.S. consumers most often use the Internet to purchase items such as books, computers and peripherals, travel, and clothes (Information Access 1999). Despite financial turbulence and increasing competitive pressure, Internet shopping sites or malls may gradually become substitutes for conventional

retailing channels, mail- or phone-order stores, catalogs, and sales forces. However, as the importance of, and competition among, Internet shopping sites increases, the quality of the sites will become important for survival and success. As Internet shopping gradually moves from a novelty to a routine way of shopping, the quality of Internet sites will play an important role in differentiating sites. High quality sites will be able to attract more browsers and shoppers than competing low quality sites because quality builds sustainable competitive assets.

In this study, we are interested in developing a way to measure the perceived quality of an Internet shopping site (referred to as SITEQUAL hereafter), that is, the quality of an Internet shopping site as perceived by consumers. There has been little research to develop a psychometrically sound measure of Internet shopping sites. In developing SITEQUAL, we emphasize three aspects of Internet shopping sites' perceived quality. First, consumers are the ultimate judges of the quality

of a site. It is their perception that we are interested in capturing, so we do not force a definition of quality on the consumers. Second, Internet site quality is described in the common consumer's vernacular rather than in formal academic language or the technical vocabulary of web designers (e.g., information content or stickiness). Third, Internet sites without shopping features (such as pure promotional or informative sites) may develop a different definition of quality and are not included in the SITEQUAL domain.

We define Internet shopping sites as web retail sites in which customers can browse, evaluate, order, and purchase a product or service. In short, Internet shopping sites are online versions of physical retail stores where all transactions and their relevant activities take place in online cyber spaces. Sites can be classified into six retail models (see Clark 1997): manufacturer sites in which manufacturers directly sell their products to customers online, skipping wholesalers and retailers (e.g., Sony.com, Compaq.com, and Hoover.com); off-line, brick and mortar retailer sites in which physical stores make their retail products available on the web (e.g., BestBuy.com, Walmart.com, and OfficeDepot.com); catalog hybrid sites that put their printed catalogs on the web (e.g., Fingerhut.com, ColumbiaHouse.com, and Christianbook.com); pure dot-com retail sites that buy products from manufacturers at wholesale and sell them online at retail without owning physical and public stores (e.g., Amazon.com, Buy.com, and eToys.com); mall sites that develop a location on the web and make money by charging retailers fees (e.g., Shopping.Yahoo.com, BizRate.com, and MySimon.com); and brokerage sites that unite buyers and sellers on the web and charge a portion of the transaction for the service (e.g., eBay.com, Ameritrade.com, and PriceLine.com). SITEQUAL is designed to capture the perceived quality of any type of web retail site.

Literature on the quality of traditional retail stores confirms that consumers use store quality as a vital extrinsic cue about the quality of the stores' products (e.g., Dodds, Monroe,

and Grewal 1991). Similarly, high quality Internet sites will attract more attention, contacts, and visits from consumers by implying that their products are of high quality. In addition, when consumers are satisfied with a high quality Internet site, they stimulate active and positive word-of-mouth communications for the site. Consumers also hear more favorable and positive comments about high quality sites through media and site evaluating agencies. Hanson (2000, p. 44) asserts:

A well-designed site has a number of beneficial impacts. It can build trust and confidence in the company; reinforce an image of competence, functionality, and usefulness; alert the visitor to the company's range of products and services; and point out local dealers, upcoming special events, and reasons to come back again.

However, little research has attempted to develop a psychometrically sound instrument to measure the perceived quality of an Internet shopping site. The lack of a measure has been a barrier to tracking and improving site quality and investigating the relationships between site quality and related variables.

Previous measures of site quality have several limitations. First, many of them are designed primarily to measure site efficiency rather than site quality. They are not consumer-based measures of quality. Examples of such measures include server logs (Trochim 1996), consumed time per page or visit (Dreze and Zufryden 1997), sales transactions (Schubert and Selz 1997), and web traffic (Lohse and Spiller 1998). Such site-centric measures may be inaccurate because a significant amount of off-server activity through local or proxy caches is never recorded (Strauss and Frost 1999). More seriously, they fail to capture consumers' cognitive and attitudinal evaluations of the sites, which are more meaningful indicators of Internet shopping behavior. Second, the measures do not show the structure of the dimensions of site quality. For example, Ho (1997) listed web sites' multiple quality features without investigating the relationships among

them. Third, the measures are mainly developed as performance metrics for general web sites rather than Internet shopping sites (e.g., Hoffman and Novak 1996). As a result, they may not be necessarily relevant to online shopping behavior because they neglect shopping-related attributes such as ordering, pricing, financial security, and payment methods. Fourth, the measures' psychometric properties have not been reported fully. For example, whether or not the retained measures make unidimensional constructs has not been checked (e.g., Alpar 1999).

The purpose of this study is to develop and validate a psychometrically rigorous instrument to measure the perceived quality of an Internet shopping site. The scale named SITEQUAL will be reliable (i.e., showing a high proportion of variance in true scores to the variance in observed scores, Nunnally and Bernstein 1994), valid (e.g., measuring what we want to measure), multidimensional (i.e., revealing distinct dimensions of the construct), and parsimonious (i.e., having a manageable number of items). Potentially, this scale will benefit Internet shopping site designers and researchers and can be used to evaluate and track the quality of Internet shopping sites. Managers may also use the measured indexes to set performance goals for their shopping sites. Also, competing sites can be evaluated using the scale. Previously, site efficiency measures, such as log analysis, customer-site interactivity, customer conversion, and retention, could be collected only for a focal site, not for competitors' sites. In addition, using the scale, researchers can examine the consequences and antecedents of site quality.

## **SCALE DEVELOPMENT**

### **Sample**

In the scale development phase, we surveyed a college student sample although we recognize the generalization of student sample results is usually believed to be limited. However, three facts may excuse the use of a student sample for this study. First, college students are among the most active online buyers. Their Internet and actual online

purchase experiences, technological advances, and innovativeness qualify them as a proper sample for online shopping research. For example, one study showed that in 1998, 45% of college students went online everyday and the average college student spent a total of \$305 for that year (Educational Marketer 1999). During the 2000 Christmas holiday season only, college students planned to spend an average of \$150 online, for a total of \$2.2 billion (PR Newswire 2000). Second, students are accepted for theory testing research in which the multivariate relationships among constructs, not the univariate differences (i.e., mean score comparisons) between samples, are being investigated (Calder, Philips, and Tybout 1981). Third, consumer behavior studies comparing students and nonstudents reveal students do not behave differently from nonstudents. Instead, students are shown to be effective surrogates for nonstudents or adults in various empirical studies that have examined, for example, country of origin, attitude toward advertising, and attitude-preference relationship (Yavas 1994), and price-quality perceptions (Lichtenstein and Burton 1989).

### **Item Generation**

We generated items based mainly on consumers' own descriptions because SITEQUAL focuses on consumers' perceptions of quality. Students in two marketing classes at a major Midwest university participated in the study. They listed the characteristics of Internet shopping sites that succeeded or failed to make them purchase in response to the following open-ended questions:

If you have ever purchased a product or service through the Internet, what characteristics of the web sites made you purchase? Provide as many characteristics as you can recall. Please be specific.

If you visited Internet shopping sites but did not make a purchase, what characteristics of the web sites made you NOT purchase? Provide as many

characteristics as you can recall. Be specific please.

Of 57 participants, 36 (63%) had purchased online. The participants described a total of 92 site characteristics. With redundant descriptions excluded, 54 descriptions were retained as the pool of candidate items of SITEQUAL. The descriptions were reworded in short, clear, Likert-type statements for evaluation.

The candidate items seemed consistent with Internet quality factors suggested in the literature. A literature review indicates the following characteristics are important to determine the quality of Internet sites: entertainment, informativeness, and organization (Chen and Wells 1999); sensory appeal, selection, substance, speed, simplicity, savings, service, and security (Pollock 2000); appearance, content, linkage, use, structure, and special feature (Abels, White, and Hahn 1997); informativeness, loading speed, accessibility, ease of navigation, and efficient transaction processing (Limayem, Khalifa, and Frini 2000); hedonic and utilitarian values (Karson 2000); social escapism, socialization, information, interactive control, economic values, security and privacy (Korgaonkar and Wolin 1999).

### **Instrumentation and Procedures**

A two-page questionnaire was developed to evaluate the 54 site characteristic descriptions, using five-point Likert scales anchored by 1 = “strongly disagree” and 5 = “strongly agree.” The descriptions were randomly ordered to avoid the systematic order effect or the cluster answering effect. Appendix 1 reports the 54 originally worded and ordered items.

The questionnaire was distributed to 94 students in three marketing classes. Participants were asked to visit and interact with three Internet shopping sites of their own selection and then immediately rate the 54 descriptions for each site. Participants were also asked to record the Internet address of each site they visited and indicate to which of 17 given product categories the site best belongs. They were given two days to visit and evaluate three sites. A total of 69 participants (73% response rate) completed the

questionnaire, which resulted in 207 evaluations as each participant evaluated three sites. Table 1 shows the classification of the responses by product category. Sites for books, music, videos, apparel and accessories, department stores, travel, and automobiles received the most visits, but no single category was dominant. This wide variety of site categories enhances the cross-site applicability of the scale.

\*\*\*\*\*  
Insert Table 1 here  
\*\*\*\*\*

### **Exploratory Factor Analysis**

To shorten the description set while maintaining the diversity of site quality perceptions, we conducted exploratory factor analysis. As a result of a careful examination of eigenvalues, proportion of variance explained, and scree test, nine distinct factors were identified. These factors exceeded one eigenvalue each, showed rapid increments in the scree curve, explained 75 percent of the variance, and, more importantly, showed a clear factor-loading pattern. Both orthogonal and oblique rotation methods produced an equivalent factor-loading matrix, confirming the stability of the factor-loading pattern.

The next step was to eliminate insignificant items. First, the items with statistically insignificant factor loadings at the 0.05 level were deleted (i.e., 0.40 or lower factor loadings for our sample size, Hair et al. 1998). Out of 54 items, 50 items survived. Second, the items with communality less than 0.50 were dropped because they did not meet acceptable levels of explanation and were poorly represented in the factor solution. Thirty-eight items were retained.

Based on the shared meaning among the items of each factor, the nine factors were labeled as aesthetic competitive value (four items), ease of use (four), design (seven), clarity of ordering (five), corporate and brand equity (five), processing speed (four), security (three), product uniqueness (two), and product quality assurance (three). The variance explained by each factor ranged from 5.3% to 10.1%. The communality of each item ranged from 51.8% to

87.2%, and the total communality reached 64.3% (24.43 items). The nine factors appeared to divide into two groups: one for vendor-related evaluations such as the products or terms of the sale, and the other for quality evaluations of the site. Although both groups of factors seemed to play an important role in online shopping, we decided to remove the vendor-related factors and keep only the factors that purely measure the quality aspects of the site itself because this study was to develop and validate the perceived quality of online shopping sites. The 20-item five factors we removed were competitive value, clarity of ordering, corporate and brand equity, product uniqueness, and product quality assurance; the retained 18-item four factors were ease of use, design, processing speed, and security. The removed factors will be useful for further research such as a broader investigation of determinants of electronic retailing success.

The nine factors can be described as follows:

**Vendor-Related Factors**

- § *Competitive value*: the competitive pricing in comparison to conventional retail stores or competing Internet shopping sites,
- § *Clarity of ordering*: the clarity of the ordering process supported by unambiguous pricing and fast delivery,
- § *Corporate and brand equity*: the name value of the site owner and the products or services on the site,
- § *Product uniqueness*: the uniqueness of the products or services on the site, such that visitors have difficulty finding the products elsewhere, and
- § *Product quality assurance*: the consumer's self-assurance of product quality obtained during the interaction with the site and not necessarily associated with direct product purchase or consumption experience.

**Quality-related Factors**

- § *Ease of use*: the ease of use and ability to search for information,
- § *Aesthetic design*: the creativity of a site with excellent multimedia and color graphics,
- § *Processing speed*: the promptness of online processing and interactive responsiveness to a consumer's requests, and
- § *Security*: the security of personal and financial information.

**Confirmatory Factor Analysis**

Even though a factor structure is found to be stable in exploratory factor analysis, in which one item is loaded on every factor, it may not be so in confirmatory factor analysis (CFA), in which one item is loaded on one factor only. Thus, to test the factor structure more rigorously, we conducted CFA and estimated the model through the LISREL 8 maximum likelihood method (Jöreskog and Sörbom 1993). This method detects the unidimensionality of each factor, which indicates the presence of a single trait or construct underlying a set of measures (Anderson and Gerbing 1988). The unidimensionality of each SITEQUAL factor would indicate sufficient convergent and discriminant validity by recognizing the factor as salient and independent of any other factor. We set a measurement model to have the retained four quality-related factors (latent variables) with the 18 selected items (manifest variables) and specified that an item be loaded on one factor only.

The measurement model turned out to be a poor model. The fit indexes were far below the acceptable levels (see Hu and Bentler 1999). The chi-square statistic was 501.63 with 129 degrees of freedom. Goodness of fit index (GFI) and adjusted goodness of fit index (AGFI) were 0.80 and 0.74, 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 0.80, 0.80, and 0.75, respectively. The

standardized root mean square residual (SRMR) was 0.085. The root mean square error of approximation (RMSEA) was 0.11.

### **Model Respecification and the Final SITEQUAL Measurement Model**

The model needed to be respecified to detect misfitting parameters and achieve a clearer factor structure with unidimensional factors only. For this purpose, CFA was used as exploratory, which was adequate for the scale development phase of this study. Variables were deleted from the variable list one by one if they showed an intolerably high modification index or residual. Large modification indexes signify the presence of factor crossloadings and error covariances, while large (positive or negative) residuals indicate the misspecification of the related items in the model (Byrne 1998). Large modification indexes or residuals, which may result from model misspecification, nonnormally distributed data, or nonlinear relationships among particular variables, affect the overall model fit (Jöreskog and Sörbom 1993).

The first respecified model was prepared by eliminating an item, Q34, which had the highest residual of 5.28. Without the item, the chi-square statistic improved from 501.63 to 399.06. The model continued to be modified until the highest residual became smaller than 3 and the highest modification index became reasonably low. In addition, although some cross-loading items may be important for measuring the overall quality of shopping sites, we removed significant cross-loading items for the following reasons. First, the purpose of this study was to develop a reliable, valid, and parsimonious scale of multiple independent dimensions, which demands a clean factor structure. Removing cross-loadings helps interpret the factors. When one item loads on one factor only, the factor structure will be simple and clear. If items significantly load on multiple factors, the factors will not be easily interpretable. Second, unless a theory suggests cross-loadings, cross-loadings might be attributed to a statistical artifact. Third, as is typical with confirmatory factor analysis (e.g., Anderson and Gerbing 1988), fit of the model is

significantly improved by removing of items with high cross-loadings.

In the process developing the final scale, the following items were dropped: “It is easy to order,” “This site is very easy to use,” and “It is easy to find products” in the ease of use dimension; “This site offers the best price” and “This site provides good deals” for the competitive value dimension; “This site is attractive” and “This site looks professional” for the aesthetic design dimension; “The web site is fast” and “This site does not ask many questions” for the processing speed dimension; and “This site has a security system on my information” the security dimension.

After a series of respecification procedures, the SITEQUAL measurement model of nine items (see Table 2) was supported by values of fit. CFI, IFI, and NNFI were 0.97, 0.98, and 0.96, respectively. SRMR was 0.044, and RMSEA was 0.054. The chi-square statistic was 32.07 with 21 degrees of freedom. GFI and AGFI were 0.97 and 0.93. These fit indexes indicate the model’s acceptable level of fit. All items were loaded on their corresponding constructs. The item loadings on their corresponding dimensions ranged from 0.51 to 0.96 as reported in Table 2. The smallest t-value of the loadings was 5.75, which exhibits the great significance of the loadings.

\*\*\*\*\*  
Insert Table 2 here  
\*\*\*\*\*

As shown in Table 3, the composite reliability estimates, which are equivalent to internal reliability coefficient alpha and are computed from LISREL results as evidence of convergent validity (Fornell and Larcker 1981), ranged from 0.72 (ease of use) to 0.78 (speed and security). These reliability coefficients were considered satisfactory in this scale development stage (Nunnally and Bernstein 1994). In addition, the average variance extracted for each dimension ranged from 0.59 (ease of use) to 0.64 (speed and security) and was greater than the squared correlation between the dimension and any other dimension, which indicates the

independence of the dimensions as evidence of discriminant validity (Fornell and Larker 1981).

\*\*\*\*\*  
Insert Table 3 here  
\*\*\*\*\*

### **SITEQUAL Index and its Applications**

The nine items of SITEQUAL can be combined and then averaged to form a composite measure, because they are almost evenly distributed among the four dimensions and no single factor is dominant. When evaluating the 207 shopping mall sites surveyed, the SITEQUAL indexes ranged from 1.25 to 5 out of 5. The mean and standard deviation of the indexes were 3.78 and 0.61, respectively. The SITEQUAL indexes of the 207 evaluations were not normally distributed with skewness of -0.68 and kurtosis of 0.89. One plausible explanation for this own is that participants visited relatively popular (and maybe high quality) sites.

The index can also be computed in terms of each dimension of SITEQUAL. As reported in Table 3, the index for individual dimensions was relatively high: ease of use (4.10), aesthetic design (3.79), processing speed (3.62), and security (3.63). The composite index of SITEQUAL was highly correlated with the individual dimensions of SITEQUAL: 0.67 with ease of use, 0.61 with aesthetic design, 0.71 with speed, and 0.69 with security. These high correlations indicate that the importance and contribution of each dimension to the overall index of quality is almost identical.

### **SCALE VALIDATION**

#### **Data Collection Procedure and Sample Profile**

A validation study of the SITEQUAL was conducted next, in which the SITEQUAL was tested using a new sample that was independent of the sample used in the scale development phase. This validation study evaluated the reliability and dimensionality of the scale and evaluated the predictive and nomological validity of the scale. We administered the questionnaire to 47 individuals

(24 males and 23 females). Each participant evaluated four Internet shopping sites of his or her own interest. On average, participants spent 27 minutes developing an understanding of each site's characteristics. A total of 187 site evaluations were received after one was eliminated due to ineligible responses. A variety of products were evaluated across 17 product categories. Major products included apparel (22% of 187 evaluations), electronics (21%), books and magazines (11%), music (9%), toys and games (9%) and travel and leisure (5%). In the last six months, the average participant purchased products 1.95 times and spent \$52 at the sites they evaluated for this study.

Data showed 100% of the participants were using the web on a daily basis, and 64% of the participants made online purchases in the last six months. Excluding non-online buyers, the participants purchased an average of 3.3 times online in the last six months. The average participant started to use the web (excluding e-mail) 4.5 years ago and started buying online 2.3 years ago; she or he used the web for 39 minutes a day, and spent \$177 on online purchases in the last six months.

#### **Reliability of the SITEQUAL**

The SITEQUAL showed acceptable psychometric properties in this validation study. First, the scale was highly reliable in the new sample, indicating the credibility of the measure. The reliability was 0.69 for ease of use, 0.76 for aesthetic design, 0.73 for speed, and 0.83 for security. In addition, the explained variance of each dimension ranged from 0.51 (aesthetic design) to 0.72 (security). Second, the scale's four-dimensional measurement model in a confirmatory manner through LISREL 8 showed satisfactory fitness. The model's chi-square statistic with 21 degrees of freedom was 23.9. RMSEA and SRMR were 0.025 and 0.040, respectively. CFI, TLI, and NNFI were each 0.99. GFI and AGFI were 0.97 and 0.94, respectively. These fit indices exhibit an excellent level of fit with the model in the new data. The item loadings to their constructs ranged from 0.58 (t-value = 7.72) to 0.99 (t-value = 11.63).

## **Validity of the SITEQUAL**

### **Internet experiences and SITEQUAL.**

Each dimension and the overall index of the scale were meaningfully related to Internet related experiences. First, participants with longer Internet usage showed more favorable perceptions of the performance of Internet shopping sites. Specifically, the number of years of Internet usage was correlated positively with ease of use ( $r = 0.18, p < 0.01$ ), aesthetic design ( $r = 0.26, p < 0.001$ ), speed ( $r = 0.27, p < 0.001$ ), security ( $r = 0.34, p < 0.0001$ ), and the overall SITEQUAL index ( $r = 0.36, p < 0.0001$ ). Second, participants with longer online shopping experiences showed less favorable perceptions of the performance of the Internet shopping sites. The number of years of online shopping was correlated negatively with ease of use ( $r = -0.38, p < 0.0001$ ), aesthetic design ( $r = -0.38, p < 0.0001$ ), speed ( $r = -0.25, p < 0.001$ ), security ( $r = -0.39, p < 0.0001$ ), and the overall SITEQUAL index ( $r = -0.49, p < 0.0001$ ). Third, participants who spent more money on online shopping showed less favorable perceptions of the performance of the Internet shopping sites. The number of dollars spent on online shopping was correlated negatively with ease of use ( $r = -0.23, p < 0.01$ ), aesthetic design ( $r = -0.37, p < 0.0001$ ), speed ( $r = -0.18, p < 0.05$ ), security ( $r = -0.22, p < 0.01$ ), and the overall SITEQUAL index ( $r = -0.36, p < 0.0001$ ).

These results show that general Internet experience lets people develop positive interest in online shopping but as they become more experienced online buyers, people tend to more critically judge online shopping sites. The gap model of service quality evaluation, in which the perceived service quality equals quality expectations subtracted from actual quality performance (Parasuraman, Zeithaml, and Berry 1988), can explain this phenomenon. Because customers with more experiences expect more improved performance over time, which is particularly true in the technology market, they tend to perceive the same level of performance as inferior.

**Construct Validity of SITEQUAL.** To evaluate the construct validity, we first generated an overall site quality index by

summarizing the average score of each SITEQUAL dimension and relating it to the individual dimensions. The correlation of the composite index with individual dimensions was significant at the 0.0001 level: 0.79 with ease of use, 0.78 with aesthetic design, 0.68 with processing speed, and 0.69 with security. This demonstrates that the scale captures consumers' overall response to the site.

Second, we related the SITEQUAL to an independent four-item measure of overall site quality perception. The overall perceived quality of a site indicates consumers' subjective judgments about the site's overall excellence or superiority. It presents overall quality rather than individual components of quality. Four items, anchored with "strongly disagree = 1" and "strongly agree = 5," were based on the items of Dodds, Monroe, and Grewal (1991). The measure is reported in Appendix 2, and its reliability coefficient was 0.90. The correlation of overall site quality with each SITEQUAL dimension was significant at the 0.0001 level: 0.68 with ease of use, 0.60 with aesthetic design, 0.48 with processing speed, and 0.50 with security. The correlation between overall site quality and the SITEQUAL index was also high ( $r = 0.77, p < 0.0001$ ). All these high correlations demonstrate the construct validity of the SITEQUAL in measuring consumers' quality perceptions of an Internet shopping site.

Third, overall site quality was regressed on the SITEQUAL dimensions. This regression analysis revealed how accurately overall site quality can be explained and predicted by each SITEQUAL dimension. To control for the effect of a site's product category on the quality evaluation of the site, we included the type of product in the multiple regression model. Because the sites were evaluated for 14 product categories, the regression model had 13 dummy variables and the intercept term was used to estimate the effect of the fourteenth product category. Table 4 reports the results. Consistent with the simple correlation results described earlier, overall site quality was a function of the SITEQUAL dimensions, which explained 63% of the variance of overall site quality. Each dimension was positively related to overall site quality, but ease of use ( $b = 0.35, p < 0.0001$ )

and aesthetic design ( $b = 0.33, p < 0.0001$ ) showed stronger relationships to overall quality than processing speed ( $b = 0.12, p < 0.5$ ) and security ( $b = 0.16, p < 0.001$ ). In the regression model, despite significant correlations among the SITEQUAL dimensions, serious multicollinearity was not detected. The variance inflation factor ranged from 1.07 to 3.24 and the condition index was 20.94. As a variance inflation factor of ten or less or a condition index of 30 or less is considered the threshold for lack of multicollinearity (see Hair et al. 1998), we concluded that the regression coefficients of our regression model were not significantly affected by other predictor variables, which indicates that the coefficients are stable.

#### **Nomological Validity of SITEQUAL.**

We evaluated the nomological validity of the scale by examining the relationships with seemingly highly related variables. We selected five relevant constructs and regressed them on the SITEQUAL. Investigated variables included attitude toward the site, site loyalty, site equity, purchase intention, and site revisit intention. We generated a two-item measure of site revisit intention for this study by borrowing and modifying other measures from previous studies. Specifically, we adopted the measure of attitude toward the site from Chen and Wells' (1999) study and the measure of online purchase intention from Limayem, Khalifa, and Frini's (2000) study. We modified Beatty and Kahle's (1988) brand loyalty measure and made a site loyalty measure to capture the overall commitment to a specific site. Finally, we modified Yoo, Donthu, and Lee's (2000) brand equity measure to fit the online shopping context, resulting in a site equity measure as an analogy of brand equity, which was expected to capture the incremental value of the site due to the site name over competing sites with identical physical characteristics. The measures are shown in Appendix 2 and all of them achieved satisfactory reliabilities, ranging from 0.77 (site loyalty) to 0.96 (purchase intention).

First, to check the general nomological validity of the SITEQUAL, we examined the correlation between its overall composite index and each nomological construct. The correlation of the composite index with the constructs was

significant at the 0.0001 level: 0.69 with attitude toward the site, 0.44 with site loyalty, 0.54 with site equity, 0.35 with purchase intention through the site, and 0.55 with site revisit intention. In addition, the correlation between individual SITEQUAL dimensions and the relevant constructs was consistent and significant, which demonstrates strong nomological validity of the scale.

Second, we conducted a series of regression analyses in which each relevant variable was regressed on the SITEQUAL dimensions and the dummy-coded product categories. Table 4 summarizes the results, which clearly shows the SITEQUAL dimensions are significant predictors of the examined variables. The coefficient of determination (that is,  $R^2$ ) was significant at the 0.0001 level, ranging from 0.31 (purchase intention) to 0.58 (attitude toward the site). Specifically, attitude toward the site was a function of ease of use ( $b = 0.32, p < 0.0001$ ), aesthetic design ( $b = 0.22, p < 0.001$ ), and security ( $b = 0.16, p < 0.0001$ ). Site revisit intention showed a pattern similar to attitude toward the site: ease of use ( $b = 0.23, p < 0.01$ ), aesthetic design ( $b = 0.28, p < 0.01$ ), and security ( $b = 0.24, p < 0.001$ ). Interestingly, purchase intention was a function of two dimensions only: ease of use ( $b = 0.32, p < 0.01$ ) and security ( $b = 0.30, p < 0.001$ ). Site loyalty was a function of ease of use ( $b = 0.27, p < 0.05$ ), processing speed ( $b = 0.16, p < 0.05$ ), and security ( $b = 0.27, p < 0.001$ ). Site equity showed results similar to site loyalty: ease of use ( $b = 0.26, p < 0.01$ ), processing speed ( $b = 0.22, p < 0.01$ ), and security ( $b = 0.33, p < 0.001$ ). Serious multicollinearity was not found in any regression model. The variance inflation factor and the condition index remained the same as in the previous regression model conducted to explain overall site quality because the identical predictor variables were used.

Across regression analyses, two dimensions (ease of use and security) were consistently more significant than the other two dimensions (aesthetic design and processing speed). This implies that ease of use and security are the most important quality criteria of an Internet shopping site that influence consumer attitudes and behaviors. Processing

speed was also an important factor, influencing site loyalty and equity, whereas aesthetic design was an additionally important factor for attitude toward the site and the site revisit intention. It was noteworthy that product categories, with the exception of electronics, generally did not affect the examined variables, which implies that the effects of the SITEQUAL dimensions may be universal across product categories.

**Applications of the SITEQUAL in Practice.** In this section, we illustrate how using the SITEQUAL can profile Internet shopping sites. We selected six shopping sites: Gap, Best Buy, eBay, Amazon, Buy, and Walmart, each of which was voluntarily selected and evaluated by six or more participants. For each site, the SITEQUAL scores were computed at the dimensional and overall levels. The overall level score was the mean of the means of the four individual SITEQUAL dimensions. Gap's shopping site was rated the best overall (mean = 4.19 out of 5), exceeding every other site in three dimensions: ease of use (4.33), design (4.44), and security (4.06), while Walmart's site was considered the worst overall (3.53), with the lowest aesthetic design (2.96) and speed (3.75). Best Buy's site, rated second, performed with consistent scores across dimensions. People showed great mistrust in Buy.com's security (3.38). Amazon.com earned medium scores across dimensions. The eBay site was perceived as the best in ease of use (4.33) but not great in security (3.66). Generally, the brick and mortar retailer sites (Gap, Best Buy, and Walmart) were rated higher in security than the pure dot-com sites (eBay, Amazon, and Buy). The sites with narrow breadth of merchandise (Gap and Best Buy) were perceived to have better design than those with broader breadth of merchandise (Buy, Walmart, eBay, and Amazon). The six sites did not show significant differences in processing speed, ranging from 3.74 (Amazon) to 3.92 (Best Buy).

## DISCUSSION

The results of the validation study show that our proposed SITEQUAL scale has appropriate reliability and validity in every aspect. The scale has four dimensions and only nine items, which can be considered

parsimonious. The scale, consistent with major factors discussed in the literature, seems to have face validity and is easy to administer. In the validation study the Internet shopping sites that had higher quality were rated higher in a variety of consumer attitudes and behaviors, such as attitude toward the site, site loyalty, site equity, purchase intention, and site revisit intention. This clearly shows that site quality, as measured by the SITEQUAL, can be directly linked to site performance.

The SITEQUAL can be used to further examine how site quality affects visitors' online behaviors, such as search patterns, site patronization, and buying decisions. Similarly, the antecedents of site quality may also be investigated using the measure. In particular, the site design factors or marketing efforts that affect each dimension of SITEQUAL need to be identified for better site management. The measure helps site managers track the SITEQUAL of individual sites on a regular basis. By assessing the individual dimensions of the SITEQUAL, managers can identify the assets and weaknesses of their own and competing Internet shopping sites.

The SITEQUAL should not be regarded as a final measure but as a starting point toward a better measure. We suggest four directions for further research. First, further testing for nomological validity needs to occur by examining the relationships with objective measures of site effectiveness, such as traffic, amount of time spent, and site-visitor interactivity. Second, surveying diverse samples may develop a more generalizable scale. In particular, a cross-cultural study will answer important questions such as if the scale is equivalent across nations, if an Internet shopping site is equivalently perceived across nations, and if each SITEQUAL dimension equivalently affects relevant consequential behavior variables across nations. Third, whether the SITEQUAL holds for both business-to-business and consumer-good shoppers needs exploration. Our participants evaluated consumer goods sites only. Fourth, how the SITEQUAL dimensions are related to the technical aspects of web design should be examined.

The purpose of this research was to develop a psychometrically sound measure of the perceived quality of Internet shopping sites. Our measure of SITEQUAL comprises nine items representing four dimensions. This study will benefit research on Internet marketing and aid in the development of high quality Internet shopping sites.

## REFERENCES

- Alpar, Paul (1999), "Satisfaction with a Web Site: Its Measurement, Factors, and Correlates," Working Paper.
- Anderson, James C. and David W. Gerbing (1988), "Structural Modeling in Practice: A Review and Recommended Two-Step Approach," *Psychological Bulletin*, 103 (3), 411-23.
- Beatty, Sharon E. and Lynn R. Kahle (1988), "Alternative Hierarchies of the Attitude-Behavior Relationship: The Impact of Brand Commitment and Habit," *Journal of Academic Marketing Science*, 16 (Summer), 1-10.
- Bentler, Peter M. (1990), "Comparative Fit Indexes in Structural Models," *Psychological Bulletin*, 107, 238-246.
- Byrne, Barbara M. (1998), *Structural Equation Modeling with LISREL, PRELIS, and SMPLIS*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Chen, Qimei and William D. Wells (1999), "Attitude toward the Site," *Journal of Advertising Research*, 39 (September/October), 27-37.
- Clark, Bruce H. (1997), "Welcome to My Parlor: The Lure of Marketing on the World Wide Web Is Great. Be Sure You Don't Get Stuck with the Wrong Approach," *Marketing Management*, 5 (Winter), 11-22.
- Dodds, William B., Kent B. Monroe, and Dhruv Grewal (1991) "Effects of Price, Brand, and Store Information on Buyers' Product Evaluation," *Journal of Marketing Research*, 28 (August), 307-319.
- Donthu, Naveen and Adriana Garcia (1999), "The Internet Shopper," *Journal of Advertising Research*, 39 (May/June), 52-58.
- Dreze, Xavier and Fred Zufryden (1997), "Testing Web Site Design and Promotional Content," *Journal of Advertising Research*, 37 (March-April), 77-91.
- Hair, Joseph F., Jr., Rolph E. Anderson, Ronald L. Tatham, and William C. Black (1998), *Multivariate Data Analysis, 5<sup>th</sup> ed.* Upper Saddle River, New Jersey: Prentice-Hall.
- Hanson, Ward (1999), *Principles of Internet Marketing*. Cincinnati, Ohio: South-Western College Publishing.
- Ho, James K. (1997), "Evaluating the World Wide Web: A Global Study of Commercial Sites," *Journal of Computer-Mediated Communication (on-line)*, 3 (1).
- Hoffman, Donna L. and Thomas P. Novak (1996), "Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations," *Journal of Marketing*, 60 (July), 50-68.
- Hu, Li-tze and Peter M. Bentler (1999), "Cutoff Criteria for Fit Indexes in Conventional Criteria Versus New Alternatives," *Structural Equation Modeling*, 6, 1-55.
- Information Access (1999), "More Ads for Online Shoppers - Especially Outside U.S.," *Marketing News*, September 27, 3.
- Jöreskog, Karl G. and Dag Sörbom (1993), *LISREL 8.02*. Chicago: Scientific Software International, Inc.
- Karson, Eric J (2000), "Two Dimensions of Computer and Internet Use: A Reliable and Valid Scale," *Quarterly Journal of Electronic Commerce*, 1 (1), 49-60.
- Korgaonkar and Wolin (1999), "A Multivariate Analysis of Web Usage," *Journal of Advertising Research*, 39 (March/April), 53-68.
- Kranhold, Kathryn (2000), "Web Sites Toot Their Horns Amid Ad Chill," *Wall Street Journal*, June 27, B1, Column 5.
- Lichtenstein, Donald R., and Burton, Scot (1989), "The Relationship Between Perceived and Objective Price-Quality," *Journal of Marketing Research*, 26 (November), 429-443.
- Limayem, Moez, Mohamed Khalifa, and Anissa Frini (2000), "What Makes Consumers Buy

- from Internet? A Longitudinal Study of Online Shopping," *IEEE Transactions on Systems, Man, and Cybernetics--Part A: Systems and Humans*, 30 (July), 421-432.
- Lohse, G. L. and P. Spiller (1998), "Quantifying the Effect of User Interface Design Features on Cyberstore Traffic and Sales." *CHI 1998 Conference Proceedings*, ACM Press.
- Montgomery, Sara Teasdale (2000), "Senior Surfers Grab Web Attention; More Boomers and Seniors Online Than 18-24 Group," *Advertising Age*, July 10, S4.
- Mullaney, Timothy J. (2000), "Clicks and Misses," *Business Week*, July 24, EB21.
- Nunnally, Jum C. and Ira H. Bernstein (1994), *Psychometric Theory*. New York, NY: McGraw-Hill.
- Parasuraman, A., Valarie A. Zeithaml, and Leonard L. Berry (1988), "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality," *Journal of Retailing*, 64 (1), 12-40.
- Peterson, Robert A., Sridhar Balasubramanian, and Bart J. Bronnenberg (1997), "Exploring the Implications of the Internet for Consumer Marketing," *Journal of the Academy of Marketing Science*, 25 (Fall), 329-346.
- Schubert, Petra and Dorian Selz (1997), "Web Assessment -- A Model for the Evaluation and the Assessment of Successful Electronic Commerce Application," *Electronic Markets (on-line)*, 7 (3).
- Strauss, Judy and Raymond Frost (1999), *Marketing on the Internet: Principles of Online Marketing*. Upper Saddle River, New Jersey: Prentice Hall.
- Trochim, William M. K. (1996), "Evaluating Websites." Downloaded from <http://trochim.human.cornell.edu/webeval/webintro/webintro.htm>.
- Yavas, Ugur (1994), "Research Note: Students as Subjects in Advertising and Marketing Research," *International Marketing Review*, 11 (4), 35-43.
- Yoo, Boonghee, Naveen Donthu, and Sungho Lee (2000), "An Examination of Selected Marketing Mix Elements and Brand Equity," *Journal of the Academy of Marketing Science*, 28 (April), 195-211.
- "Online Survey by edu.com Reveals College Students Plan to Spend \$2.2 Billion Online This Holiday Season," *PR Newswire*, November 21, 2000, p. 2987.

**Table 1**  
**Evaluated Internet Shopping Sites by Category: the Scale Development Phase**

---

<i>Site Category</i>	<i>Responses</i>
1. Apparel and Accessories	35
2. Art and Collectibles	1
3. Auctions and Outlets	14
4. Books, Music and Videos	49
5. Computers	9
6. Consumer Electronics	3
7. Department Stores	21
8. Flowers and Gifts	8
9. Food and Drink	3
10. Health and Beauty	8
11. Home and Garden	3
12. Home Office and Business	3
13. Jewelry and Accessories	2
14. Kids, Toys, Games and Babies	7
15. Sports and Fitness	16
16. Travel and Auto	19
17. Comprehensive Site	7

---

N = 207

**Table 2**  
**SITEQUAL (Scale of Perceived Quality of Internet Shopping Site) Items:**  
**A LISREL Completely Standardized Solution**

<i>Dimensions and Items</i>	<i>Factor Loadings</i>	<i>t-value</i>
<b>Ease of Use</b>		
Q5 This site is convenient to use.	0.96	7.72
Q7 It is easy to search for information.	0.51	5.75
<b>Aesthetic Design</b>		
Q20 This site is colorful.	0.77	9.15
Q19 This site is creative.	0.60	7.48
Q30 This site shows good pictures of the products.	0.51	6.47
<b>Processing Speed</b>		
Q27 It is easy to access the results.	0.83	11.02
Q26 This site has quick process.	0.77	10.41
<b>Security</b>		
Q3 This site ensures me of security.	0.65	7.37
Q25 I am confident of security with this site.	0.93	8.90

**Table 3**  
**Intercorrelations Among the SITEQUAL Dimensions:**  
**A LISREL Completely Standardized Solution**

<i>Dimensions</i>	<i>Ease of Use</i>	<i>Aesthetic Design</i>	<i>Speed</i>	<i>Security</i>
Ease of Use	1.			
Aesthetic Design	0.37****	1.		
Processing Speed	0.45****	0.30***	1.	
Security	0.21**	0.24**	0.44****	1.
Reliability Coefficient	0.72	0.74	0.78	0.78
Variance Explained for	0.59	0.50	0.64	0.64
Mean	4.10	3.79	3.62	3.63
Standard Deviation	0.86	0.83	0.95	0.98

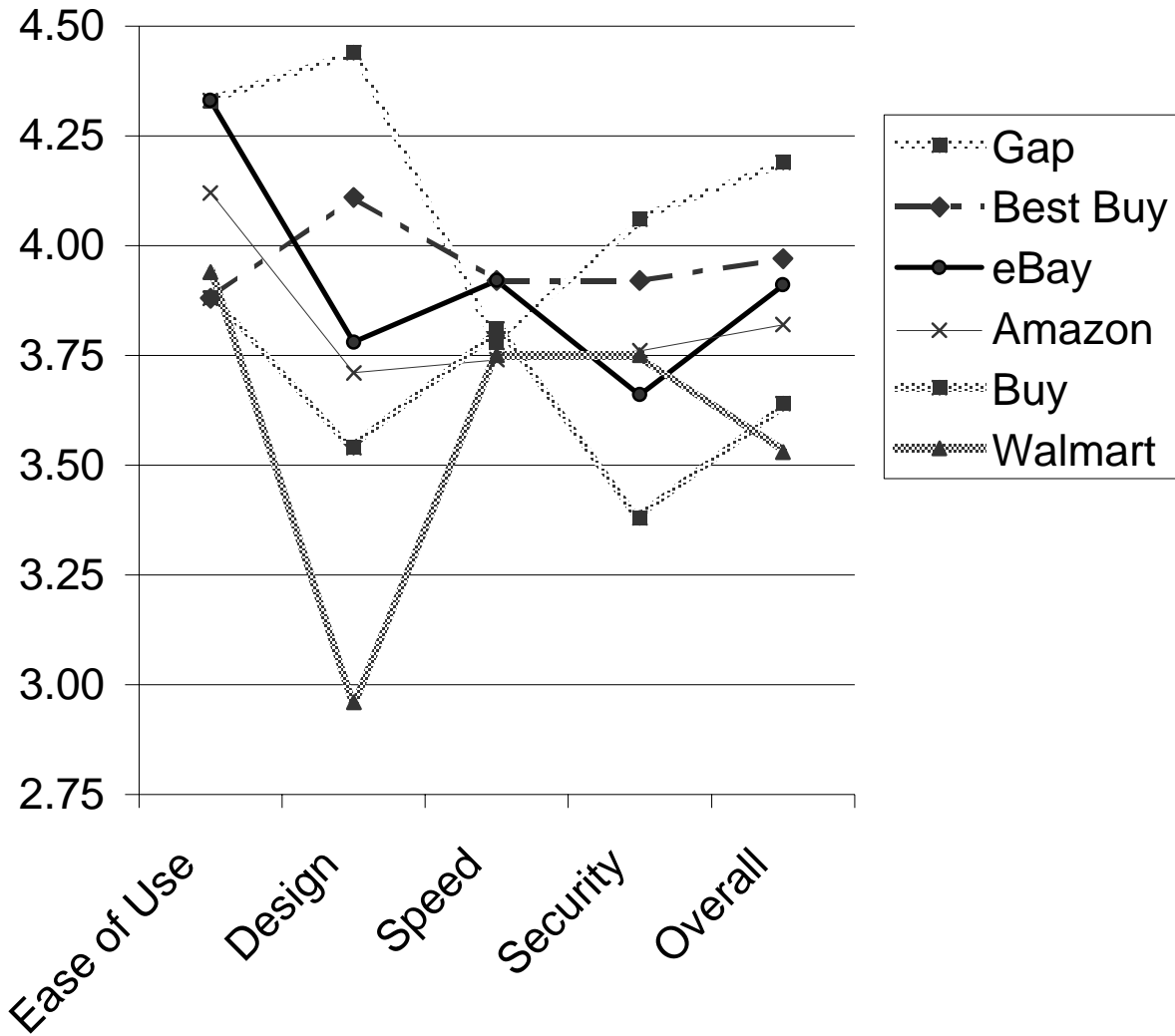
\* p < 0.05. \*\* p < 0.01. \*\*\* p < 0.001. \*\*\*\* p < 0.0001

**Table 4**  
**The Effects of SITEQUAL Dimensions on Relevant Constructs:**  
**Results of Multiple Regression Analysis**

<i>Variable</i>	<i>Overall Site Quality</i>	<i>Site Loyalty</i>	<i>Site Equity</i>	<i>Attitude Toward Site</i>	<i>Site Revisit Intention</i>	<i>Purchase Intention</i>
Intercept	0.39	0.82*	0.62	0.69*	1.29**	1.67**
<i>Product Category</i>						
Apparel	-0.10	-0.24	-0.23	-0.41*	-0.29	-0.25
Auctions	-0.37	-0.26	-0.25	-0.68*	-0.92*	-0.36
Books and magazines	-0.01	-0.42	-0.01	-0.34*	-0.25	-0.81*
Computer hardware	-0.11	-0.48	-0.06	-0.46	-0.35	-1.29*
DVDs and videos	-0.34	0.44	0.18	-0.05	-0.14	0.27
Electronics	-0.28*	-0.84**	-0.91***	-0.61***	-0.85***	-1.20***
Food and drink	0.00	-0.32	-0.02	-0.36	-1.15**	-1.61**
Gifts and flowers	-0.37	-0.61	-0.08	-0.45	-0.56*	-1.00*
Health and beauty	-0.35	-0.77	-0.86*	-0.64*	-0.91*	-0.73
Home and garden	-0.26	-1.42	-0.04	-0.48	0.01	-0.54
Music	-0.17	0.02	-0.18	-0.06	0.12	-0.32
Sports and hobbies	-0.13	-0.19	-0.18	-0.30	-0.23	-0.43
Toys and games	-0.34	-0.21	0.32	-0.50*	-0.54	-0.57
<i>SITEQUAL Dimensions</i>						
Ease of use	0.35****	0.27*	0.26**	0.31****	0.23**	0.32**
Aesthetic design	0.33****	-0.07	-0.01	0.22****	0.28**	-0.13
Processing speed	0.12*	0.16*	0.22**	0.06	-0.00	-0.01
Security	0.16****	0.27***	0.33****	0.24****	0.24***	0.30***
R <sup>2</sup>	0.63****	0.33****	0.46****	0.58****	0.44****	0.31****

\* p < 0.05. \*\* p < 0.01. \*\*\* p < 0.001. \*\*\*\* p < 0.0001. n = 187.

**Table 5**  
**Profiles of Selected Internet Shopping Sites with the SITEQUAL Scores**



## **Appendix 1**

### **The Originally Worded and Ordered Items**

---

- |   |  |
|---|--|
| Q1. This site is attractive.  | Q30. This site shows good pictures of the products.                  |
| Q2. The products' benefits are described well.                            | Q31. This site promises fast delivery.                               |
| Q3. This site insures me of security.                                     | Q32. The shipping and handling charge is reasonable.                 |
| Q4. The delivering time schedule is clear.                                | Q33. This site is impressive.  |
| Q5. This site is convenient to use.                                       | Q34. This site has a security system to protect my information.      |
| Q6. This site provides the best products.                                 | Q35. This site provides a variety of products.                       |
| Q7. It is easy to search for information.                                 | Q36. This site has a legitimate reputation.                          |
| Q8. This site offers the best price.                                      | Q37. This site has familiar brand names.                             |
| Q9. This site sells an unique product that I could not pick up in stores. | Q38. The products are cheaper in this site than in stores.           |
| Q10. It is all right to use a credit card in this site.                   | Q39. This site has payment methods other than credit cards.          |
| Q11. It is easy to order.   | Q40. The web site is fast.   |
| Q12. This site is very easy to use.                                       | Q41. It is easy to manage this site.                                 |
| Q13. The prices in this site are cheaper.                                 | Q42. I can evaluate the quality of the merchandise.                  |
| Q14. I get detailed information on products.                              | Q43. This site requires a reasonable amount of personal information. |
| Q15. The price is cheaper than in stores.                                 | Q44. The prices are not confusing.                                   |
| Q16. The products on this site are tough to find elsewhere.               | Q45. I can interact with the products.                               |
| Q17. Logging onto this site is possible any time.                         | Q46. This site is a pleasure.  |
| Q18. It is possible to get quick shipments.                               | Q47. A purchase decision can be made in a few links.                 |
| Q19. This site is creative.   | Q48. The site looks professional.                                    |
| Q20. This site is colorful.   | Q49. The company owning the site is legitimate.                      |
| Q21. It is easy to find products.   | Q50. I can feel the products.  |
| Q22. This site provides good deals.                                       | Q51. I can be sure of the quality of products.                       |
| Q23. This site has rich information of the products.                      | Q52. This site is organized well.                                    |
| Q24. This site has easy step-by-step instructions.                        | Q53. This site does not ask many questions.                          |
| Q25. I am confident in security with this site.                           | Q54. This site has a return policy.                                  |
| Q26. This site has quick process.   |  |
| Q27. It is easy to access results.  |  |
| Q28. In this site, I can find what I need.                                |  |
| Q29. The credit card security is great.                                   |  |
-

## Appendix 2

### The Measures Used to Evaluate Construct and Nomological Validity

---

#### **Overall Site Quality** (*Reliability coefficient = 0.90*)

1. This site is of high quality.
2. The likely quality of this site is extremely high.
3. This site must be of very good quality.
4. This site appears to be of very poor quality. (r)

#### **Attitude Toward the Site** (*Reliability coefficient = 0.83*)

1. This site makes it easy for me to build a relationship with the company.
2. I would like to visit this site again in the future.
3. I am satisfied with the service provided by this site.
4. I feel comfortable in surfing this site.
5. I feel surfing this site is a good way to spend my time.
6. Compared with other shopping sites, I would rate this one as one of the best.

#### **Site Loyalty** (*Reliability coefficient = 0.77*)

1. I consider myself to be loyal to this site.
2. This site would be my first choice.
3. I will not shop on other sites as long as I can access this site.

#### **Site Equity** (*Reliability coefficient = 0.92*)

1. It makes sense to buy on this site instead of any other site, even if they are the same.
2. Even if another site has the same features as this site, I would prefer to buy on this site.
3. If there is another site as good as this site, I prefer to buy on this site.
4. If another site is not different from this site in any way, it seems smarter to purchase on this site.

#### **Purchase Intention** (*Reliability coefficient = 0.96*)

1. I will definitely buy products from this site in the near future.
2. I intend to purchase through this site in the near future.
3. It is likely that I will purchase through this site in the near future.
4. I expect to purchase through this site in the near future.

#### **Site Revisit Intention** (*Reliability coefficient = 0.78*)

1. I am likely to revisit this site in the near future.
2. I am encouraged to revisit this site in the near future.

---

(r) Reverse coded.