Developing a Scale to Measure the Interactivity of Web Sites*

Yuping Liu

Marketing Area

Department of Business Administration

College of Business and Public Administration

Old Dominion University

Norfolk, VA 23529

Phone: (757) 683-6551

Email: YXXLiu@odu.edu

*This paper will appear in *Journal of Advertising Research*.

DEVELOPING A SCALE TO MEASURE THE INTERACTIVITY OF WEB SITES

ABSTRACT

This paper describes the development and validation of a scale measuring the interactivity of Web sites. Three studies were conducted to verify the factor structure, content validity, discriminant validity, and reliability of the scale. Results from the studies showed that interactivity comprises three correlated but distinct dimensions: active control, two-way communication, and synchronicity. The multidimensional scale showed a high level of validity and reliability and yielded consistent ratings among both experienced and inexperienced Internet users.

DEVELOPING A SCALE TO MEASURE THE INTERACTIVITY OF WEB SITES

INTRODUCTION

From the beginning, the Internet has stood out as being highly interactive. The easy modifiability of contents and the ability for instant information transmission has given users control and two-way communication opportunities that have not been possible with traditional media. Utilizing this interactive nature of the Internet, companies can now communicate more efficiently with consumers on a one-to-one basis. They can also better gauge consumer interests in their offerings.

Being such a central characteristic of the Internet medium, it is surprising how little interactivity is understood. Although people generally assume interactivity to be a desirable attribute, research on interactivity effects has produced ambiguous results. Whereas some studies confirmed the positive impact of interactivity on consumer response such as attitude toward the ad (Cho and Leckenby, 1999; Yoo and Stout, 2001), other studies have found little or even negative effect of interactivity (Bezjian-Avery, Calder, and Iacobucci, 1998; Coyle and Thorson, 2001; Sundar, Hesser, Kalyanaraman, and Brown, 1998). A closer look at these studies suggests that the conflicting results may be partly due to the lack of uniform conceptualization and operationalization of interactivity. Holding different understanding about what interactivity is, researchers have manipulated or measured the construct in various ways (see Liu and Shrum 2002 for a review). It is natural, then, that there have been no conclusive results about how interactivity influences online communication.

This lack of consistent understanding of interactivity can impede the effective use of the Internet as a marketing communication channel. For example, it is uncertain whether interactivity really facilitates persuasion and whether companies should make the extra effort to make their marketing messages more interactive. To answer these questions and to better understand the interactive nature of the Internet, it is essential that interactivity can be accurately and consistently gauged. An accurate measure of interactivity not only can quantify the construct and thus make it possible to find the precise relationship between interactivity and the dependent variables, but also can help ensure the effective operationalization of interactivity in experimental studies. Accordingly, the current research sets out to design such an interactivity measure. A series of studies were conducted to develop a scale assessing the interactivity of Web sites and to verify the latent structure and the validity and reliability of the scale.

CONCEPTUAL DEFINITION

A variety of interactivity definitions can be found in the literature. These definitions have mainly focused on two distinct aspects of interactivity: reciprocal communication and control. Definitions focusing on the reciprocal nature of an exchange argue that an interactive communication should satisfy two conditions. First, the interaction should allow two-way flow of information, and the information being exchanged in a sequence should closely relate to each other (Alba et al., 1997; Rafaeli and Sudweeks, 1997). Second, such an exchange of information should happen quickly. When one communicating party sends out a piece of information, he or she should be able to receive a fast response (Alba et al., 1997; Wu, 1999).

In contrast, several other definitions of interactivity have suggested control as the core component of the construct (Bezjian-Avery, Calder, and Iacobucci, 1998; Jensen, 1998; Rogers, 1995; Steuer, 1992). In an interactive communication, participants in the communication should be able to exert control on the information exchanged (Jensen, 1999; Rogers, 1995). This includes both information sent and information received. For example, an interactive Web site should allow visitors to control the information flow on the site and to customize the messages they receive according to their communication goals. In the case of a mediated communication, participants should also have control over the communication medium.

Both control and reciprocal communication are important aspects of online interactivity. Control helps ensure a reciprocal exchange that satisfies the needs of all communicating parties, while reciprocal communication provides an effective channel for exerting control. Melding the two aspects, Liu and Shrum (2002) defined interactivity as "the degree to which two or more communication parties can act on each other, on the communication medium, and on the messages and the degree to which such influences are synchronized." They proposed three dimensions of interactivity: active control, which describes a user's ability to voluntarily participate in and instrumentally influence a communication; two-way communication, which captures the bi-directional flow of information; and synchronicity, which corresponds to the speed of the interaction. The current interactivity scale was developed based on this conceptualization. Here an interactive communication is defined as *a communication that offers individuals active control and allows them to communicate both reciprocally and synchronously*.

EXISTING INTERACTIVITY MEASURES

Several measures of interactivity have been used in empirical studies, ranging from the one-item "relative interactivity" scale by Shankar, Smith, and Rangaswamy (2000) to the ten-item scale by Wu (1999). These existing measures are limited in two aspects. First, few of the measures have been constructed through a formal scale development process. As a result, the nature of the scales is hardly known. For example, it is unclear whether the scales are measuring the intended construct, whether they are unidimensional or multidimensional, or how the respondents' characteristics would influence the final ratings. Without knowing these, the validity and reliability of the scales cannot be established.

Second, many of the scales are contaminated with user-response variables. For example, Wu's (1999) perceived interactivity scale contains such affective response items as "I was delighted to be able to choose which link and when to click". The interactivity scale used in Cho and Leckenby (1999) is also confounded with behavioral intention, such as "I would bookmark this site for future usage" and "I would be willing to provide my personal information for the advertiser." Although such affective responses or behavioral intentions may be related to interactivity, they are also contingent upon individual users' characteristics, such as their Internet usage habits and their reactions to interactivity. For example, two users may perceive the same level of interactivity in an online ad. But one of them may be more concerned with privacy than the other and may not be willing to provide personal information to the advertiser. As a result, the same level of perceived interactivity will result in two different ratings using Cho and Leckenby's (1999) scale.

To avoid such confusions, the current research considers interactivity to be a perceptual level construct. It can lead to affective or behavioral responses, but it should be separated from those consequences.

INITIAL ITEM POOL GENERATION AND REVISION

Based on the above conceptualization and on an extensive review of related literature, an initial pool of twelve items was developed for each of the three dimensions of interactivity – active control, two-way communication, and synchronicity. Several of the items used by McMillan (2000) were included in this initial pool. Special care was taken to ensure that the scale does not contain any attitudinal or behavioral intention items.

Critiques of these items were then sought from colleagues familiar with the research topic. They were given the definition of interactivity and a description of the three dimensions and were asked to identify: (1) any incompatibility between an item and the dimension it is supposed to measure; (2) any set of items that do not fully capture the dimension they are supposed to measure; (3) and any ambiguity in the wording of the items. The items were revised based on these critiques. The resulting items were formatted into seven-point semantic differential scales anchored at "strongly disagree" and "strongly agree". Items for the three dimensions were randomized and interspersed in the questionnaire used in the following studies. Seventeen undergraduate business students were then recruited to rate several Web sites on interactivity using the scale. They were asked to describe any difficulties they had with completing the scale. The scale was then revised to incorporate the feedback received from these participants.

STUDY 1 AND STUDY 2: ITEM PURIFICATION

Overview

Two studies were conducted to refine the items in the scale. The questionnaire was administered to two samples of 42 and 87 undergraduate business students. The samples represented Internet users with a wide variety of Internet usage patterns and experiences, from using the Internet for one hour a week with 2 years Internet experience to using the Internet for more than 40 hours a week with 10 years of Internet experience. Furthermore, among the 87 participants in the second study, thirty-seven were nontraditional students. These students represent an older population than traditional students, and most of them work full-time during the day and take courses at night. The age of all participants ranged from 19 to 40 years old. Detailed participant demographic and Internet usage information is provided in Table 1.

Insert Table 1 about here

Procedure

In the first study, participants were asked to browse and rate consecutively three existing Web sites on interactivity. The Web sites used were home.com from the former broadband Internet service provider @Home, the online store of the retailer J. C. Penney, and quakeroatmeal.com from Quaker Oats. The order in which participants went through the Web sites was randomized. In the second study, each participant rated one of two Web sites developed for a fictitious portable audio company. Portable audio product category was chosen through pretests to be a relevant product for the sample population.

In both studies, participants were also asked to provide general demographic and computer/Internet usage information at the end of the questionnaire.

Item Purification

To evaluate the items, corrected item-to-total correlations and pairwise correlations between the items were calculated for both samples. An item was deleted if any of the following was true for either sample: (1) its item-to-total correlation was below .50 (Bearden, Netemeyer, and Teel, 1989); (2) its interitem correlation (the correlation between two items within a dimension) did not exceed .30 (Robinson, Shaver, and Wrightsman, 1991); and (3) the item correlated more strongly with items in other dimensions than with items in the same dimension (Bearden, Netemeyer, and Teel, 1989). The items within each of the three dimensions were also factor analyzed. Items whose factor loading did not reach .50 in either of the two samples were dropped from the scale (Bearden, Hardesty, and Rose, 2001). Fifteen items were retained from the above analysis with four items for active control, six items for two-way communication, and five items for synchronicity. These remaining items are listed in Table 2.

Insert Table 2 about here

STUDY 3: LATENT STRUCTURE ANALYSIS

Procedure

An additional study was carried out to verify the underlying structure of the items retained from the previous analysis. Eighty undergraduate business students participated in the study, among which 47 were nontraditional students (see Table 1 for detailed sample demographics). Participants were assigned to one of the two Web sites for a

fictitious portable audio company. The two Web sites were the same as the ones used in the second study. Participants were asked to browse the Web site and then rate the site on interactivity using the refined scale. Several open-ended questions were also added at the end of the questionnaire asking them what they liked or disliked about the Web site.

Confirmatory Factor Analysis

A series of confirmatory factor analysis was conducted on the resulting data. The proposed factor structure of interactivity is a second-order model in which the three factors form a second-order construct. This is mathematically equivalent to a three-factor correlated model, which was fitted using LISREL 8.5. To compare the model's performance with alternative structures, several other models were also estimated. One alternative model was a three-factor uncorrelated model assuming the three dimensions to be distinct and independent constructs. Three two-factor correlated models were also developed by combining every possible pair of the three dimensions. These two-factor models assume that two of the three dimensions actually belong to the same underlying factor, making interactivity a two-dimensional construct. Also estimated were a one-factor model treating interactivity as a unidimensional construct and a null model that assumes no systematic structure in the data.

Table 3 displays the goodness-of-fit indices of all six models. The hypothesized three-factor correlated model performed the best among the six models. The chi-square for the hypothesized model was 123.99 ($df = 87$), and the relative chi-square (dividing the chi-square by the degree of freedom) was 1.43, less than two as recommended by Carmines and McIver (1981) and the lowest among the six models. The hypothesized model is also the only model that satisfies the criteria of a good model fit on all other

indices, with both the Comparative Fit Index and the Non-Normed Fit Index at least .90 (Jaccard and Wan, 1996) and RMSEA less than .08 (Browne and Cudek, 1993). Furthermore, as shown in Table 2, the standardized factor loadings of the items estimated in the analysis all exceeded .50. These results suggest that the hypothesized three-factor correlated model is the best representation of the data.

Insert Table 3 about here

SCALE VALIDITY AND RELIABILITY

Reliability

Cronbach's alpha was calculated for each of the three dimensions of interactivity (Cronbach, 1951). A reliable measure should have an alpha value of .70 or more (Nunnally, 1978). The average Cronbach's alphas in the current studies were as follows: .75 for active control, .86 for two-way communication, and .86 for synchronicity, all exceeding the .70 threshold. The average reliability indices calculated through structural equation modeling were also high with .81 for active control, .90 for two-way communication, and .89 for synchronicity (Werts, Linn, and Jöreskog, 1974).

Fornell and Larcker (1981) have also recommended calculating the average variance extracted for a construct as an indicator of reliability. It measures the percentage of total variance of the data accounted for by the construct. The average variances extracted for the three factors across the three studies were .51 for active control, .59 for two-way communication, and .60 for synchronicity (see Table 4). All of them exceeded the .50 threshold recommended by Fornell and Larcker (1981).

Insert Table 4 about here

Discriminant Validity

Three tests were conducted to examine the discriminant validity of the scale. The first test compares the squared pairwise correlation between factors and the average variance extracted for each factor (Fornell and Larcker, 1981). To establish discriminant validity, the average variance extracted for a factor should be higher than all the squared correlations involving the factor. As shown in Table 4, all three average variances extracted in the analysis are larger than the squared pairwise correlations for all three studies, suggesting discriminant validity of the factors.

The second test of discriminant validity compares the chi-square statistics among the different models in the latent structure analysis. If the scale possesses discriminant validity, each constrained model (models with more factors) should result in a significantly improved chi-square from the less constrained models (models with fewer factors) (Anderson and Gerbing, 1988). Table 3 shows that the improvement in chi-square in the series of models was significant at all levels. The smallest chi-square difference between the one-factor model and the two-factor models was 22.74 ($p < .001$); and the improvement in chi-square from each of the two-factor models to the hypothesized three-factor correlated model was 45.73 ($p < .001$), 128.76 ($p < .001$), and 24.00 ($p < .001$) respectively. This suggests that treating the individual dimensions as distinct factors is superior to lumping the dimensions together. In other words, the three dimensions possess enough discriminant validity to be treated as distinct factors.

The third test of discriminant validity involves examining the correlation between each pair of the factors (Anderson and Gerbing, 1988). If the two factors are indeed

distinct, the correlation between them should be less than one. This was tested by constructing a 95% confidence interval for each correlation. To reject the null hypothesis that the two factors are not distinct, the confidence interval should not include one. In study 3, the confidence intervals were (.09, .37) for active control--two-way communication correlation, (.41, .63) for active control--synchronicity correlation, and (.19, .42) for two-way communication--synchronicity correlation. None of the confidence intervals included one. Study 1 and Study 2 yielded similar results, again suggesting the discriminant validity of the scale.

Known Group Validity

The two Web sites used in Study 2 and Study 3 were designed to possess different levels of interactivity. To achieve this, feature pairs that fulfill the same function but are different on interactivity were used. For example, one Web site had a banner ad on its entry page (more control), whereas the other site featured a pop-up ad (less control). The more interactive Web site also had an online feedback form in its customer service section (more two-way communication), whereas the less interactive site only provided phone numbers and email addresses of the customer service department (less two-way communication). In addition, the two sites were hosted on two different servers, one of which was expected to be faster than the other. It was therefore expected that the ratings of the two Web sites should differ significantly on the three interactivity dimensions. This was supported by the current data. As shown in Table 5, the more interactive Web site received higher ratings on all three dimensions than did the less interactive Web site.

Insert Table 5 about here

Content Validity

The qualitative responses participants provided on what they liked or disliked about the Web site in Study 3 were examined to verify the content validity of the scale. First, the responses were content analyzed and coded as pertaining to active control, two-way communication, synchronicity, or other (see Table 6 for examples of the responses in each category). Two marketing doctoral students served as independent judges, and interjudge agreement index was 90%. The responses the two judges did not agree upon were coded based on a discussion between the two judges and on consultation with the author. For each participant, a score was then derived for each of the interactivity dimensions by subtracting the number of unfavorable responses for a dimension from the number of favorable responses for that dimension. Correlation coefficients were obtained between these scores and the ratings obtained from the interactivity scale. If the scale is measuring what it is supposed to measure, a high correlation between the two groups of scores should be expected. The results from the study showed a correlation of .40 for active control, .43 for two-way communication, and .54 for synchronicity. All of the correlations are significant ($p < .01$), suggesting the content validity of the scale.

Insert Table 6 about here

Relationship with Personal Variables

Ideally, the interactivity scale should produce similar ratings of a given Web site for experienced as well as inexperienced Internet users. To verify this, participants in the three studies were asked how many years they had been using the Internet, their Internet usage per week, how comfortable they were with using the Internet and using computer

in general, and their diversity of experience with the Internet. Each of these variables was then correlated with the participants' ratings of the Web sites. No significant correlation was found between any of these variables and the ratings of the Web sites, suggesting that the scale is a stable measure of interactivity across people with diverse online experiences. This is consistent with Yoo and Stout's (2001) finding that users' Internet skills did not influence their perceived interactivity of a Web site.

COMPARING TRADITIONAL WITH NONTRADITIONAL STUDENTS

Previous research suggests that traditional students behave quite differently from both nontraditional students and adult consumers (James 2001). In a scale development study, however, Chen (2002) found little difference between traditional student population and the general consumer population. To see whether the traditional and nontraditional student participants in the current research responded similarly to the interactivity scale and to the Websites, several additional analyses were conducted. Overall, the analyses showed similar responses to the scale but significantly different responses to the Websites from the two types of participants.

To compare the two groups' responses to the scale, a separate factor analysis of the scale items was run for each group. As shown in Table 2, the patterns of factor loadings are quite similar between the traditional and the nontraditional student groups. A separate correlation analysis for each group also yielded similar correlations between the interactivity dimensions: for traditional students, the average correlations were .30 between active control and two-way communication, .38 between two-way communication and synchronicity, and .66 between active control and synchronicity; for the nontraditional student group, the correlation coefficients were .38, .40, and .57

respectively. Furthermore, as shown in Table 7, the scale was reliable for both samples. These results suggest that the underlying structure of the scale is similar for both types of participants and that it is a valid measure for both populations.

Insert Table 7 about here

Although the two types of participants responded similarly to the interactivity scale itself, MANOVA analysis of participants' response to the websites revealed one interesting difference. While the two groups provided similar ratings of the websites on active control and two-way communication, nontraditional student participants rated the websites consistently lower on synchronicity (M = 4.24 and 4.97 for the less and more interactive websites respectively) than did traditional student participants (M = 5.05 and 5.18; F = 4.17, p = .05). Although the current studies do not explain why such differences exist, it may be due to the difference in time perception between the two groups. Recall that synchronicity refers to the speed of interaction. Such judgment of speed is determined both by the actual waiting occurred during browsing and by individuals' subjective perception of the waiting (Dellaert and Kahn 1999; Hornik 1984). As nontraditional students often have to deal simultaneously with education, family, and a full-time job, they tend to have a higher sense of time urgency and bear a higher unit cost of time. Thus, they are more likely to pay close attention to time and are more sensitive to delay than traditional students. The current finding concurs with the view that traditional students may react differently to marketing stimuli than nontraditional students and adult consumers (James 2001). The ability of the current scale to distinguish between the two groups further indicates the accuracy of the scale. Given that

many adult consumers lead a busy family and work life, they are also likely to be sensitive to delay. It is therefore important for companies to provide instantaneous service to consumers on the Internet.

DISCUSSION

The current research developed a measure of interactivity based on the multidimensional conceptualization of the construct in Liu and Shrum (2002). The interactivity scale exhibited a high level of validity and reliability. It was able to differentiate the levels of interactivity of different Web sites and corresponded well with users' qualitative responses to the Web sites. Results also confirmed that active control, two-way communication, and synchronicity are three independent yet correlated dimensions of interactivity. Although the current research only involved student participants, the sample represented a diverse group of Internet users with various Internet usage history and experiences. It also included both traditional as well as nontraditional students. The similar responses of the two groups to the scale indicate the general applicability of the scale to different populations. Given that nontraditional students closely resemble adult consumers (James 2001), the validity and reliability of the current scale are likely to extend to the general Internet user population as well.

The interactivity scale developed here can be applied in both marketing practice and scholarly research. Companies can use the scale to evaluate their Web sites. Because of the multi-dimensional nature of the scale, it can help companies identify individual problems with their Web sites. For example, does the Web site try to control consumers too much? Does the site give consumers enough opportunities for two-way communication? Is the site responding to consumer requests fast enough? Finding these

problems with the individual dimensions helps pinpoint the deficiencies in a company's Web site and reveal ways of better utilizing the interactive potential of the Internet. The scale can also be used in academic research to understand how interactivity affects consumers' response to an online communication. Given the importance of interactivity in the online medium, the current understanding of interactivity is highly inadequate. A more precise measure of the construct can help uncover more accurate relationship between interactivity and the dependent variables. It can also help verify the widely accepted but unconfirmed positive nature of interactivity.

As online marketing goes beyond Web sites to other online tools such as e-mails and online advertisements, it would be desirable to measure the interactivity of these other online marketing tools using a more universal interactivity scale. Future research can expand the current scale to all forms of online communication. Such a universal scale would allow convenient comparison of these online marketing communication tools on their interactivity levels and can vastly deepen our understanding of how to utilize the Internet for marketing and advertising purposes.

REFERENCES

Alba, J., J. Lynch, B. Weitz, C. Janiszewski, R. Lutz, A. Sawyer, and S. Wood, "Interactive Home Shopping: Consumer, Retailer, and Manufacturer Incentives to Participate in Electronic Marketplaces." *Journal of Marketing* 61, 3 (1997): 48-53.

Anderson, J. C. and D. W. Gerbing, "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach," *Psychological Bulletin* 103, 3 (1988): 411-423.

Bearden, W. O., R. G. Netemeyer, and J. E. Teel, "Measurement of Consumer Susceptibility to Interpersonal Influence." *Journal of Consumer Research* 15, 4 (1989): 473-481.

----------, D. M. Hardesty, and R. L. Rose, "Consumer Self-Confidence: Refinements in Conceptualization and Measurement." *Journal of Consumer Research* 28, 1 (2001): 121-134.

Bezjian-Avery, A., B. Calder, and D. Iacobucci, "New Media Interactive Advertising vs. Traditional Advertising." *Journal of Advertising Research* 38, 4 (1998): 23-32.

Browne, M. W. and R. A. Cudeck, "Alternative Ways of Assessing Model Fit." In *Testing Structural Equation Models,* K. A. Bollen and J. S. Long, eds. Newbury Park, CA: Sage, 1993.

Carmines, E. G. and J. P. McIver, “Unobserved Variables.” In *Social Measurement: Current Issues,* G. W. Bohrnstedt and E. F. Borgatta, eds. Beverly Hills, CA: Sage, 1981.

Chen, Qimei, “Attitude Toward the Site II: New Information.” *Journal of Advertising Research* 42, 2 (2002): 33-45.

Cho, C., and J. D. Leckenby, "Interactivity As a Measure of Advertising Effectiveness: Antecedents and Consequences of Interactivity in Web Advertising." In *Proceedings of the 1999 Conference of the American Academy of Advertising*, M. S. Roberts, eds. Gainesville, FL: University of Florida, 1999.

Coyle, J. R. and E. Thorson, “The Effects of Progressive Levels of Interactivity and Vividness in Web Marketing Sites.” *Journal of Advertising* 30, 3 (2001): 65-77.

Cronbach, L. J., “Coefficient Alpha and the Internal Structure of Tests.” *Psychometrika,* 31 (1951): 93-96.

Dellaert, B. G. C. and B. E. Kahn, “How Tolerable is Delay?: Consumers’ Evaluation of Internet Websites after Waiting.” *Journal of Interactive Marketing*, 13, 1 (1999): 41-54.

Fornell, C. and D. F. Larcker, “Evaluating Structural Equation Models with Unobserved Variables and Measurement Error.” *Journal of Marketing Research* 18, 1 (1981): 39-50.

Hornik, J., “Subjective vs. Objective Time Measures: A Note on the Perception of Time in Consumer Behavior.” *Journal of Consumer Research* 11, 1 (1984): 615-618.

Jaccard, J. and C. K. Wan, *LISREL Approaches to Interaction Effects in Multiple Regression*. Thousand Oaks, CA: Sage Publications, 1996.

James, William L., “Just Say No to Student Samples.” *Journal of Advertising Research* 41, 5 (2001): 63-71.

Jensen, J. F., "Interactivity: Tracing a New Concept in Media and Communication Studies." *Nordicom Review* 19, 1 (1998): 185-204.

Liu, Y. and L. J. Shrum, “What Is Interactivity and Is It Always Such a Good Thing? Implications of Definition, Person, and Situation for the Influence of Interactivity on Advertising Effectiveness.” *Journal of Advertising* 31, 4 (2002): 53-64.

McMillan, S. J., "Interactivity Is in the Eye of the Beholder: Function, Perception, Involvement, and Attitude toward the Web Site." In *Proceedings of the 2000 Conference of the American Academy of Advertising,* Mary Alice Shaver, ed. East Lansing, MI: Michigan State University, 2000.

Nunnally, J., *Psychometric Theory*, 2nd ed. New York: McGraw-Hill, 1978.

Rafaeli, S., and F. Sudweeks, "Networked Interactivity." *Journal of Computer Mediated Communication* 2, 4 (1997), available at http://www.ascusc.org/jcmc/vol2/issue4/rafaeli.sudweeks.html [last accessed on September 9, 2002].

Robinson, J. P., P. R. Shaver, and L. S. Wrightsman, "Criteria for Scale Selection and Evaluation." In *Measures of Personality and Social Psychological Attitudes*, J. P.

Robinson, P. R. Shaver, and L. S. Wrightsman, eds., San Diego, CA: Academic Press, 1991.

Rogers, E. M., *Diffusion of Innovations*, 4th ed., New York: The Free Press, 1995.

Shankar, V., A. K. Smith, and A. Rangaswamy, "The Relationship between Customer Satisfaction and Loyalty in Online and Offline Environments." Pennsylvania State University eBusiness Research Center Working Paper 02-2000, 2000.

Steuer, J., "Defining Virtual Reality: Dimensions Determining Telepresence." *Journal of Communication* 42, 4 (1992): 73-93.

Sundar, S. S., K. M. Hesser, S. Kalyanaraman, and J. Brown, "The Effect of Website Interactivity on Political Persuasion." Paper presented at the meeting of the International Association for Media and Communication Research, Glasgow, UK, 1998.

Werts, C. E., R. L. Linn, and K. G. Jöreskog, “Interclass Reliability Estimates: Testing Structural Assumptions,” *Journal of Educational and Psychological Measurement* 34, 1 (1974): 25-33.

Wu, G., "Perceived Interactivity and Attitude Toward Web Sites." In *Proceedings of the 1999 Conference of The American Academy of Advertising,* Marilyn S. Roberts, ed. Gainesville, FL: University of Florida, 1999.

Yoo, C. Y. and P. A. Stout, “Factors Affecting Users' Interactivity with the Web Site and the Consequences of Users' Interactivity.” In *Proceedings of the 2001 Conference of the American Academy of Advertising,* Charles R. Taylor, ed. Villanova, PA: American Academy of Advertising, 2001.

Table 1

Sample Demographics and Internet Usage Profile

	Study 1	Study 2	Study 3
Gender			
Female	52%	56%	46%
Male	48%	44%	54%
Age			
Under 25 years old	81%	64%	56%
26 to 35 years old	17%	29%	25%
36 to 45 years old	2%	7%	18%
Over 45 years old	0%	0%	1%
Years using the Internet			
Under 2 years	5%	11%	0%
2 to 4 years	40%	34%	25%
5 to 6 years	48%	39%	38%
Over 6 years	7%	15%	38%
Time spent on the Internet per week			
Under 5 hours	33%	34%	38%
6 to 20 hours	57%	47%	48%
21 to 40 hours	7%	16%	15%
Over 40 hours	2%	2%	0%

Table 2

Interactivity Scale Items and Standardized Factor Loadings

Items	Standardized Factor Loadings		
	All	Traditional Students	Nontraditional Students
Active Control			
I felt that I had a lot of control over my visiting experiences at this Web site	.57	.55	.69
While I was on the Web site, I could choose freely what I wanted to see	.78	.84	.73
While surfing the Web site, I had absolutely no control over what I can do on the site*	.67	.70	.66
While surfing the Web site, my actions decided the kind of experiences I got	.50	.50	.51
Two-Way Communication			
The Web site is effective in gathering visitors' feedback	.71	.64	.73
This Web site facilitates two-way communication between the visitors and the site	.75	.72	.74
It is difficult to offer feedback to the Web site*	.61	.53	.68
The Web site makes me feel it wants to listen to its visitors	.71	.76	.71
The Web site does not at all encourage visitors to talk back*	.61	.74	.59
The Web site gives visitors the opportunity to talk back	.67	.71	.66
Synchronicity			
The Web site processed my input very quickly	.57	.60	.50
Getting information from the Web site is very fast	.83	.93	.75
I was able to obtain the information I want without any delay	.89	.92	.88
When I clicked on the links, I felt I was getting instantaneous information	.82	.79	.83
The Web site was very slow in responding to my requests*	.62	.71	.57

*These items are reverse scaled.

Table 3

Model Fit Indices for Six Competing Models

Competing Models	χ^2	Df	Relative χ^2	χ^2 Difference	CFI	NNFI	RMSEA
Null model	524.34**	105	5.00	NA	NA	NA	NA
One-factor model	275.49**	90	3.06	248.85**	.56	.48	.20
Two-factor correlated model (AC and TW combined)	169.72**	89	1.91	105.77**	.81	.77	.11
Two-factor correlated model (TW and SY combined)	252.75**	89	2.84	22.74**	.61	.54	.19
Two-factor correlated model (AC and SY combined)	147.99**	89	1.66	127.50**	.86	.83	.10
Three-factor uncorrelated model	144.64**	90	1.61	NA	.87	.85	.09
Three-factor correlated model	123.99*	87	1.43	45.73**[a] 128.76**[b] 24.00**[c]	.91	.90	.07

Note: AC = active control, TW = two-way communication, SY = Synchronicity
a Chi-square difference over the two-factor model (AC and TW combined)
b Chi-square difference over the two-factor model (TW and SY combined)
c Chi-square difference over the two-factor model (AC and SY combined)
* $p < .01$ ** $p < .001$

Table 4

Standardized Correlation Matrix and Average Variance Extracted

Standardized Correlation Between Factors:			
	Study 1	Study 2	Study 3
Active Control and Two-Way Communication	.38 [.14]	.45 [.20]	.23 [.05]
Two-Way Communication and Synchronicity	.32 [.10]	.48 [.23]	.31 [.10]
Active Control and Synchronicity	.51 [.26]	.68 [.46]	.52 [.27]
Average Variance Extracted for Each Factor			
	Study 1	Study 2	Study 3
Active Control	.53	.50	.51
Two-Way Communication	.54	.60	.62
Synchronicity	.57	.51	.73

*The numbers in brackets are squared pairwise correlations.

Table 5

Mean Interactivity Scores of the Two Web Sites used in Study 2 and Study 3

	Active Control	Two-Way Communication	Synchronicity
Study 2			
Less Interactive Web Site	5.15[a]	4.13[a]	5.08[a]
More Interactive Web Site	5.72[b]	4.78[b]	5.54[b]
Study 3			
Less Interactive Web Site	5.34[A]	4.06[a]	4.56[a]
More Interactive Web Site	5.69[B]	4.86[b]	5.09[b]

Note: Within a column, mean scores in the same study with different upper case letter superscripts were significantly different at $p < .10$; mean scores with different lower case letter superscripts were significantly different at $p < .05$.

Table 6

Examples of User Responses for the Three Interactivity Dimensions

Dimensions	**Examples of User Responses**
Active Control	"It gave you the control of what you want to look at." (favorable) "Ad kept popping up – bothering me." (unfavorable)
Two-Way Communication	"The support section is set up very well." (favorable) "I dislike the fact that the site is somewhat indifferent to public opinion." (unfavorable)
Synchronicity	"It is very quick in loading up pages. This is nice because you don't have to wait long." (favorable) "The site also loaded fairly slow." (unfavorable)
Other	"Its range of products was too limited."

Table 7

Average Cronbach's Alpha of the Scale for Nontraditional and Traditional Students

	Active Control	Two-Way Communication	Synchronicity
Traditional Students	.71	.84	.89
Nontraditional Students	.75	.84	.83