



Contents lists available at ScienceDirect

Information & Management

journal homepage: www.elsevier.com/locate/im



Designing usable online stores: A landscape preference perspective

Younghwa Lee^{a,*}, Kenneth A. Kozar^{b,1}

^aSchool of Business, University of Kansas, Lawrence, KS 66045, United States

^bLeeds School of Business, University of Colorado at Boulder, Boulder, CO 80309, United States

ARTICLE INFO

Article history:

Received 25 January 2008

Received in revised form 7 September 2008

Accepted 9 November 2008

Available online xxx

Keywords:

Website usability

Theoretical model of website usability

Kaplan's landscape preference model

Structural equation modeling

ABSTRACT

Developing an explanatory theoretical model of website usability is pivotal for understanding usable website design. Such a model would explain and help predict the effects of website usability on online purchases, but few studies have been devoted to such model development. As an exploratory effort, we adopted and extended Kaplan's landscape preference model by including factors of legibility, coherence, variety, and mystery, and examined their effect on cognitive and affective appraisals and their impact on purchase intention. A field survey with 495 online customers using two different categories of websites was conducted to validate the research model. A multi-group analysis with gender and age was performed to cross-validate it. Our findings demonstrated that the proposed model explained a large amount of the variance of purchase intention, invariant across different subgroups. Key implications for theory and practice are discussed.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Usable websites are pivotal for e-business success. They create a positive attitude toward an online store, increases stickiness and revisit rates, and can stimulate online purchases [35] while providing benefits by reducing web development and maintenance costs. However, current websites still have numerous usability problems [7]. This study focused on developing a model of usability as a way to understand web design factors and their effect on the perceptions of online consumers. Such a model could also provide guidelines for the design of a usable website and the metrics needed to evaluate its usability and for comparing it with those of its competitors.

We proposed and validated a new model of website usability by using Kaplan's landscape preference model [12]. It posits that people develop a positive attitude toward a scene which has coherent, legible, various, and mysterious features. We believe that the model can be applied successfully in a website design. Online customers develop a positive attitude toward a website when they can understand structure, content, and features quickly (coherent); easily navigate it (legibility); enjoy vivid and dynamic images (variety); and be stimulated with features inspiring their curiosity (mystery). A website is an emulation of a physical store and visitors

with a mental representation of it are likely to prefer and revisit it [41]. Successful online stores have features found in the brick and mortar world.

2. Literature review

2.1. Identify website usability factors

Considerable effort has been expended to identify website usability factors, as well as examining their effects on online consumer perceptions or behaviors across multiple disciplinary areas including marketing, human–computer interaction (HCI), and IT acceptance. Researchers in marketing [44] have developed online consumer behavior models to investigate factors affecting online purchases, including marketing, demographic, personality, and usability variables. HCI research generally took an engineering approach for an objective assessment of website design quality. For example, Gehrke and Turban [10] suggested page loading, navigation efficiency, download time, successful search rate, error rates, task completion time, and frequency of cursor movement as usability factors. Website design experts have specified content quality, fun, productivity, and relevance, navigation, response time, and credibility [29]. Based on industry guidelines from companies like IBM and Microsoft and competitions such as the Webby awards, other usability factors were added including content, ease-of-use, promotion, made-for-the-medium, and emotion [42]. However, as the influence of website design factors on online customers' purchases was realized, IS researchers investigated their exclusive effects on online purchasing [8,46].

* Corresponding author. Tel.: +1 785 864 7559; fax: +1 785 864 5328.

E-mail addresses: gabelee@ku.edu (Y. Lee), kenneth.kozar@colorado.edu (K.A. Kozar).

¹ Tel.: +1 303 492 8347; fax: +1 303 492 5962.

Collectively, our understanding of usability factors and their influence on online consumer behavior has been enriched as shown in Appendix A. However, because of the lack of fundamental theories of website usability, the studies had proposed different sets of website usability factors; used different terminology and scope; and investigated different nomological networks between usability factors and online customer perceptions. Therefore, the previous studies did not provide coherent lenses to observe and explain the distinct contribution of website usability on online purchase behavior.

2.2. Lack of a theoretical model of website usability

There are few theoretical frameworks relating the factors and outcomes of website designs and no underlying model or theory of the process exists to be confirmed or refuted. Usability researchers [28] indicated that there is a lack of the theoretical models of website usability proven empirically or theoretically, resulting in the prevalence of incorrect or inapplicable usability guidelines. Such models would be useful because design decisions provide too many alternatives to test by trial and error. A strong theoretical model could reduce the alternatives to a manageable number.

Although previous theoretical models (e.g., [15,42]) have provided useful insights, there is room for improvement because some did not fully investigate the causal network of variables, attitude, and behavioral intention; some did not conduct analyses to test the goodness-of-fit of the proposed models; some used many single item constructs; and finally some were not tested in Western countries. Therefore, a theoretical model which resolved such problems would be of value.

2.3. Kaplan’s landscape preference model

Kaplan’s model examined physical environments in an attempt to develop design patterns that incorporated end users’ use of environmental cues, making it easier to process information and function effectively and enjoyably. The model explains landscape preference using a 2 × 2 perceptual structure, as shown in Table 1.

The model uses the concepts of understanding and exploring. *Making sense* refers to the concern to comprehend, to keep one’s bearings, to understand what is going on in the immediate here and now and often in some larger world as well, while *involvement* refers to the concern to figure out, to learn, to be stimulated [14]. Making sense relates to the operations in the environment, which takes in stimuli to make it easier to characterize and summarize (coherence) and easier to map (legibility). The supportive environment of involvement contains rich landscape components (variety). Involvement also is related to the process of engaging and sustaining one’s interest in an environment (mystery). In addition, the model explains landscape preference based on two- and three-dimensional visual perception and depicts humans as cognitive creatures who can decide on future possibilities of present landscape choices. The decision process consists of both immediate and future calculations, which can occur quickly but sequentially.

Immediate (two-dimensional) perceptions are related to coherence and variety, whereas future (three-dimensional) perceptions are related to legibility and mystery [13]. The primary level of the model represents immediate perceptions of the

Table 1
 Structure of landscape preference.

	Making sense	Involvement
Two-dimensional	Coherence	Variety
Three-dimensional	Legibility	Mystery

elements in the landscape, which allow rapid assessment based on a surface examination. Settings that are orderly (coherent) increase the individual’s ability to make sense of the landscape, and richness of the elements (variety) enhance involvement. This immediate assessment is followed by an inference of what is deeper within a landscape. Its components help visitors to stay oriented (legibility) and provide perceived expected curiosity and fancy (mystery) in understanding an unfamiliar landscape and encourage further navigation.

2.4. Developing a theoretical model of website usability

Kaplan’s model provides a conceptual foundation for developing a theoretical model of website usability. First, perceptions of physical space can share common properties with those of cyberspace. Second, the perception development process of seeing physical space is similar to that of online visitors. People who visit a particular landscape develop their preference after visiting, entering into, walking around, and comparing it with other landscapes. Similarly, online consumers develop their preference toward a particular website after visiting, searching and navigating it, and then comparing it with other sites. Through this process, people perceive whether an interesting space or cyberspace has coherent, various, legible, and mysterious features. Third, the visitors at an interesting physical space combine their perceptions from the direct and inferred features. This can also be true in visiting cyberspace. Online visitors develop their direct perceptions quickly by looking at the home page of the website. They further develop their perceptions by navigating and experiencing a wealth of design components in the underlying web pages.

However, Kaplan’s model only captures the molar-level perceptions of website usability. Individuals develop their perceptions toward space by combining their molecular-level perception (“trees”) with molar-level perception (“forest”). The perceptions from objective and tangible design components (e.g., white space in a page, use of Sans Serif font, presence of security features) are molecular-level, while the abstract and global perceptions of the website design (e.g., coherence) are molar-level. By only assessing the molar-level perceptions, the model provides limited practical guidelines to practitioners, although it can help designers develop a fundamental conceptual foundation for designing a usable website.

Although the proposed model captures individuals’ hedonic perceptions of the website by use of affective appraisal and mystery, it is expanded by including more hedonic variables (e.g., enjoyment, playfulness) to better address the effect of web usability in the context of website design. Consequently the model is appropriate for designing a usable utilitarian e-business website that evaluates the influence of molar-level perceptions of website design factors on online purchases.

3. Hypotheses development

To answer the fundamental question of whether landscape preference factors can be operationalized as website preference factors in a website context and whether they significantly influence online consumers’ perceptions and behavior, this study utilizes Kaplan’s model to propose hypotheses. Our main assumption is that a website developed with legible, coherent, diverse and mysterious usability factors can stimulate a positive attitude from both a cognitive and affective perspective which triggers online purchase intention. Legibility represents ease of navigating of a site. It helps users acquire more of the information they are seeking, since the information is easier to find [25]. Diverse search engines, intelligent agents, and landmarks are examples of legibility; then online consumers experience a light

cognitive load, low error rates, and less disorientation and will stay longer at sites creating enjoyment and satisfaction. Thus, online customers will be pleased when they can navigate the website easily and quickly to find the information they want. They also will perceive usefulness when they can find the product or information efficiently and effectively. Therefore, we hypothesized that

H1a. Legibility of websites significantly influences cognitive appraisals.

H1b. Legibility of websites significantly influences affective appraisals.

Coherence in virtual space is to the ability of a website to provide consistent and orderly contents, structures, and multi-media components. Online consumers easily grasp the organization of a website by providing a common look and feel to each page, and this positively affects their attitude towards the website. Websites often use cascading style sheets to provide the same structure, color, font, and frames across web pages; visitors see the same menu bar, navigation bar, company logo, search engine, font, and white space in the same location across pages. This results in faster search, less disorientation, and ease of learning which triggers a positive cognitive appraisal [30]. Therefore, we hypothesized

H2. Coherence of websites significantly influences cognitive appraisals.

Variety in virtual space involves to the ability of a website to provide diverse website components that create vivid interaction and communication with its customers. Various website design components help customers develop a sense of presence in a virtual environment, resulting in greater engagement. Web designers use multiple design components to increase this variety perception. For example, some provide diverse search mechanisms with a number of options classifying products based on brand, price, bestseller, service provider, and whether the product is used or new. Other sites provide customers with diverse ways to see the products (showing front, side, back, top view, etc.) or video clips providing 360° rotation. Diverse communication channels, such as a live chat with customer support or discussion boards help online customers and provide vivid experiences. People have more enjoyment and fun when they visit such websites. At the same time, people will perceive usefulness or effectiveness when they can gather rich information by using several options and can see information or images of the product of interest before making their purchase decision. Therefore, we hypothesized:

H3a. Variety of websites significantly influences cognitive appraisals.

H3b. Variety of websites significantly influences affective appraisals.

In virtual space, *mystery* refers to the ability of a website to invoke curiosity and interest and stimulate further viewing. For example, 'bonus', 'hot buy', 'learn more', 'compare price', or 'try it' signs may make online customers curious, and this may positively affect their appraisal [20]. Therefore, we hypothesized that:

H4. Mystery of websites significantly influences affective appraisals.

Cognitive appraisals refer to evaluations based on beliefs and knowledge structures and affective appraisals refer to those based on emotions, feelings and reactions [4]. Both appraisals affect customers' attitude when they evaluate products; they compare their affective and utilitarian value separately before purchases,

and thus successful product designers should carefully develop the site to promote both values.

Recently, hedonic system features, such as playfulness, fun, and excitement, have become important in developing sites. For example, Te'eni [39] indicated that the site should have both cognitive and affective appraisals, so as to build a more accurate representation of actual behavior. On this basis, we divided attitude into the two forms and hypothesized that cognitive and affective appraisals significantly influenced purchase intention. This lead to:

H5. Cognitive (H5a) and Affective (H5b) appraisals significantly influence Online Purchase Intention.

4. Research methodology

Development of measures requires careful analysis, as they become the building blocks for generating valid relationships among the variables. The development of instrument items to measure website usability allows e-business managers to make comparisons, benchmark performance, and plan improvements. In our study, multi-staged scientific instrument development and validation procedures [37] were used. In addition, by analyzing the data using SEM, nomological networks between endogenous and exogenous variables were examined.

4.1. Instrument development

To develop the instrument items for coherence, legibility, variety, and mystery we first conducted a literature review of previous studies that performed similar testing. However, most prior studies had used single item measures and thus we reviewed published and generally accepted concepts of the landscape preference constructs and generated a set of 24 items to measure coherence, legibility, variety, and mystery. Five experts were selected and provided the definition of each construct and a related item description. Then they were asked to place the items in one of the four constructs or a 'not applicable' category. After completion of two-rounds of this instrument purification processes, a total of 16 items had been selected. The content validity of all instrument items was then tested and the wording, item order, content, and format of the questionnaire was examined and modified. We conducted an exploratory factor analysis with an oblique rotation and found that they were grouped well into four constructs with a greater than 1 Eigen value. Each item was formatted into a 7-point Likert-type scale.

A total of 41 undergraduate students from an e-business class in a large Midwestern US university, having participated in at least one prior online purchase, were recruited to pretest the instrument. Without informing the participants of the goal of the study, the participants were asked to navigate Amazon.com before completing a questionnaire. Participants were motivated by giving them an opportunity to win a sweepstakes and receive class participation points. Appendix B shows the instrument items employed in the main survey.

4.2. Main survey

We selected an online electronics site (Amazon.com) and three online travel sites (Orbitz.com, Travelocity.com, and Expedia.com) as targets. These were top-ranked websites in Alexa Global 500 rankings of average consumer visits and were well recognized by our student subjects. We selected well-known websites because the theoretical model measured website usability and needed enough subjects having familiarity with the sites to assess the quality of web design components quickly. Prospective subjects

Table 2
Demographics of target subjects.

Category		Online electronics (n = 184)	Percentage	Online travel (n = 294)	Percentage
Gender	Males	104	57%	178	61%
	Females	80	43%	116	39%
Age	Below 20	2	1%	5	2%
	20–29	111	60%	182	62%
	30–39	63	34%	75	26%
	Over 40	8	4%	32	11%
Average online experience (at electronics sites or travel sites)	Years < 1	10	5%	11	4%
	1 ≤ years < 3	14	8%	155	53%
	3 ≤ years < 5	85	46%	123	42%
	6 ≤ years < 7	56	30%	5	2%
	7 or more years	19	10%	0	0%
Yearly online purchases	1 ≤ times < 2	33	18%	52	18%
	2 ≤ times < 3	34	18%	86	29%
	3 ≤ times < 4	65	35%	75	26%
	4 ≤ times < 5	29	16%	44	15%
	5 or more times	23	13%	37	13%

experienced less stress when they knew the site. Also, the need for a small training time for subjects to understand design features helped reduce the total time to conduct the study. Further using the same target sites allowed us to compare our findings with those of some previous studies.

The population of the main study was online purchasers or individuals who had experienced at least one online purchase in the past at either an online electronics or travel site. The study was advertised in local newspapers, online bulletin boards, and students taking summer-semester classes at a large Midwestern US university resulting in a total of 900 potential subjects (371 for the online electronics site and 529 for online travel sites). After distributing the questionnaires, 184 usable responses were gathered for the online electronics site (after removing 18 incomplete ones), resulting in a net response rate of 49.5% (184/371), while 294 usable questionnaires were gathered for the online travel sites (after removing 23 incomplete ones) for a net response rate of 55.5% (294/529). Demographics of the subjects for the online electronics and online travel sites are shown in Table 2.

Participants navigated the website following an online purchase scenario provided by us before they completed the questionnaire. The goal of the scenario was to stimulate participants' memory of the website in assessing its legibility, coherence, variety, and mystery features. An example of the scenario is shown in Appendix C. In order to ensure serious engagement while site visiting, the participants were asked to answer a customized list of questions. There were no significant individual differences. Some of the questions were

1. What was today's recommended item on the first page? [(a) watch, (b) mp3 player, (c) camcorder]
2. What was the last minute deal? [(a) vacation to Mexico, (b) vacation to Hawaii, (c) vacation to Caribbean]

Participation was voluntary and the process took approximately 30 min. Subjects were compensated by being entered into a sweepstake. Student subjects also received class participation points.

5. Results

Structural equation modeling (SEM) was used to perform both measurement and structural model analysis simultaneously. The analysis validated the psychometric properties of the measures, and was used to investigate nomological network relationships

between constructs in the model. Data were analyzed using AMOS 7.0. We adopted a two-step approach to cross-validate the model. Amazon.com data was used as a calibration sample and the online travel sites data was used as the validation sample.

5.1. Calibration sample analysis: Amazon.com

5.1.1. Measurement model analysis

A confirmatory factor analysis was conducted to validate the psychometric properties of the instrument. They were measured by examining whether the measurement model had an acceptable goodness-of-fit and investigating its unidimensionality, convergent and discriminant validity, and reliability. Overall goodness-of-fit for the model was confirmed. The $\chi^2/d.f.$ was 1.47, which is below the desired threshold of 3.0. The RMSEA was 0.051, which is below the 0.08 cut-off. All NFI (0.906) and CFI (0.967) were above their corresponding cut-off value of 0.90. These results suggested that the measurement model adequately fit the data.

Convergent validity was evaluated using three criteria: (1) all indicator factor loadings (λ) should be significant at $p < 0.05$ and exceed 0.7, (2) composite reliabilities should exceed 0.7, and (3) average variance extracted (AVE) by each construct should exceed the variance due to measurement error for that construct [9]. As shown in Table 3, all factor loadings exceeded 0.7 and were significant at $p = 0.001$. Composite reliabilities ranged between 0.843 and 0.922 and AVE values were well above the cut-off value of 0.50, which is greater than variance due to measurement error. Therefore, all three conditions for convergent validity were met.

Discriminant validity was assessed by constraining the estimated correlation parameters (ϕ_{ij}) between constructs to 1.0 and then performing a chi-squared difference test on the values obtained for the constrained and unconstrained models. The chi-squared differences between these models were significant at $p < 0.05$, demonstrating clear discriminant validity among these constructs (detailed results will be provided upon request). As a supplementary assessment of discriminant validity, inter-construct correlations were used (see Table 4). All constructs were found to have a stronger correlation with their own measures than to those of other constructs. All the correlations between the constructs were less than 0.7 and less than the square root value of average variance extracted shown in the diagonal, representing appropriate discriminant validity. Finally, *reliability* was examined using Cronbach α and all constructs showed a value of over 0.84, indicating a high reliability of items used for each construct.

Table 3
Results of measurement model analysis.

Construct	Items	Loadings	Cronbach's α	Composite reliabilities	Average variance extracted
Legibility	LEG1	0.799	0.900	0.900	0.694
	LEG2	0.896			
	LEG3	0.817			
	LEG4	0.816			
Coherence	COH1	0.800	0.896	0.897	0.685
	COH2	0.829			
	COH3	0.833			
	COH4	0.847			
Variety	VAR1	0.852	0.904	0.905	0.703
	VAR2	0.840			
	VAR3	0.813			
	VAR4	0.849			
Mystery	MYS1	0.918	0.920	0.922	0.747
	MYS2	0.869			
	MYS3	0.826			
	MYS4	0.841			
Affective appraisals	AFF1	0.781	0.847	0.843	0.642
	AFF2	0.810			
	AFF3	0.813			
Cognitive appraisals	COG1	0.857	0.915	0.912	0.723
	COG2	0.872			
	COG3	0.841			
	COG4	0.830			
Purchase intention	PI1	0.854	0.842	0.845	0.732
	PI2	0.857			

5.1.2. Structural model analysis

Fig. 1 shows the result of the structural model analysis, including the R^2 and path loadings for all hypothesized relationships. The model was found to explain online purchase intention, successfully confirming that it is appropriate for website design. The analysis provided two results: the goodness-of-fit of the comprehensive model and all relationships among constructs based on the hypotheses. The fit statistics ($\chi^2/d.f. = 1.510$, $NFI = 0.900$, $CFI = 0.963$, $RMSEA = 0.053$) indicated that the model provided a good fit to the data. All goodness-of-fit statistics were above their cut-off values.

Most components of H1–H5 showed significant influence on Cognitive and Affective Appraisals. Legibility (H1a: $\lambda = 0.431$,

$p < 0.001$), Coherence (H2: $\lambda = 0.307$, $p < 0.001$), and Variety (H3a: $\lambda = 0.176$, $p < 0.05$) strongly influenced Cognitive Appraisals and explained its large variance ($R^2 = 0.635$). Variety (H3b: $\lambda = 0.470$, $p < 0.001$) and Mystery (H4: $\lambda = 0.357$, $p < 0.001$) showed strong influence on Affective Appraisals ($R^2 = 0.615$), while Legibility (H1b: $\lambda = 0.019$, $p > 0.05$) did not. As hypothesized in H5, both Cognitive Appraisals and Affective Appraisals were found to have significant effect on online purchase intention. Especially, Cognitive Appraisals (H5a: $\lambda = 0.506$, $p < 0.001$) showed stronger influence on Purchase Intention than Affective Appraisals (H5b: $\lambda = 0.314$, $p < 0.001$). Both Cognitive Appraisals and Affective Appraisals explain 66.3% of variances of Purchase Intention.

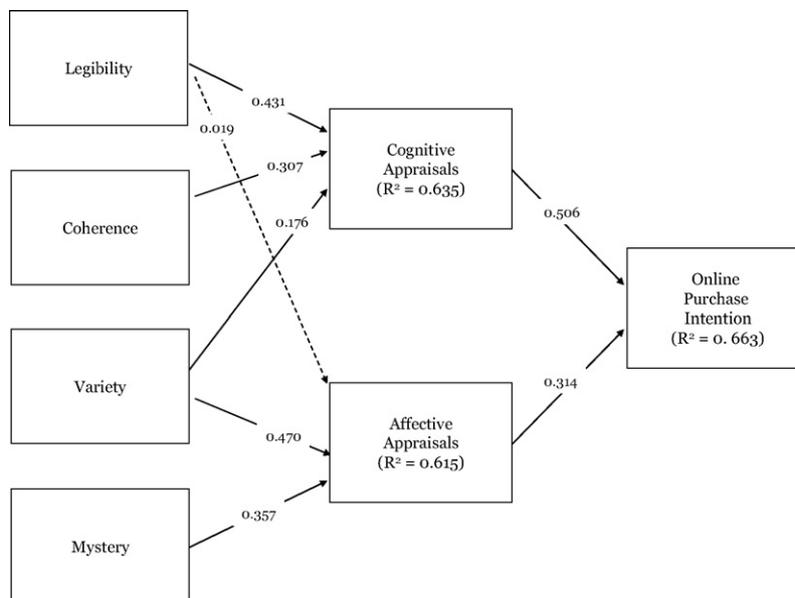


Fig. 1. Structural model analysis (Amazon.com).

Table 4
Inter-construct correlation matrix.

	1	2	3	4	5	6	7
1. Legibility	0.833						
2. Coherence	0.556	0.827					
3. Variety	0.572	0.528	0.838				
4. Mystery	0.547	0.575	0.681	0.864			
5. Affective Appraisal	0.446	0.473	0.661	0.626	0.801		
6. Cognitive Appraisal	0.657	0.614	0.591	0.460	0.384	0.850	
7. Purchase Intention	0.580	0.599	0.514	0.531	0.483	0.658	0.855

Values shown in the main diagonal represent the square root of AVE.

5.2. Validation sample analysis: online travel websites

The proposed model was cross-validated using the data from three online travel websites. The t-test for demographic variables between subjects of these websites showed no significant differences (e.g., $F_{2,291} = 0.113, p > 0.05$), hence data were pooled to increase the statistical power. The psychometric properties of the measurement instruments were measured based on a confirmatory factor analysis and found to have good psychometric properties. Convergent and discriminant validity also met the recommended criteria (see [2]). All the constructs showed over 0.80 Cronbach α , indicating that high reliability was achieved. The analysis of the structural equation model was conducted and found to have a good fit to the data ($\chi^2/d.f. = 2.294, NFI = 0.909, CFI = 0.946, RMSEA = 0.066$) (Table 5).

As shown in Fig. 2, hypothesized relationships except legibility-affective appraisals ($\gamma = 0.025, p > 0.05$) were significant at $p < 0.05$ and thus were supported.

5.3. Common method variance test

Since our study utilized self-reported purchase intention measures, the common method effect was investigated using the correlated uniqueness model to validate conclusions about the identified relationships between constructs [26]. By investigating the chi-square difference ($\chi^2/d.f.$) and model fit between the uncorre-

Table 5
Results of measurement model analysis (online travel sites).

Construct	Items	Loadings	Cronbach's α	CFR	AVE
Legibility	LEG1	0.807	0.892	0.893	0.677
	LEG2	0.874			
	LEG3	0.757			
	LEG4	0.848			
Coherence	COH1	0.876	0.934	0.935	0.782
	COH2	0.877			
	COH3	0.901			
	COH4	0.884			
Variety	VAR1	0.903	0.907	0.909	0.716
	VAR2	0.832			
	VAR3	0.796			
	VAR4	0.849			
Mystery	MYS1	0.879	0.899	0.901	0.695
	MYS2	0.838			
	MYS3	0.807			
	MYS4	0.808			
Affective Appraisals	AFF1	0.768	0.882	0.885	0.722
	AFF2	0.847			
	AFF3	0.926			
Cognitive Appraisals	COG1	0.818	0.909	0.909	0.715
	COG2	0.871			
	COG3	0.862			
	COG4	0.829			
Purchase Intention	PI1	0.872	0.840	0.846	0.733
	PI2	0.840			

Table 6
Common method variance test.

	Recommended value	Uncorrelated model	Correlated model
$\chi^2/d.f.$	<3.00	407.3/260	387.9/246
GFI	>0.90	0.858	0.864
AGFI	>0.80	0.822	0.821
NFI	>0.90	0.896	0.901
CFI	>0.90	0.959	0.961
RMSEA	<0.08	0.056	0.056

lated model and the correlated model, the significance of the method effect was determined. A significant effect does not exist if there is no chi-square difference and no significant model fit improvement. Table 6 shows the comparison: no significant method effect was found. First, there was no significant chi-square difference ($\chi^2/d.f. = 407.3/260 - 387.9/246 = 1.386, p > 0.05$). Second, the model fit was not improved. Therefore, there was no significant common method variance, indicating that the identified relationships between constructs were not strongly affected by the method effect.

6. Discussion

Our findings show that the model had strong psychometric properties and explained most of the variance of attitudes and purchase intention of online consumers, indicating that the model could be considered as an alternative model in evaluating the usability of a website.

Both cognitive and affective appraisals significantly influenced a customer's intention to purchase online. Compared to prior IS, successful websites have had to incorporate many hedonic features to meet users' need for enjoyment, fun, and arousal. Also, a strong influence of legibility, coherence, variety, and mystery on the appraisals was noticed, indicating that most preference constructs were appropriate determinants of a positive attitude toward a website, suggesting that a blend of legible, coherent, diverse, and mysterious website design components and content can successfully communicate with online consumers and invoke positive appraisals. Legibility was found to be the strongest factor. Both variety and mystery were found to have significant impact on affective appraisals suggesting that website designers should provide relevant content that stimulate consumers' fun, excitement, and curiosity.

Analysis was conducted to examine whether the models and instruments were invariant across different subgroups. Appendix D provides the detailed findings of the analysis. However, several interesting differences between groups were found. For age, except for the relationship of Coherence-Cognitive Appraisals ($\lambda = 0.069, p > 0.05$) of the 30 or below age-group and the relationship of Variety-Cognitive Appraisals ($\lambda = 0.121, p > 0.05$), Variety-Affective Appraisals ($\lambda = 0.185, p > 0.05$), and Legibility-Affective Appraisals ($\lambda = 0.051, p > 0.05$) of over 30 age-group, all nomological networks were validated across different age groups. Interestingly, affective appraisals showed stronger influence on purchase intention than cognitive appraisals for the 30 or below age-group, but the result was reversed for the over 30 age-group. This implies that older people made more conservative decisions. In addition, for the younger group, variety was the strongest factor affecting cognitive appraisals, while for the older group, mystery was the strongest.

For gender, all hypothesized relationships were significant except Legibility-Cognitive Appraisals ($\lambda = 0.131, p > 0.05$) and Legibility-Affective Appraisals ($\lambda = 0.142, p > 0.05$) for the male group and Legibility-Cognitive Appraisals ($\lambda = 0.225, p > 0.05$) and Variety-Affective Appraisals ($\lambda = 0.136, p > 0.05$) for the female group. Cognitive appraisals showed a stronger influence on purchase intention than affective appraisals on males, while this was reversed for females. Males may be more materialistic and task-oriented while females may be more emotional. Mystery was

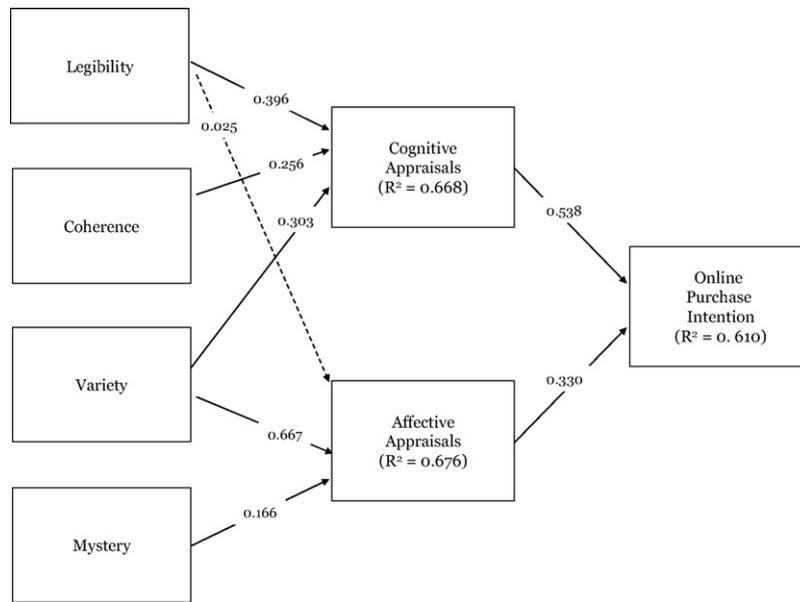


Fig. 2. Structural model analysis: online travel sites.

the strongest factor affecting female affective appraisals, whereas variety was strongest for males.

7. Contributions

Our study had several limitations. First, the two sectors of websites used do not necessarily represent all e-business domains. Second, although we conducted a common method bias test and confirmed no significant self-selection bias present, objective online purchase data should have been gathered to prevent self-selection bias. Third, by selecting well-known and well-organized websites equipped with similar web design features, we could not capture much variation among the sites. Finally, this we overlooked several website usability constructs (e.g., content and security) because of the theoretical boundary of the original model.

Despite its limitations, this study could aid by providing useful knowledge for usable website design. From a theoretical perspective, the study has at least three implications. There were few theoretical frameworks for usable website design, resulting in a lack of a conceptual base for designers. By including new usability dimensions such as legibility and coherence, the new model provided a new way to understand the unrevealed effect of website usability on perceptions of online consumers. This study also seeks to enrich our understanding of the effect of different website

usability factors on an individual’s online purchases by investigating the perceptual differences between men and women and between older and younger consumers. Given the extensive role of websites in e-commerce and the increasing presence of women and older consumers, understanding gender and age differences could prove essential in developing a better website that meets the needs of all online consumers.

Our study also has practical implications. Management and designers of e-businesses sites can evaluate the usability level of their own website using the model. Companies can also conduct a questionnaire survey from representative customers to evaluate the usability level of their website using our instrument. At the same time, the company can ask its website developers to evaluate the website and learn from any differences from those of its competitors in understanding relative strengths and weaknesses. Given this comparison, companies can make strategic and resource allocation decisions on how to improve their websites.

In summary, our study proposed and validated a new theoretical lens to help in understanding the effect of website usability on online purchases.

Appendix A. Summary of previous studies of website usability variables

Field	Ref.	IV: Website usability factors	DV			
			Attitude	Intention	Purchase/use	Others
Marketing	[4]	Informativeness, organization, entertainment	✓			Satisfaction/loyalty
	[11]	Information quality, security, website functionality, customer relationships, responsiveness				
	[22]	Consistent menu bars, number of levels, store entrances, lists with buttons, lists of pictures, feedback, FAQ			✓	
	[34]	Coherence, Web page involvement	✓	✓		
HCI	[38]	Fast presentation, ease of navigation, uncluttered screens				Satisfaction
	[16]	Accessibility, product-search, shopping-basket-handling, product overview, shop response time		✓	✓	
	[19]	Font type, line length				Performance
	[29]	Navigation, response time, credibility, content				
	[33]	Simplicity, multimedia				Satisfaction/performance
	[36]	Ease of use, readability, content quality, fun, productivity, completeness, relevance				Web quality
	[40]	Interactivity, effectiveness, efficiency	✓			
	[45]	Ease of use, sense of presence, usefulness				Satisfaction

Appendix A (Continued)

Field	Ref.	IV: Website usability factors	DV			
			Attitude	Intention	Purchase/use	Others
IT adoption	[1]	System quality, information quality	✓	✓		
	[17]	Value-added search mechanisms, challenging, shopping enjoyment, usefulness, ease of use		✓		
	[18]	Accuracy, user involvement, convenience, ease of use		✓		
	[21]	Information, learning, playfulness, system quality, service quality	✓			
	[27]	information relevance, timeliness, reliability, scope, access, usability, navigation, interactivity				Satisfaction
	[31]	Download delay, navigability, site content, interactivity, responsiveness		✓	✓	Satisfaction
	[32]	Information content, design		✓		
	[35]	Ease of use, navigation	✓	✓		
Usability	[3]	Site design, information, empathy, usability				Web quality
	[5]	Navigation design, visual design, information design				Satisfaction/loyalty/trust
	[15]	Firmness, convenience, delight				Satisfaction/loyalty
	[23]	Tailored information, visual appeal, intuitive operations, ease of understanding, response time		✓		
	[24]	Interactivity (active control, two-way communication, synchronicity)				Satisfaction
	[28]	Content, functionality, structure,		✓		
	[43]	Content, ease of use, promotion, made-for-the-medium, emotion			✓	
	[46]	Personalization, structure, navigation, layout, searchability, fast response				Satisfaction

Appendix B. Instrument items

B.1. Affective appraisals

Visiting the website is

- AFF1. Dull...Exciting
- AFF2. Pleasant...Unpleasant
- AFF3. Enjoyable...Unenjoyable

B.2. Cognitive appraisals

- COG1. The website is effective to achieve my goals
- COG2. The website is convenient to achieve my goals
- COG3. I feel comfortable using the website to achieve my goals
- COG4. The website is helpful to achieve my goals

B.3. Purchase intention

Assuming that you visit a web portal (or an online travel site) like this, please indicate how strongly you disagree or agree with the following statements.

- PI1. I intend to purchase products or services from the website.
- PI2. I predict I would purchase products or services from the website

B.4. Legibility

- LEG1. It is clear where I can go in the website.
- LEG2. It is easy to get around the whole website.
- LEG3. It does not take much time to figure out a way of moving around the website.
- LEG4. I can always figure out where I am.

B.5. Coherence

- COH1. Each component of the website is well related to each other.

- COH2. Components of the website work well together.
- COH3. Each component of the website seems to hang together.
- COH4. Each component of the website helps each other to provide better comprehension to visitors.

B.6. Variety

- VAR1. The website has too many distractions, making it confusing (Reversed).
- VAR2. The website does not contain enough components to interest me (Reversed).
- VAR3. The website contains a good variety of components that keep me involved.
- VAR4. I feel drawn in by the variety of information or components the website offers.

B.7. Mystery

- MYS1. The website makes me feel there is something interesting to explore.
- MYS2. As I navigate through the website, more curiosity inspires me.
- MYS3. I expect that the website will provide interesting things to invoke my curiosity as I explore around.
- MYS4. I feel I will find interesting things if I navigate more.

Appendix C. An example of online purchasing scenario at Amazon.com

To refresh your memory with past online purchases, we ask that you navigate the Amazon.com website while following the instructions below. Please read through the instructions briefly and then proceed to www.Amazon.com. Again, this should take no more than 15 min.

Assume you are a customer who wants to buy a new cellular phone through Amazon.com. Following the guidelines below proceed through the complete purchase process but do not actually purchase a cellular telephone.

1. Review the Amazon.com homepage for a couple of minutes. Make note of the look and feel of the site, the content provided on the homepage and the various menus and hyperlinks. Make yourself familiar with the structure, features, and design of Amazon.com.
2. Using whatever means you choose, locate the cellular phones.
3. Taking no more than 5 min, select your cellular phone using whatever means you would normally use were you to actually be considering the purchase of the telephone. You may compare features, images, service plans, service providers, and rebate information of the various offerings.
4. Once you have selected a particular phone for further review, click through the various customer images, read the product description, any special promotions, and customer reviews.
5. Click Help to review available customer support information.
6. Select the phone of your choice and proceed to Checkout. Enter all information that you need to complete the checkout process except the credit card information. If you are a member of Amazon.com, login to your account during the checkout process. If you are not a member of Amazon.com, please register as a new member by filling out online application. Continue the checkout by providing address, shipping method, and any other requested information until you reach to the 'Place Order' screen.
7. Return to the Amazon.com homepage.
8. Review the 'Privacy Notice', 'Return Policy' and 'Conditions of Use'.
9. If any remaining time from your allotted 15 min is left, continue to check whether features (e.g., content, structure, color, buttons, hyperlinks, etc) of this website are easy-to-learn, consistent, concise, and readable.

Appendix D. Multi-group analysis

The multi-group analysis has two main advantages: (1) it allows a test of the generalizability of measurement items by comparing item-factor loadings across different groups, and (2) the structural weights are directly comparable by using equivalent measurement where observed scores from different groups are in the same scales. The multi-group analysis recently has been adopted by researchers to validate the generalizability of well-known constructs including end-user computing satisfaction and models including the technology acceptance model [6]. The multi-group analysis consists of two sequential processes: (1) a comprehensive assessment of item-factor loadings and model-data fit for each subgroup, and (2) structural weights invariance test.

D.1. Item-factor loadings and model-fit analysis

The item-factor loadings and model-fit analysis were conducted across gender and age. As shown in Table C-1, the item-factor loadings remained high across different groups (from 0.734 to 0.914). Adequate model-fit was also found across different groups as shown in Table C-2. Although NFI scores slightly below 0.90 for subjects over 30 years old and male and female, they have other model-fit indices that could be considered adequate and excellent item-factor loadings. Based on these findings, we can conclude that the data for each group fit the model sufficiently well to further process the multi-group analysis of structural invariance.

D.1.1. Item-factor loadings across different groups

Construct	Items	Age		Gender	
		30 or Below (n = 265)	Over 30 (n = 184)	Male (n = 169)	Female (n = 194)
Legibility	LEG1	0.858	0.821	0.858	0.820
	LEG2	0.767	0.796	0.810	0.814
	LEG3	0.878	0.875	0.897	0.879
	LEG4	0.803	0.796	0.816	0.789
Coherence	COH1	0.906	0.856	0.870	0.813
	COH2	0.879	0.840	0.854	0.844
	COH3	0.856	0.836	0.849	0.834
	COH4	0.889	0.834	0.828	0.779
Variety	VAR1	0.777	0.741	0.734	0.739
	VAR2	0.804	0.804	0.796	0.801
	VAR3	0.817	0.797	0.811	0.827
	VAR4	0.864	0.805	0.854	0.878
Mystery	MYS1	0.802	0.743	0.763	0.764
	MYS2	0.757	0.769	0.766	0.745
	MYS3	0.872	0.820	0.865	0.840
	MYS4	0.888	0.832	0.888	0.856
Affective Appraisals	AFF1	0.964	0.829	0.861	0.845
	AFF2	0.909	0.746	0.874	0.902
	AFF3	0.885	0.742	0.836	0.861
Cognitive Appraisals	COG1	0.836	0.848	0.850	0.860
	COG2	0.879	0.874	0.853	0.841
	COG3	0.845	0.875	0.849	0.881
	COG4	0.863	0.870	0.767	0.783
Purchase intention	PI1	0.868	0.864	0.846	0.857
	PI2	0.855	0.863	0.886	0.896

D.1.2. Fit Index across different groups

Fit Index	Age		Gender	
	30 or Below	Over 30	Male	Female
χ^2 /d.f.	1.595	1.842	1.478	1.424
RMSEA	0.053	0.073	0.059	0.052
NFI	0.908	0.864	0.885	0.885
CFI	0.963	0.932	0.959	0.962

D.2. Multi-group analysis of structural invariance

The first step of a multi-group analysis of structural invariance is to develop an equal pattern baseline model. As shown in Model I in Table C-3, this equal pattern model has good model fit indicating that the same patterns of parameters could fit the data for all different groups. The second step is to investigate measurement equivalence by forcing each item-factor loading to be equal across the different groups resulting in Model 2. Model 2 also has adequate model-fit. The chi-square difference test between Model 2 (χ^2 /d.f. = 1942.1/11027) and Model 1 (χ^2 /d.f. = 1885.9/1003) was conducted and found that it is insignificant (56.2 with 24 d.f.). This implies that the measurement model is invariant across subgroups.

D.2.1. Measurement equivalence

Model	Model	χ^2	d.f.	p value	RMSEA	NNFI	CFI
Model 1	Equal patterns	1885.9	1003	.000	0.033	0.885	0.942
Model 2	Factor loadings invariant	1942.1	1027	.000	0.033	0.882	0.940

D.3. Standardized structural weights

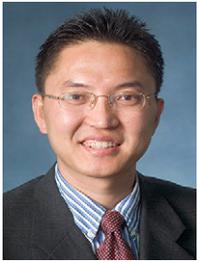
The standardized structural weights for the endogenous and exogenous variables are shown in Table C-4 for all different groups. Each model showed a strong goodness of fit explaining a large amount of variance in cognitive and affective appraisals and purchase intention. However, there were significant differences in the structural weights among different groups. For age, except for the relationship of Coherence-Cognitive Appraisals ($\lambda = 0.069, p > 0.05$) for the 30 and below age-group and the relationship of Variety-Cognitive Appraisals ($\lambda = 0.121, p > 0.05$), Variety-Affective Appraisals ($\lambda = 0.185, p > 0.05$), and Legibility-Affective Appraisals ($\lambda = 0.051, p > 0.05$) for over 30 age-group, all other nomological networks were validated across different age groups. For gender, all hypothesized relationships were significant except Legibility-Cognitive Appraisals ($\lambda = 0.131, p > 0.05$) and Legibility-Affective Appraisals ($\lambda = 0.142, p > 0.05$) for the male group and Legibility-Cognitive Appraisals ($\lambda = 0.225, p > 0.05$) and Variety-Affective Appraisals ($\lambda = 0.136, p > 0.05$) for the female group.

D.3.1. Standardized structural weights: age and gender

D.V.	I.V.	Age		Gender	
		30 or Below	Over 30	Males	Females
Purchase Intention	Cognitive Appraisals	0.390	0.492	0.490	0.339
	Affective Appraisals	0.605	0.452	0.425	0.610
Cognitive Appraisals	Legibility	0.365	0.444	0.131	0.225
	Coherence	0.069	0.439	0.557	0.235
	Variety	0.408	0.121	0.191	0.317
Affective Appraisals	Variety	0.220	0.185	0.399	0.136
	Mystery	0.359	0.402	0.361	0.448
	Legibility	0.266	0.051	0.142	0.355

References

- [1] T. Ahn, S. Ryu, I. Han, The impact of web quality and playfulness on user acceptance of online retailing, *Information and Management* 44, 2007, pp. 263–275.
- [2] J.C. Anderson, D.W. Gerbing, Structural equation modeling in practice: a review and recommended two-step approach, *Psychological Bulletin* 103 (3), 1988, pp. 411–423.
- [3] S.J. Barnes, R.T. Vidgen, Data triangulation and web quality metrics: a case study in e-government, *Information and Management* 43, 2006, pp. 767–777.
- [4] Q. Chen, W.D. Wells, Attitude toward the site, *Journal of Advertising Research* 39, 1999 (September/October), pp. 27–37.
- [5] D. Cyr, Modeling website design across cultures: relationships to trust, satisfaction and e-loyalty, *Journal of Management Information Systems* 24 (4), 2008, pp. 47–72.
- [6] X. Deng, W.J. Doll, A.R. Hendrickson, J.A. Scazzero, A multi-group analysis of structural invariance: an illustration using the technology acceptance model, *Information and Management* 42, 2005, pp. 745–759.
- [7] J. Ethier, P. Hadaya, J. Talbot, J. Cadieux, B2C web site quality and emotions during online shopping episodes: an empirical study, *Information and Management* 43, 2006, pp. 627–639.
- [8] X. Fang, C.W. Holsapple, An empirical study of website navigation structures' impacts on website usability, *Decision Support Systems* 43, 2007, pp. 476–491.
- [9] C. Fornell, V.F. Larcker, Evaluating structural equation models with unobservable variables and measurement error, *Journal of Marketing Research* 18, 1981, pp. 39–50.
- [10] D. Gehrke, E. Turban, Determinants of successful web site design: relative importance and recommendations for effectiveness, in: *Proceedings of the 31st Hawaii International Conference on Information Systems*, 1999.
- [11] C.I. Ho, Y.L. Lee, The development of an e-travel service quality scale, *Tourism Management* 28 (6), 2007, pp. 1434–1449.
- [12] R. Kaplan, S. Kaplan, *The Experience of Nature: A Psychological Perspective*, Cambridge University Press, Cambridge, England, 1989.
- [13] S. Kaplan, Aesthetics, affect, and cognition: environmental preference from an evolutionary perspective, *Environment and Behavior* 19, 1987, pp. 3–32.
- [14] S. Kaplan, *Perception and Landscape: Conceptions and Misconceptions*, Cambridge University Press, 1988.
- [15] J. Kim, J. Lee, K. Han, M. Lee, Business as buildings: metrics for the architectural quality of internet businesses, *Information Systems Research* 13, 2002, pp. 239–254.
- [16] U. Konradt, H. Wandke, B. Balazs, T. Christophersen, Usability in online shops: scale construction, validation and the influence on the buyers' intention and decision, *Behaviour & Information Technology* 22 (3), 2003, pp. 165–174.
- [17] M. Koufaris, Applying the technology acceptance model and flow theory to online consumer behavior, *Information Systems Research* 13 (2), 2002, pp. 205–223.
- [18] Z. Liao, M.T. Cheung, Internet-based e-banking and consumer attitudes: an empirical study, *Information and Management* 39 (4), 2002, pp. 283–295.
- [19] J. Ling, P. van Schaik, The influence of font type and line length on visual search and information retrieval in web pages, *International Journal of Human-Computer Studies* 64 (5), 2006, pp. 395–404.
- [20] J.A. Litman, Curiosity and the pleasures of learning: wanting and liking new information, *Cognition and Emotion* 19 (6), 2005, pp. 793–814.
- [21] C. Liu, K.P. Arnett, Exploring the factors associated with web site success in the context of electronic commerce, *Information and Management* 38, 2000, pp. 23–33.
- [22] G.L. Lohse, P. Spiller, Internet retail store design: how the user interface influences traffic and sales, *Journal of Computer Mediated Communication* 5 (2), 1999.
- [23] E.T. Loiacono, R.T. Watson, D.L. Goodhue, WebQual: an instrument for customer evaluation for web sites, *International Journal of Electronic Commerce* 11 (3), 2007, pp. 51–87.
- [24] P.B. Lowry, T. Spaulding, T. Wells, G. Moody, K. Moffit, S. Madariaga, A theoretical model and empirical results linking website interactivity and usability satisfaction, in: *Proceedings of the 39th Hawaii International Conference on System Sciences*, Big Islands, HI, 2006.
- [25] S. Machlis, Site redesigns keep it simple, *Computerworld* 32 (43), 1998, p. 25.
- [26] H.W. Marsh, M. Bailey, Confirmatory factor analyses of Multitrait-Multimethod data: a comparison of alternative models, *Applied Psychological Measurement* 15, 1991, pp. 47–70.
- [27] V. McKinney, K. Yoon, F.M. Zahedi, The measurement of web-consumer satisfaction: an expectation and disconfirmation approach, *Information Systems Research* 13 (3), 2002, pp. 296–315.
- [28] S. Mithas, N. Ramasubbu, M.S. Krishnan, C. Fornell, Designing web sites for customer loyalty across business domains: a multilevel analysis, *Journal of Management Information Systems* 23 (3), 2006, pp. 97–127.
- [29] J. Nielsen, *Designing Web Usability*, New Riders, Indianapolis, IN, 2000.
- [30] A.A. Ozok, G. Salvendy, How consistent is your web design? *Behaviour and Information Technology* 20 (6), 2001, pp. 433–447.
- [31] J.W. Palmer, Web site usability, design, and performance metrics, *Information Systems Research* 13 (2), 2002, pp. 151–167.
- [32] C. Ranganathan, S. Ganapathy, Key dimensions of business-to-consumer web sites, *Information and Management* 39 (6), 2002, pp. 457–465.
- [33] P.-L.P. Rau, Q. Gao, J. Liu, The effect of rich web portal design and floating animations on visual search, *International Journal of Human-Computer Interaction* 22 (3), 2007, pp. 195–216.
- [34] A.E. Schlosser, A. Kanfer, Interactivity in commercial web sites: implications for web site effectiveness, in: *Proceedings of the Individual Differences in Consumer Behavior on the web*, Special Topic Session, 2001.
- [35] J. Song, F.M. Zahedi, A theoretical approach to web design in e-commerce: a belief reinforcement model, *Management Science* 51 (8), 2005, pp. 1219–1235.
- [36] J.M. Spool, T. Scanlon, W. Schroeder, C. Snyder, T. DeAngelo, *Web Site Usability: A Designer's Guide*, Morgan Kaufmann Publishers, San Francisco, CA, 1999.
- [37] D. Straub, Validating instruments in MIS research, *MIS Quarterly* 13, 1989, pp. 147–169.
- [38] D.M. Szymanski, R.T. Hiss, E-satisfaction: an initial examination, *Journal of Retailing* 76 (3), 2000, pp. 309–322.
- [39] D. Te'eni, A cognitive-affective model of organizational communication for designing it, *MIS Quarterly* 25 (2), 2001, pp. 251–312.
- [40] H.-H. Teo, L.-B. Oh, C. Liu, K.-K. Wei, An empirical study of the effects of interactivity on web user attitude, *International Journal of Human-Computer Studies* 58 (3), 2003, pp. 281–305.
- [41] B. Tversky, Structures of mental spaces: how people think about spaces, *Environment & Behavior* 35, 2003, pp. 66–80.
- [42] V. Venkatesh, R. Agarwal, Turning visitors into customers: a usability-centered perspective on purchase behavior in electronic channels, *Management Science* 52 (3), 2006, pp. 367–382.
- [43] V. Venkatesh, V. Ramesh, Web and wireless site usability: understanding differences and modeling use, *MIS Quarterly* 30 (1), 2006, pp. 181–206.
- [44] M. Wolfinbarger, M.C. Gilly, EtailQ: dimensionalizing, measuring, and predicting Etail quality, *Journal of Retailing* 79, 2003, pp. 183–198.
- [45] S. Yoon, J. Laffey, H. Oh, Understanding usability and user experience of web-based 3D graphics technology, *International Journal of Human-Computer Interaction* 24 (3), 2008, pp. 288–306.
- [46] M. Zviran, C. Glezer, I. Avni, User satisfaction from commercial web sites: the effect of design and use, *Information and Management* 43, 2006, pp. 157–178.



Younghwa Lee is an Assistant Professor of Information Systems at the University of Kansas School of Business. He received his PhD from University of Colorado/Boulder in 2005. His research interest is in website usability, technology acceptance, and IT ethics and security. He is an ICIS 2003 Doctoral Consortium fellow. He has published in *Communications of the ACM*, *Data Base*, *Decision Support Systems*, *Journal of Organizational Computing and Electronic Commerce* among others.



Kenneth A. Kozar is a Professor of Information Systems at the University of Colorado/Boulder Leeds School of Business. His interests lie in the area of human and organizational impacts of technology. He has published in a number of journals, served two terms as an associate editor of the *MIS Quarterly*, and was the chair of the Society for Information Management's International Paper Award Competition.