Developing and validating an instrument for measuring user-perceived web quality

By: Adel M. Aladwani, Prashant C. Palvia

Aladwani, A.M., and Palvia, P. "Developing and Validating an Instrument for Measuring User-Perceived Web Quality." Information & Management. Vol. 39, 2002, pp. 467-476.

Abstract:
Many of the instruments to measure information and system quality were developed in the context of mainframe and PC-based technologies of yesteryears. With the proliferation of the Internet and World Wide Web applications, users are increasingly interfacing and interacting with web-based applications. It is, therefore, important to develop new instruments and scales, which are directly targeted to these new interfaces and applications. In this article, we report on the development of an instrument that captures key characteristics of web site quality from the user's perspective. The 25-item instrument measures four dimensions of web quality: specific content, content quality, appearance and technical adequacy. While improvements are possible, the instrument exhibits excellent psychometric properties. The instrument would be useful to organizations and web designers as it provides an aggregate measure of web quality, and to researchers in related web research.
Keywords: Internet; Electronic commerce; Web site quality; Instrument development; Web measures; Web design

Article:

1. Introduction

The last few years have witnessed a technological revolution fueled by the wide spread use of the Inter-net, web technologies and their applications. The literature is replete with accounts about organizations from various types and sizes integrating web technologies into their operations [9,20,27,34,44,47]. This wide interest in the subject is attributed to the fact that organizations are becoming aware of the possible implications of the Internet on their work. Potentialities of web applications are remarkable leading many organizations to spend awesome amounts of money on these technologies. Using web technologies, an organization can reach out to customers and provide them with not only general information about its products or services but also the opportunity of performing interactive business transactions. Organizations investing in web technologies and applications are looking forward to realizing the benefits of these investments, however, this would not be possible without an appropriate tool for measuring the quality of their web sites.

Construct measurement in general and in the context of web technologies and applications in particular is a challenging task, hence, deserves more attention from researchers interested in this phenomenon. However, web site quality measurement is neither simple nor straightforward. Web quality is a complex concept; therefore, its measurement is expected to be multi-dimensional in nature. Current research on web quality seems to pay less attention to construct identification and measurement efforts. Only limited academic research exists, but it is fragmented and usually only discusses the meaning of some aspects of web quality. At the practitioner level, several commercial ranking systems are available to rate web sites according to certain quality attributes, e.g. Web Awards (www.webaward.org) and The Web Awards (www.thewebawards.com). These ratings, however, lack clarity in terms of criteria used and the ranking methodology. The problem is exacerbated by the paucity of systematic and empirically derived systems for rating the quality of web sites. Today no multi-item scale is available to measure users' perceived web quality. The provision of such a scale will further enhance management's ability to exploit the full potential of the Internet.

The objective of this study is to develop an instrument to measure perceived web quality from the perspective of Internet users. This instrument and proposed scale would be valuable to researchers and practitioners interested in designing, implementing, and managing web sites. The rest of the paper is organized as follows. The next section reviews prior research related to quality in IS/IT and more specifically, web quality. The following sections and its various subsections describe the measurement and the various steps involved in scale development. The final section discusses applications of the proposed scale.

2. The concept of quality in IS research

Quality is not a new concept in information systems management and research. Information systems practitioners have always been aware of the need to improve the information systems function so it can react to external and internal pressures and face the critical challenges to its growth and survivability [1]. Moreover, information systems scholars have been concerned with definitions of quality in information systems research. Information systems researchers have attempted to define data quality [21], information quality [23], software/system quality [35], documentation quality [14], information systems function service quality [22], and global information systems quality [31]. More recently, there has been some effort to define quality in the context of the Internet [25]. However, the web quality concept remains underdeveloped.

Web quality is a vastly undefined concept. For the most part, existing scientific research discusses the meaning of some aspects of web quality in a descriptive manner without delineating its major dimensions or providing tested scales to measure it. For example, Liu and Arnett [25] named such quality factors as accuracy, completeness, relevancy, security, reliability, customization, interactivity, ease of use, speed, search functionality, and organization. Huizingh [19] focused on two aspects of web quality: content and design. Wan [45] divided web quality attributes into four categories: information, friendliness, responsiveness, and reliability. Rose et al. [36] provided a theoretical discussion of technological impediments of web sites. The authors highlighted the importance of factors such as download speed, web interface, search functionality, measurement of web success, security, and Internet standards. Misic and Johnson [29] suggested such web-related criteria as finding contact information (e.g. e-mail, people, phones, and mail address), finding main page, speed, uniqueness of functionality, ease of navigation, counter, currency, wording, and color and style.

Olsina et al. [33] specified quality attributes for academic web sites. These authors took an engineering point of view and identified factors such as cohesiveness by grouping main control objects, direct controls permanence, contextual controls stability, etc. Bell and Tang [5] identified factors such as access to the web, content, graphics, structure, user friendliness, navigation, usefulness, and unique features. Another useful stream of research is by Ho [18] and Sakaguchi et al. [38] where they investigate key website characteristics on the purpose and value dimensions. While the purpose dimension relates directly to the contents of the site, the value dimension relates more to the quality aspects.

The trade press and Internet sources also discussed some aspects of web quality attributes. For example, Schacklett [39] proposed nine tips for improving web site quality, including effective use of graphics and colors, 24/7 web site accessibility, and ease of web site use and navigation. Levin [24] offered tips to help a company with web site design including fast web page download, web page interactivity, and current content, among other factors. Wilson [48] recommended avoiding seven mistakes relevant to web site design. Furthermore, based on their own personal experience, Barron et al. [4] recommended 39 guidelines relevant to web site graphics, text, links, page size and length, and multimedia.

Three conclusions can be drawn from the above review. First is that past web quality research, albeit useful, is fragmented and focuses only on subsets of web quality. For example, Rose et al. [36] list six factors, and Bell and Tang [5] mention eight factors. Misic and Johnson’s [29] study was more extensive, but it missed several critical factors such as web security, availability, clarity, and accuracy, to name a few. Liu and Arnett [25] list 11 items and two dimensions of web quality—information and system quality. Like the other studies, several important quality dimensions are missing from the authors’ list. Second, past research lacks rigor when it comes to measuring the web quality construct. In some cases, an ad hoc web quality tool was suggested [5,25,29].

However, it is not clear to the reader what is the domain of the measured construct or what refinement, validation, and normalization procedures were employed. For example, Liu and Arnett's scale included items about information to support business objectives, empathy to customers' problems, and follow-up services to customers. These items loaded on the same factor, which they called "information quality". In addition, double barreled items can be found in their scale, e.g. security and ease of use. Third, the majority of the suggested web quality attributes and scales are relevant to web designers than to web users; like for instances, the ideas and scales proposed by Liu and Arnett [25] and Olsina et al. [33].

The previous discussion underscores the fact that the web quality construct lacks a clear definition and web quality measurement is still in its infancy. With this background, we embark on developing a sound instrument to measure user-perceived web quality.

3. The measurement process

A number of psychometric researchers have pro-posed several procedural models to help other researchers develop better scales for their studies, e.g. [3,7]. Applying these concepts, MIS researchers have developed several instruments, e.g. the end user computing satisfaction by Doll and Torkzadeh [11] and the microcomputer playfulness instrument by Webster and Martocchio [46]. Straub [42] described a process for creating and validating instruments in IS research, which includes content validity, construct validity and reliability analyses. Three generic steps common in all these models include (1) conceptualization, (2) design, and (3) normalization. Conceptualization focuses on content validity, involves such activities as defining the construct of interest and generating a candidate list of items from the domain of all possible items representing the construct. The second step, design focuses on construct validity and reliability analysis. It pertains to the process of refining the sample of items from the previous step to come up with an initial scale, deciding on such operational issues as question types and question sequence, and pilot-testing the initial scale that has been developed in the preparation stage. The third and last step concerns the effort to normalize the scale that has been developed. It involves the important steps of subsequent independent verification and validation. Unfortunately, this step is omitted in many scale development efforts. In the conduct of these steps, several analytical techniques can be used such as factor analysis and reliability analysis as we will describe next.

3. 1. Conceptualization

The first step in our measurement process is conceptualization. This step involves delimiting the domain of the construct and generating sample items representing the concept under consideration. In order to ensure content validity, the instrument needs to draw representative items from a universal pool [8]. In the present study, we define perceived web quality as users' evaluation of a web site's features meeting users' needs and reflecting overall excellence of the web site. This definition is important to delimit the domain of web quality and determine relevant literature from which the researcher can generate sample items for the web quality construct.

Overall, our review of the academic literature and relevant trade press articles identified three dimensions of web quality: technical adequacy, web content, and web appearance; and yielded 102 representative items. The sample items were initially assessed using a Delphi method. Two information systems scholars were asked to evaluate the items and make changes to eliminate repetitive items, technical/non-user oriented items, and sub-attributes of higher level attributes. After three evaluation rounds, 55 web quality attributes remained in the list. The experts deleted items such as: relative versus absolute links, invalid internal links, invalid external links, make paragraph text flush left, disclaimer note, quality of link phrase, color of hyperlinks, what's new feature, counter availability, long domain name, time and date, vertical scrolling, horizontal scrolling, errors free site, etc. Table 1 summarizes web quality dimensions and sample items.

3.2. Design

The second step in our model is scale design. As mentioned earlier, the focus here is on construct validity and reliability analysis. These are in essence operational issues and investigate whether the measures chosen are true constructs describing the event or merely artifacts of the methodology itself [6,8]. Several tests were conducted in order to refine the instrument. We started out by arranging the selected items in a questionnaire format in preparation for data collection. The items were measured using a seven-point scale ranging from (1) ‘‘extremely not important’’ to (7) ‘‘extremely important’’. Following Churchill’s [7] recommendations, we subjected the instrument to a two-stage data collection and refinement procedure. The first stage was used for design and the second stage for normalization.

In first stage of data collection, the 55-item instrument was administered to student web users enrolled in three different sections of an introductory information systems class at a business school. A total of 104 web users participated in our study. Of these, we collected usable responses from 101 web users. All of the students in the sample were from 18 to 21 years of age. Approximately, 64% of the respondents are females, and 36% are males. Of the students that participated, a majority were business majors.

We first computed reliability coefficients of the scales using Cronbach’s alpha. The alpha values for technical adequacy, web content, and web appearance came as 0.77, 0.70 and 0.59, respectively. Reliability tests suggested that screening the data along Churchill’s recommendations would improve reliability levels. We screened the collected data by discarding items that showed very low corrected item-total correlations, i.e. <0.40. After several screening attempts, 30 items remained in our pool of items. Reliability levels for the reduced web quality dimensions came as 0.89, 0.86 and 0.81 for technical adequacy, web content, and web appearance, respectively.

Next, we factor analyzed the 30-item instrument to examine the dimensionality of the construct. In our quest for a stable factor structure, we followed an iterative procedure that began with submitting the items to a factor analysis procedure with varimax rotation. Hair et al. [16] suggest that item loadings >0.30 are considered significant, >0.40 are more important, and >0.50 are considered very significant. There are no accepted “absolute” standards for the cut-offs; the choice is based on judgment, purpose of the study, and prior studies. Since our goal is to examine the most significant loadings in interpreting the factor solution, we decided to use a cut-off point of 0.50 for item loadings and eigenvalue of 1. After the first iteration, we examined items loadings and eliminated items that did not meet the loading cut-off or loaded on more than one factor. We then resubmitted the remaining items to another round of factor analysis. The process went on until we reached a meaningful factor structure. The factor analysis revealed seven factors with eigenvalue of ≥ 1 (see Table 2). The scree test, however, indicated that a four-factor solution was appropriate. At the end of the factor analysis procedure, 25 items remained.

Table 1
Major web quality dimensions

Dimension	Sample items	Sample support references
Technical adequacy	Security; ease of navigation; broadcast services; limited use of special plug-ins; search facilities; anonymity; availability; valid links; reliability; browser sniffing; personalization or customization; speedy page loading; interactivity; ease of access; multi-language support; protected content; bookmark facility	[25]; [10]; [13]; [15]; [24]; [33]; [36]; [43]; [12]
Web content	Usefulness of content; completeness of content; clarity of content; uniqueness of content; broadness of content; originality of content; currency of content; conciseness of content; accuracy of content; finding contact info.; finding people without delay; finding site maintainer; finding links to relevant sites; finding firm’s general info.; finding products/services details; finding customers’ policies; finding customer support; finding FAQ list; finding free services; using limited registration forms; finding online help; diversity of content; finding free info	[25]; [26]; [29]; [37]; [41]; [4]; [5]; [17]
Web appearance	Attractiveness; distinctive hot buttons; changing look; organization; proper use of fonts; proper use of colors; proper use of graphics; graphics-text balance; proper use of multimedia; style consistency; proper choice of page length; good labeling; text-only option; proper use of language/style; color consistency	[19]; [30]; [39]; [28]; [29]; [40]; [2]; [4]; [17]

The results showed that there were four web quality dimensions not three, as proposed earlier. It was found that web content is not unidimensional, but comprises two dimensions: specific content and content quality. Specific content (five items) reflected concerns related to finding specific details about products/services, customer support, privacy policies, and other important information. Content quality consisted of five items and dealt with such attributes as information usefulness, completeness, accuracy, and so on and so forth. Another round of reliability tests resulted in alpha values of 0.91, 0.91, 0.85 and 0.87, for technical adequacy, specific content, content quality, and appearance, respectively.

Table 2

Principal component analysis with varimax rotation—first study

	Component						
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
Security	0.70	0.05	0.14	0.22	0.06	0.38	−0.06
Ease of navigation	0.74	0.07	0.13	0.12	0.15	0.19	0.17
Broadcast services	0.12	0.19	0.32	0.07	0.44	0.06	0.47
Search facilities	0.76	0.15	0.09	0.14	0.05	0.04	−0.03
Availability	0.78	0.08	0.08	0.10	−0.19	0.01	−0.09
Valid links	0.77	−0.12	0.08	0.13	0.00	0.12	0.11
Reliability	0.21	−0.03	−0.10	−0.18	0.01	0.70	0.37
Browser sniffing	0.25	−0.06	0.25	−0.04	0.71	0.01	0.09
Personalization or customization	0.69	−0.02	0.02	0.01	0.10	−0.03	0.08
Speed of page loading	0.78	0.11	0.24	0.01	0.20	−0.15	−0.09
Interactivity	0.69	0.07	0.10	0.01	−0.19	−0.11	0.44
Ease to access the site	0.72	0.13	0.20	0.11	0.22	0.02	−0.07
Usefulness	0.13	0.14	−0.03	0.77	0.08	−0.01	−0.02
Completeness	0.21	0.10	0.08	0.71	0.05	−0.17	−0.03
Clarity	0.02	0.18	0.12	0.81	−0.08	−0.04	0.06
Currency	0.07	0.08	0.16	0.73	−0.12	0.10	−0.12
Conciseness	0.04	0.17	0.15	0.78	0.08	0.06	−0.06
Accuracy	0.16	−0.02	−0.02	0.66	−0.20	−0.02	0.24
Finding contact information	0.07	0.89	0.03	0.12	0.04	−0.02	−0.30
Finding firm's general information	0.06	0.85	−0.02	0.15	−0.08	0.08	0.22
Finding products/services	0.07	0.89	0.03	0.12	0.04	−0.02	−0.30
Finding customers' policies	0.11	0.77	−0.06	0.17	−0.08	−0.13	0.12
Finding customer support	0.04	0.81	−0.05	0.14	−0.01	0.14	0.33
Attractiveness	0.26	0.01	0.74	0.02	−0.05	−0.07	0.19
Organization	0.24	0.05	0.70	0.09	0.06	0.19	−0.12
Proper use of fonts	0.07	−0.08	0.85	0.15	0.04	−0.04	−0.05
Proper use of colors	0.08	0.04	0.82	0.06	−0.05	0.10	0.00
Proper use of multimedia	0.13	−0.09	0.81	0.11	0.05	0.06	0.07
Style consistency	−0.02	0.10	0.29	0.13	−0.68	0.05	0.10
Good labeling	0.03	0.04	0.31	0.03	−0.05	0.70	−0.20
Eigenvalue	5.28	3.83	3.69	3.63	1.46	1.37	1.18
Variance explained	0.176	0.127	0.123	0.121	0.049	0.046	0.039

3.3. Normalization

To verify and validate the four dimensions of perceived web quality, the 25-item instrument was tested using another independent data set. A sample of 127 students web users enrolled in four different sections of an introductory information systems class at a business school were asked to participate in second study and to indicate their agreement or disagreement with the 25 statements. The users were assigned to four different groups to evaluate the sites of a bank (25 students), a bookshop (31 students), a car manufacturer (34 students), and an electronics retailer (37 students). The goal was to examine reliability of the proposed instrument across web types. Before completing the questionnaire, the users were asked to read a brief overview of the study and were instructed to navigate through the assigned web site for some time. At the end of this exercise, we collected usable responses from all the participants. Of the students that participated in second study, a majority were business majors and 61% were females. All of the students in the second sample, like in the first sample, were from 18 to 21 years of age.

Reliability scores were 0.92,0.94,0.88 and 0.88, for technical adequacy, specific content, content quality, and appearance, respectively. The overall reliability of the 25-item scale was 0.91. Cronbach's alpha of the same scale for the bank subgroup is 0.95, for the bookshop subgroup is 0.91, for the electronics retailer subgroups is 0.70, and for the car manufacturer sub-group is 0.92. These scores are above the reliability cut-off points suggested by Nunnally [32].

Testing for convergent and discriminant validity allowed us to further assess construct validity. Convergent validity was first examined by submitting each scale's items to a principal component analysis procedure with varimax rotation to explore the underlying dimensions of the construct. The usual cut-off point of 0.50 for item loading and eigenvalue of 1 were used in second study like first study. All sub-scales of perceived web quality showed adequate convergent validity (Table 3). In every factor analysis run, each scale's items converged cleanly on the same factor representing these items. The results of the factor analysis procedure were consistent with the results of the first study. Table 3 shows that there are four factors explaining 67% of the variance in perceived web quality.

Table 3
Principal component analysis with varimax rotation—second study

	Component			
	Factor 1	Factor 2	Factor 3	Factor 4
Security	0.73	0.08	0.27	0.11
Ease of navigation	0.80	0.13	0.14	0.06
Search facilities	0.77	0.20	0.17	−0.02
Availability	0.75	0.08	0.17	0.02
Valid links	0.81	−0.04	0.16	0.08
Personalization or customization	0.73	0.03	0.05	0.01
Speed of page loading	0.77	0.18	0.04	0.15
Interactivity	0.75	0.04	0.08	0.09
Ease of accessing the site	0.69	0.26	0.11	0.14
Usefulness	0.20	0.28	0.72	0.10
Completeness	0.23	0.07	0.74	0.07
Clarity	0.12	0.10	0.85	0.08
Currency	0.09	0.09	0.79	0.06
Conciseness	0.08	0.22	0.77	0.14
Accuracy	0.21	0.04	0.70	0.07
Finding contact information	0.10	0.91	0.13	0.12
Finding firm general information	0.16	0.87	0.15	0.08
Finding products/services details	0.09	0.90	0.13	0.12
Finding customers' policies (e.g. dispute policies)	0.14	0.82	0.18	0.04
Finding customer support	0.19	0.83	0.13	0.09
Attractiveness	0.20	0.08	0.05	0.76
Organization	0.16	0.17	0.08	0.74
Proper use of fonts	0.01	0.05	0.17	0.84
Proper use of colors	−0.01	0.11	0.04	0.85
Proper use of multimedia	0.08	−0.01	0.13	0.85
Eigenvalue	5.49	4.13	3.85	3.44
Variance explained	0.219	0.165	0.154	0.137

We also used the multitrait-multimethod matrix (MTMM) approach [6] to evaluate the convergent and discriminant validity of the instrument. Convergent validity determines whether associations between scales of the same group are higher than zero and large enough to proceed with discriminant validity tests. In the present case, for every single variable, the correlations in the validity diagonal (i.e. items of the same variable) are higher than zero. All 71 correlations between items within the variables were significant at the 0.01 level. The smallest within-variable correlations for the various variables are technical adequacy: 0.436, specific content: 0.442, content quality: 0.655, and appearance: 0.515.

We examined discriminant validity for each item by counting the number of times an item correlates higher with items of other variables than with items of its own variable. For instance, the lowest own-variable correlation for technical adequacy is 0.436, and none of the correlations of technical adequacy with items of other variables are >0.436, i.e. number of violations is zero. Campbell and Fiske suggest that: for discriminant validity, the number of violations should be <50% of the potential comparisons. We retain items where violations are <50% as per this criterion, and reject items where violations are >50%. Discriminant validity results showed that all of the tests exceeded the benchmark and were at 0% violations and this was true for all four sub-scales. Thus, the factor and scale structure reported in the first study was corroborated in the second study.

As further evidence of the validity of the perceived web quality (PWQ) construct and its four dimensions, we examined the relationship between the construct scale ratings and users' overall quality rating (OQR) for a wellknown web site. The web users in second study were also asked to evaluate the overall quality of the web site on a three-point scale: (1) fair, (2) good and (3) excellent. Table 4 summarizes our results and reports the correlation matrix along with variables' means and standard deviations. The four dimensions of perceived web quality correlated significantly with each other and with the overall index of perceived web quality: PWQ. The highest correlation between the variables comprising the four dimensions perceived web quality was between technical adequacy and content quality (Pearson's $r = 0.38$), whereas the lowest correlation was that between technical adequacy and appearance (Pearson's $r = 0.23$). In addition, the four dimensions of perceived web quality correlated significantly with users' overall quality rating for the web site; the association between OQR and technical adequacy was the highest (Pearson's $r = 0.73$), whereas the association between OQR and specific content was the lowest (Pearson's $r = 0.30$). The results clearly give further credence to the sound psychometric properties of the instrument.

Table 4
Correlations among constructs and descriptive statistics

	Appearance	Specific content	Content quality	Technical adequacy	Perceived web quality	Overall quality rating
Specific content	0.23**					
Content quality	0.25**	0.35**				
Technical adequacy	0.23**	0.32**	0.38**			
Perceived web quality	0.57**	0.65**	0.76**	0.72**		
Overall quality rating	0.43**	0.30**	0.53**	0.73**	0.73**	
Mean	4.37	4.42	4.38	4.49	4.41	2.28
S.D.	1.10	1.09	1.08	1.19	0.71	0.61

** $P < 0.01$.

Table 5
The user-perceived web quality instrument

	Strongly disagree				Strongly agree		
1. _____'s web site looks secured for carrying out transactions (e.g. uses SSL, digital certificates, etc.)	1	2	3	4	5	6	7
2. _____'s web site looks easy to navigate through	1	2	3	4	5	6	7
3. _____'s web site has adequate search facilities	1	2	3	4	5	6	7
4. _____'s web site is always up and available	1	2	3	4	5	6	7
5. _____'s web site has valid links (hyperlinks)	1	2	3	4	5	6	7
6. _____'s web site can be personalized or customized to meet one's needs	1	2	3	4	5	6	7
7. Web pages load fast in _____'s web site	1	2	3	4	5	6	7
8 _____'s web site has many interactive features (e.g. online shopping, etc.)	1	2	3	4	5	6	7
9. _____'s web site is easy to access (i.e. has a reflective and widely registered name)	1	2	3	4	5	6	7
10 The content of _____'s web site is useful	1	2	3	4	5	6	7
11 The content of _____'s web site is complete	1	2	3	4	5	6	7
12. The content of _____'s web site is clear	1	2	3	4	5	6	7
13. The content of _____'s web site is current	1	2	3	4	5	6	7
14 The content of _____'s web site is concise	1	2	3	4	5	6	7
15. The content of _____'s web site is accurate	1	2	3	4	5	6	7
16. In _____'s web site, one can find contact information (e.g. e-mail addresses, phone numbers, etc.)	1	2	3	4	5	6	7
17. In _____'s web site, one can find firm's general information (e.g. goals, owners)	1	2	3	4	5	6	7
18. In _____'s web site, one can find details about products and/or services	1	2	3	4	5	6	7
19. In _____'s web site, one can find information related to customers' policies (e.g. privacy and dispute details)	1	2	3	4	5	6	7
20. In _____'s web site, one can find information related to customer service	1	2	3	4	5	6	7
21. _____'s web site looks attractive	1	2	3	4	5	6	7
22. _____'s web site looks organized	1	2	3	4	5	6	7
23. _____'s web site uses fonts properly	1	2	3	4	5	6	7
24. _____'s web site uses colors properly	1	2	3	4	5	6	7
25. _____'s web site uses multimedia features properly	1	2	3	4	5	6	7

The final 25-item instrument for user-perceived web quality is shown in Table 5. We believe that the instrument, having undergone extensive evaluation and validation, represents significant progress towards the development of a standard instrument for measuring perceived web quality. Moreover, the instrument is precise and easy to use. It can be utilized to evaluate web quality at an aggregate level. The model/instrument could also serve as a starting point for a detailed evaluation of web sites.

4. Conclusion and implications

Past web quality research has focused on general description of some specific aspects of web quality and paid little attention to construct identification and measurement efforts. In this study, we moved beyond descriptive and narrative evidence to empirical evaluation and verification by developing a multi-dimensional scale for measuring user-perceived web quality. The results of the two-phased investigation uncovered four dimensions of perceived web quality (technical adequacy, specific content, content quality, and appearance) and provided evidence for the psychometric properties of the 25-item instrument.

Another contribution is that while past web quality research focuses mostly on the perspectives of web developers and designers, the current study targets the web users. In this era of intense competition and customer responsiveness, the users are major stakeholders and should not be ignored.

The limitations of the study include those customarily associated with instrument-building and survey methods. However, the extensive testing and validation improved the internal validity, and using several groups of subjects improved the external validity and generalizability of the instrument to a larger population. Nevertheless, instruments are always subject to further improvement and we encourage fellow re-searchers to do so.

The web quality model/instrument has practical as well as theoretical and research applications. In terms of practical applications, a validated instrument provides an important tool for assessing the quality of web site. The Internet is hosting hundreds of millions of web sites varying widely in terms of quality. The scales might be used to assess the quality of a given web site. This could be carried out at the overall quality level using the 25-item instrument or at a specific quality dimension level, e.g. using a sub-scale of one of the four dimensions of perceived web quality. This evaluation may provide a fast and early feedback to the firm. If the firm finds itself lacking in any of the dimensions, then it may do a more detailed analysis and take necessary corrective actions. The four dimensions and the 25-items of the instrument may also be used by web site designers in a proactive manner. The dimensions and the items may be considered explicitly in the site design. Additionally, a firm may want to assess the relative importance of the four quality dimensions and the specific items in its own context. While this study provided their relative importance based on its own sample, each firm is unique based on its products, services, customers, business strategies, etc. Such an evaluation would facilitate the design of a quality web site in the first place.

References

[1] A.M. Aladwani, Implications of some of the recent improvement philosophies for the management of the information systems organization, Industrial Management and Data Systems 99 (1), 1999, pp. 33–39.

[2] D. Angelo, S.K. Little, Successful web pages: what are they and do they exist? Information Technology and Libraries 17 (2), 1998, pp. 71–81.

[3] E.R. Babbie, The Practice of Social Research, 6th Edition, Wadsworth Publishing Company, Belmont, CA, 1992.

[4] A. Barron, B. Tompkins, D. Tai, Design Guidelines for the World Wide Web, Last Updated: 13 February 1998, http:// fcit.coedu.usf.edu/workshops/wwwdesg.htm.

[5] H. Bell, N. Tang, The effectiveness of commercial Internet web sites: a user's perspective, Internet Research 8 (3), 1998, pp. 219–228.

[6] D.T. Campbell, D.W. Fiske, Convergent and discriminant validation by the multitrait-multimethod matrix, Psychological Bulletin 56 (2), 1959, pp. 81–105.

[7] G.A. Churchill, A paradigm for developing better measures of marketing constructs, Journal of Marketing Research 16 (1), 1979, pp. 64–73.

[8] L.J. Cronbach, Test validation, in: R.L. Thorndike (Ed.), Educational Measurement, American Council on Education, Washington, DC, 1971, pp. 443–507.

[9] W.L. Currie, The Global Information Society, Wiley, Chichester, UK, 2000.

[10] G. Dempsey, R. Sussman, A hands-on guide for multilingual web sites, Communication World 16 (6), 1999, pp. 45–47.

[11] W.J. Doll, G. Torkzadeh, The measurement of end-user computing satisfaction, MIS Quarterly 12 (2), 1988, pp. 259– 274.
[12] J. Dysart, Interactivity makes for wonderful web sites, Frozen Food Age 47 (3), 1998, pp. 62–63.
[13] D. Gehrke, E. Turban, Determinants of successful web site design: relative importance and recommendations for effectiveness, in: Proceedings of the 32nd HICSS, Hawaii, 1999.
[14] L.A. Gemoets, M.A. Mahmood, Effect of the quality of user documentation on user satisfaction with information systems, Information and Management 18 (1), 1990, pp. 47–54.
[15] D. Hager, C. Kibler, L. Zack, The basics of user-friendly web design, Journal for Quality and Participation 22 (3), 1999, pp. 58–61.
[16] J.F. Hair, R.E. Anderson, R.L. Tatham, W.C. Black, Multi-variate Data Analysis With Readings, 3rd Edition, Macmillan, New York, NY, 1992.
[17] R.P. Heath, Design a killer web site, American Demo-graphics, 1997, pp. 50–55.
[18] J. Ho, Evaluating the World Wide Web: A Global Study of Commercial Sites, Last Updated: 1997, http://www.ascus-c.org/jcmc/vol3/issue1/ho.html.
[19] E.K. Huizingh, The content and design of web sites: an empirical study, Information and Management 37 (3), 2000, pp. 123–134.
[20] R. Kalakota, A.B. Whinston, Electronic Commerce: A Manager's Guide, Addison-Wesley, Reading, MA, 1997.
[21] D. Kaplan, R. Krishnan, R. Padman, J. Peters, Assessing data quality in accounting information systems, Communications of the ACM 41 (2), 1998, pp. 72–78.
[22] W.J. Kettinger, C.C. Lee, Perceived service quality and user satisfaction with the information services function, Decision Sciences 25, 1994, pp. 737–766.
[23] W.R. King, B. Epstein, Assessing information system value, Decision Sciences 14 (1), 1983, pp. 34–45.
[24] G. Levine, 10 steps to building a successful web site, Bobbin 40 (8), 1999, pp. 61–63.
[25] C. Liu, K.P. Arnett, Exploring the factors associated with web site success in the context of electronic commerce, Information and Management 38 (1), 2000, pp. 23–33.
[26] F. Longwell, Effective web sites can facilitate work site sales, National Underwriter 103 (49), 1999, pp. 27.
[27] M. Ma, Agents in e-commerce, Communications of the ACM 42 (3), 1999, pp. 78–80.
[28] A.I. Mechitov, H.M. Moshkovich, S.H. Underwood, Empirical study of university web sites, in: Proceedings of the 30th DSI, New Orleans, Louisiana, 1999.
[29] M.M. Misic, K. Johnson, Benchmarking: a tool for web site evaluation and improvement, Internet Research 9 (5), 1999, pp.383–392.
[30] R. Monti, Better web design boosts sub stales, Folio: The Magazine for Magazine Management 29 (2), 2000, pp. 54.
[31] K.G. Nelson, Global information systems quality: key issues and challenges, Journal of Global Information Management 4, 1996, pp. 4–14.
[32] J. Nunnally, Psychometric Theory, 2nd Edition, McGraw-Hill, New York, NY, 1978.
[33] L. Olsina, D. Godoy, G.J. Lafuente, G. Rossi, Specifying quality characteristics and attributes for web sites, First ICSE Workshop on Web Engineering (WebE-99), Los Angeles, CA, 1999.
[34] S. Poon, P.M.C. Swatman, An exploratory study of small business Internet commerce issues, Information and Management 35 (1), 1999, pp. 9–18.
[35] A. Rai, H. Al-Hindi, The effects of development process modeling and task uncertainty on development quality performance, Information and Management 37 (6), 2000, pp. 335–346.
[36] G. Rose, H. Khoo, D.W. Straub, Current technological impediments to business-to-consumer electronic commerce, Communications of AIS 1 (Article 16), 1999.
[37] J.P. Rowell, G.R. Kessler, R. Berke, Five keys to a web site that's more than just a pretty face, Medical Marketing and Media 34 (5), 1999, pp. 80–84.
[38] T. Sakaguchi, P.C. Palvia, B.D. Janz, Business practices on the world wide web: a comparison of Japanese and US web sites, 2000, submitted for publication.
[39] M. Schacklett, Nine ways to create a retail environment on your web site, Credit Union Magazine (2000) 12-13.
[40] L. Schmeiser, Dress it up, Macworld 16 (11), 1999, pp. 100– 103.

[41] P. Schubert, D. Selz, Web assessment—measuring the effectiveness of electronic commerce sites going beyond traditional marketing paradigms, in: Proceedings of the 32nd HICSS, Hawaii, 1999.
[42] D.W. Straub, Validating instruments in MIS research, MIS Quarterly 13 (2), 1989, pp. 147–169.
[43] P. Tarasewich, Issues in web site design and usability, in: Proceedings of the 30th DSI, New Orleans, Louisiana, 1999.
[44] T.S.H. Teo, M. Tan, An empirical study of adopters and non-adopters of the Internet in Singapore, Information and Management 34 (6), 1998, pp. 339–345.
[45] H.A. Wan, Opportunities to enhance a commercial web site, Information and Management 38 (1), 2000, pp. 15–21.
[46] J. Webster, J.J. Martocchio, Microcomputer playfulness: development of a measure with workplace implications, MIS Quarterly 16 (2), 1992, pp. 201–226.
[47] J.C. Westland, T.H.K. Clark, Global Electronic Commerce: Theory and Case Studies, MIT Press, Cambridge, MA, 1999.
[48] R. Wilson, Seven Debilitating Diseases of Business Web Sites, Last Updated: 1 February 1998, http://www.wilson-web.com/articles/7diseases.htm.