Available online at www.sciencedirect.com
SCIENCE DIRECT®
Information & Management 41 (2004) 543–560
INFORMATION & MANAGEMENT
www.elsevier.com/locate/dsw

The conceptualization and empirical validation of web site user satisfaction

Steve Muylle[a,*], Rudy Moenaert[b], Marc Despontin[c]

[a]*Vlerick Leuven Gent Management School, Reep 1, 9000 Gent, Belgium*
[b]*Tias Business School, University of Tilburg, Tilburg, The Netherlands*
[c]*Vrije Universiteit Brussel, Brussel, Belgium*

Received 5 September 2001; received in revised form 24 November 2002; accepted 23 June 2003

Abstract

This article addresses the concern for effective web site design by means of the conceptualization and empirical validation of a *web site user satisfaction* construct. Based on IS success theory, hypermedia design theory, a qualitative exploratory pilot study, and a quantitative online critical incident technique, we introduce and define the construct of web site user satisfaction, explore its dimensionality, provide empirical validation of the construct and its underlying dimensionality, develop a standardized instrument with desirable psychometric properties for measuring WUS, and explore the measure's theoretical and practical application.

Keywords: User satisfaction; Web design; Scale development

1. Introduction

> Learning to program has no more to do with designing interactive software than learning to touch-type has to do with writing poetry. (Ted Nelson [22])

Perhaps no other factor has changed the face of IS as significantly as the swift emergence and adoption of the World Wide Web (WWW). At the very heart of the migration from traditional to online commercial environments are the interactions between individual users and companies' web sites. Consumers and business people alike are increasingly turning to the WWW to search for information on goods and services, e.g., in 1999, 55 million out of the estimated 92 million Internet users in North America searched for and compared prices and features of products and services on the Internet, while an estimated 28 million users actually bought online [19].

Given the role a web site plays as the interface between a user and a company, effective web site design must be seen as a prerequisite for successfully instituting an Internet commerce initiative. Indeed, effective interface design is a research priority in electronic commerce by the National Science Foundation: ''Graphical User Interface implemented in the WWW is recognized as one of the deciding factors that make electronic commerce possible'' [66].

* Corresponding author. Tel.: +32-9-210-98-77; fax: +32-9-210-98-75.
E-mail address: steve.muylle@vlerick.be (S. Muylle).

0378-7206/$ – see front matter
doi:10.1016/S0378-7206(03)00089-2

This relates directly to users' inability to find and access desired information both across the web and within commercial web sites [18]. Even though industrial literature and information on the practical implementation of usability studies in web site design (e.g., Nielsen's work at http://www.useit.com) are available to assist organizations in the effective design of their web sites, science-based theory and guidelines are, as of yet, scarce [24,48]. Yet, a valid and empirically reliable measure that captures users' evaluation of the organization's web site is still needed: researchers could then achieve greater precision in formulating and testing research hypotheses, while managers would be able to assess the effectiveness, strengths, and weaknesses of a web site in a reliable and valid way.

Therefore, the objective of our research was to contribute to management scholars' and practitioners' understanding of effective web site design by:

1. introducing and defining a *web site user satisfaction* construct;
2. identifying the underlying dimensions of the construct;
3. operationalizing a measure to assess an individual's web site user satisfaction;
4. empirically validating the construct; and
5. exploring the theoretical and managerial implications of the findings.

It should be noted that we focus on *goal-directed behavior* on the WWW given that "the optimal design of a computer-mediated environment (CME) site (e.g., a web site) differs according to whether the behavior is goal-directed or experiential" and that "goal-directed behavior is relevant for task-specific use of a CME, such as prepurchase deliberation." As such, our article concentrates on the needs of users searching a web site for a specific company, product, or service. It excludes the web site evaluation of users merely surfing or browsing the web.

2. The web site user satisfaction construct

In order to introduce a theoretically grounded construct of web site user satisfaction, two major streams of research were perused. Reviews of literature on:

- hypermedia design, to describe the user environment;
- IS success, to investigate the user satisfaction concept.

2.1. Hypermedia design

The literature on hypermedia design was appropriate, given that the WWW is generally considered to be the first and current commercialized implementation of a hypermedia CME [46]. One of the key features distinguishing such an environment from traditional commercial ones is the hands-on direct interaction with computer-mediated sites utilizing hypermedia [67,78,80].

While most people have been introduced to hypermedia through the web or Apple's Hypercard System, its conceptual foundation can be traced back to Bush who proposed the idea of a personal "memory extension" system, the "memex" in 1945 [15]. This system would allow the user to store and consult information, based on trails of association by which the human mind is supposedly working: "It [the human mind] operates by association. With one item in its grasp, it snaps instantly to the next that is suggested by the association of thoughts, in accordance with some intricate web of trails carried by the cells of the brain." Nelson, a student of Bush, coined the terms "hypertext" and "hypermedia" in 1965 [20]. He described hypertext as "nonsequential writing-text that branches and allows choices to the reader," echoing the trails of association paradigm envisioned by Bush [9].

Hypermedia extends the definition of hypertext to include media types other than text [12,86]. In such a node-and-link model, nodes are the holders of multimedia content and links are the means by which the content is structured. Therefore, two key aspects of hypermedia design concern content, i.e., the information in the application, and its structure. A third key design issue is presentation [42,58,65,76], including appearance of content nodes and hyperlinks, location of navigational aids, and button layouts [51].

2.2. The IS literature on MIS success

Although MIS are different from web sites, literature on MIS success contains concepts that can be readily applied to research on user satisfaction with

web sites [26]. In the literature six dimensions of MIS success have been identified: (1) system quality, (2) information quality, (3) system use, (4) user satisfaction, (5) individual impact, and (6) organizational impact [27]. Because of the difficulty in measuring the impact of a MIS on individual decision making processes and resulting organizational performance, user satisfaction and system use have been the most widely used constructs. Moreover, since system use is not always an appropriate measure of system success (e.g., in the case of mandatory use) [56], and user satisfaction stresses the importance of *perceived* system and information quality rather than technical quality [53], user satisfaction has become the prevailing construct [61]. As such, the underlying assumption in adopting user satisfaction as a measure of MIS success is that there is a positive relationship between this construct and individual and organizational performance [83]. These key relationships have been tested and supportive evidence found [43,49,50,74].

Various measures for evaluating user satisfaction with an organization's MIS have been developed [4,25,30–32,41,55,63,70,82,91]. Although these measures use different definitions of user satisfaction, they all share the idea that the user provides some form of evaluative response. Furthermore, based on McGuire [60] and Ajzen [1], Melone suggests that user satisfaction is embodied in user attitudes. For instance, this notion of user attitude can be found in the definition of Doll and Torkzadeh who define end-user computing satisfaction as ''the affective attitude towards a specific computer application by someone who interacts with the application directly.'' The definition also refers to an end-user computing environment instead of the more traditional data processing environment.

2.3. Definition of web site user satisfaction

Using the theoretical developments in the IS success literature, *web site user satisfaction* may be considered a measure of the utility of an organization's web site in the end-user's decision making process and may be defined as *the attitude toward the web site by a hands-on user of the organization's web site*. In addition to emphasizing the need for direct interaction, the definition also contains the notion of user attitude towards an organization's web site.

The structure of the web site user satisfaction construct can be viewed as a set of items relating to the user's satisfaction with the hypermedia content, structure, and presentation.

3. Exploratory pilot study

Electronic commerce initiatives are presenting new challenges for IS managers. The hypermedia design and MIS success literature may not be sufficient to monitor all factors affecting web site user satisfaction. Therefore, an *exploratory pilot study* was conducted to determine the items important in the user–web site interface.

3.1. Data collection

The pilot study was conducted in collaboration with a leading international courier company that handles more than 260,000 shipments on an average day and serves over 675,000 destinations worldwide. Professional users of a Western-European country site of this company were asked to visit a limited set of courier companies' web sites and, during a 1-h time period, using computer facilities provided by the research team analyze its usefulness, i.e., whether it satisfied their needs. The following instruction was given:

> As a professional user of courier companies you want to use the Internet. You decide to visit the web sites of the leading companies in this sector and to evaluate each web site with respect to your needs.

Different users visited web sites in a variety of countries (mostly Belgium, The Netherlands, the UK, and the US) pursuing different search objectives. For instance, a secretary was looking for information (price, transit time, procedure) on sending a shipment to a particular destination, while a distribution manager was finding out about the strategies, strengths, and weaknesses of the courier companies. Thus, a distribution manager typically compared the web sites of a set of courier service firms both globally and in-country.

A sample consisting of 14 professional users of courier services was used. The respondents were sampled for heterogeneity in order to obtain a deep insight into the factors that affected their satisfaction

with the web site [21]. The sample size was in line with the principles of case-based research methodology, involving analytic generalization (i.e., replication of conceptual understanding) as opposed to statistical generalization of results to a larger universe in quantitative research [89]. Furthermore, ''with high complexity, a study with more than 15 cases or so can become unwieldy'' [62].

In order to reveal the evaluative thoughts of the respondents, two distinct forms of protocol analysis were used [37,38] and the navigational behavior of each of the respondents was filmed. This was done unobtrusively (so that it would not influence task performance [87]). Data was collected as follows:

- *Method 1*. Verbal reporting by the respondent while visiting the web sites.
- *Method 2*. Later verbal reporting by the respondent following task completion, while the respondent watched a replay of the video record of his/her navigational behavior.

In order to counterbalance each method's potential threats to protocol validity half of the respondents were assigned to Method 1 while the others were assigned to Method 2. During the research session, the researcher was careful not to interfere; no prompt was given unless the respondent lapsed into silence, when a nondirective comment was made (e.g., ''keep on talking'').

The respondents assigned to *Method 1* were asked to report their thoughts in a general way, i.e., ''Report everything that goes through your head while visiting the web sites, think aloud.'' A short training session was held. Immediately following the research session questions about satisfaction with the web sites and intention to revisit them were asked [88]. Protocol study using data collection Method 1 took approximately 1.5 h.

The respondents assigned to *Method 2* were free to complete the task without having to think aloud. Immediately following task completion and while watching the replay the respondent was asked to report ''everything that went through your head while visiting the web sites.'' Because the respondent was asked to retrospectively report his/her evaluative thoughts, the duration of this type of protocol study was longer: approximately 2.5 h.

Table 1
Threats to protocol validity

Sources of invalidity	Method 1 (concurrent reporting)	Method 2 (retrospective reporting)
Reactivity		
Disruption	Moderate threat	Minor threat
Response time	Moderate threat	Minor threat
Nonveridicality		
Omission	Minor threat	Moderate threat
Commission	Minor threat	Moderate threat

3.2. Validity of the exploratory study

The data collection process was guided by concerns for validity. Russo et al. [73] point to ''reactivity'' and ''nonveridicality'' as major problems. Reactivity pertains to the threat that the thinking aloud may change the way the task is done. As such, the process may be disrupted and/or response time may be prolonged. Nonveridicality includes errors of omission (the verbal protocol may be incomplete because some thoughts are omitted) and errors of commission (the respondent may report mental activities that did not occur). With respect to the exploratory pilot study, the threats to protocol validity are summarized in Table 1.

3.2.1. Reactivity threat

Concurrent protocol analysis (Method 1) may cause disruption of natural processing activities and also prolong response time. However, when respondents internally pace the message content and perform the task at hand in a distraction-free (minimal interference), low time–pressure situation, reactivity is reduced (moderate threat) [36]. Moreover, given the waiting time following a course of action, response time is not believed to have a substantial impact on the performance of the task (moderate threat). Because retrospective protocol analysis (Method 2) does not require the respondent to think aloud while performing the primary task, the reactivity threat is negligible.

3.2.2. Nonveridicality threat

In general, the validity of retrospective protocols suffers from forgetting (omission) and fabrication (commission) of mental events [68]. Therefore, in case of retrospective protocol analysis, verbal reporting immediately following task completion, while

many thoughts are still in active memory, is recommended. Moreover, the respondent's recall error may be minimized by showing the respondent the replay of his/her search behavior while he or she was performing the task at hand (moderate nonveridicality threat) [71]. Conversely, in case of concurrent verbal reporting, errors of omission and errors of commission only pose a minor threat to protocol validity, because the respondent thinks aloud while visiting the web sites.

3.3. Dimensions underlying WUS

Transcripts were made of each interview. This resulted in approximately 300 pages of interview material, which was analyzed through thematic content analysis [47] using the open coding method defined by Strauss and Corbin [81]. The objective of the analysis was to elicit relevant dimensions pertaining to web site user satisfaction from the respondents. In addition, 14 h of video material were collected. The video material was used to support the open coding of the transcripts. In order to reconcile a holistic approach with theoretical parsimony, the findings of this exploratory research study were compared with studies on MIS success.

Using this process, 11 dimensions underlying web site user satisfaction were derived. The first five dimensions can be viewed as directly corresponding to well-established dimensions of user satisfaction and information utility constructs. Indeed, for each of these dimensions, a corresponding dimension is found in the IS literature on user satisfaction:

1. *Information relevancy*. The degree to which the hands-on user perceives the web site information content to meet the needs.
2. *Information accuracy*. The extent to which the information content is perceived to be precise.
3. *Information comprehensibility*. The extent to which the user can easily decode and understand the information in the web site.
4. *Information comprehensiveness*. The perceived completeness of the information in the web site.
5. *Ease of use*. The extent to which the web site is perceived to be user friendly.

In addition to these well-established dimensions, six dimensions were discovered from our exploratory work. These specifically pertain to user–web site interaction, and have no direct counterpart in the literature:

6. *Entry guidance*. The user's perception of the quality of the start page (the homepage) in guiding him or her to the appropriate spot in the web site.
7. *Web site structure*. The perceived structural cohesion of the site, i.e., the user's perception of the manner in which the different parts of the web site are linked up.
8. *Hyperlink connotation*. The extent to which the user can easily decode and correctly interpret the hyperlinks in the site.
9. *Web site speed*. The degree to which the user perceives the site to be slow or fast.
10. *Layout*. The perceived look of the web site in terms of color scheme, backgrounds, fonts, and images.
11. *Language customization*. The degree to which the national language is tailored to the user.

3.4. Dimensional structure

From a user-oriented perspective, the design of the content, its structure, and presentation of a hypermedia application is likely to affect the user's satisfaction [77]. As such, the hypermedia design framework can be related to the items of the web site user satisfaction construct. For each of the design issues, a corresponding user-oriented dimension of WUS can be hypothesized, as depicted in Fig. 1. The proposed structure differs from the structure derived from perusing the hypermedia design and the IS success literature. Indeed, the structure proposed in the hypermedia design literature was found to be too coarse-grained. Similarly, the perceived system and information quality dimensions were found to contain some of the web environment addressed here. As such, the structure not only builds upon dimensions well-established in the literature, but also adds novel dimensions that emerged from our exploratory work.

Information. This corresponds to the user's satisfaction with the web site nodes in which the information is contained. Furthermore, in line with Zmud's study of the dimensionality of information, the information relevancy, accuracy, comprehensibility, and comprehensiveness dimensions can be seen as reflecting user satisfaction.

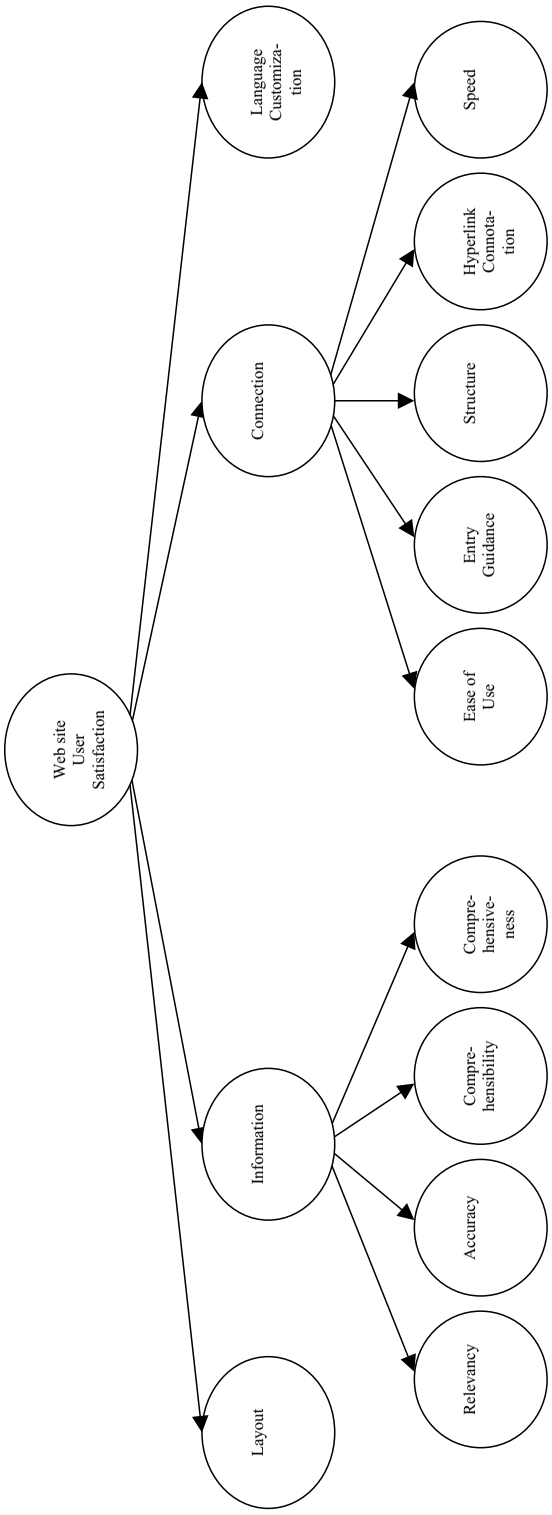

Fig. 1. Dimensional structure of web site user satisfaction.

Connection. This concerns a user's satisfaction with the connection, i.e., the way the content is organized and linked, i.e., how information is divided, how they are linked, and which semantic relationships are used. Based on these design issues, factors such as ease of use, entry guidance, web site structure, hyperlink connotation, and web site speed are assumed to reflect a user's satisfaction with the web site connection. Web site speed is included to underline the fact that it must be high enough and that the bandwidth (i.e., "the amount of network capacity required to support a connection" [54]) is sufficient because many web pages must often be traversed in order to connect to the desired content.

Layout. This deals with the presentation of hypermedia: the layout dimension identified in the exploratory pilot study. Hypermedia accommodates many of the visual qualities of traditional print (e.g., colors, textures, graphic elements) but goes far beyond this by possibly adding other layout features such as animations, moving three-dimensional typography, and hybrid image-text juxtapositions. Since all these elements may facilitate or hamper the goal-directed search behavior of web site users, the hypermedia should be seen in a way that helps users pursue their search objectives [45]. In order to further investigate the absence of lower order layout dimensions the Internet marketing literature was surveyed. While support for the layout dimension was found in web uses and gratifications [33,34], web advertising [7,16,29,44] and web site quality [90], none of the studies identified lower order layout dimensions.

Language. Given the global span of the WWW, the national language(s) in which the information is offered may influence a user's satisfaction level with a commercial site [72]. If the user is addressed in a language that is not his or her mother tongue, satisfaction may be adversely affected. This is particularly significant in the multilingual Western-European setting where the exploratory pilot study was conducted. Again, no lower order dimensions are hypothesized. Also, our review of the Internet marketing literature did not provide any indication of lower order language customization dimensions.

In summary, the following propositions were formulated:

Proposition 1. *Web site user satisfaction relates positively to the user's perceived utility of the information content, the connection, the layout, and the language customization of the web site.*

Proposition 2. *User satisfaction with the information content reflects the user's perception of the relevancy, accuracy, comprehensibility, and comprehensiveness of the information content in the web site.*

Proposition 3. *User satisfaction with the connection embraces the user's perception of the ease of use, entry guidance, structure, hyperlink connotation, and speed of the web site.*

4. Online critical incident technique

In order to validate the propositions empirically and to develop a standardized measure of WUS, data was collected using an online critical incident technique [39].

4.1. Data collection

The data was gathered in a stepwise manner. First, a personalized invitation e-mail was sent to a sample of Internet users. This explained the general goal of our study and solicited the respondent's participation. In order to increase the probability of a favorable response, a lottery with five $50 cash prizes was promised and a feedback report was offered. The invitation e-mail also listed eight subject categories (company information, flights, courier services, computers, travel, mortgages, cars, CD-recorders) from which the respondent was asked to choose. These categories were selected to support the generalizability of the findings by representing a broad spectrum of information search contexts.

Upon a positive reply to the invitation, a second e-mail containing specific instructions (as well as a questionnaire) was sent to the respondent. This provided a search objective for a particular web site to which the Uniform Resource Locator was given. The search objective was chosen to fit the field of interest indicated by the respondent. For each of the eight subject categories, various search objectives and web sites were used in order to further enhance the generalizability of the study.

After having carried out the task, the respondent was asked to complete an e-mail questionnaire. The first part

included a set of 60 items, formatted as 5-point Likert scales, tapping the proposed dimensions of the WUS construct. This pool of 60 items was generated on the basis of the results of the pilot study and MIS user satisfaction measures. Approximately half of the items were negatively stated in order to reduce "yea-" saying tendencies [28]. The order of the 60 items was randomized, i.e., the items were not presented in the order of the proposed dimensions. This was done to improve reliability and validity [2]. The second part of the questionnaire contained a 7-point perceived overall satisfaction item, as well as questions about sociodemographic and webographic characteristics (e.g., amount of use of the web per week and whether the respondent had visited the web site before).

The e-mail was sent to a sample of 5922 valid e-mail addresses, randomly drawn from a large e-mail list provided by a company operating a major Western-European portal web site. 1248 Internet users agreed to participate in the study. The e-mail containing the tailored instruction and questionnaire was then sent to these 1248. After 2.5 weeks a reminder e-mail was sent to those people who had not yet responded. A total of 837 respondents provided usable answers to the questionnaire; this is a 14.1% return of the original mailing or 67.1% of the people that agreed to cooperate. People who originally chose not to reply or ignored the second and subsequent e-mails were not contacted again in order to avoid *e-mail spamming* [17].

The sample of 837 respondents consisted of 755 male (90.2%) respondents. The average age was 34.9 years with a standard deviation of 10.6 years. On average, the respondents said that they surf 8 h per week on the Internet. A standard deviation of 8.2 h was observed. 83.4% of the respondents had not previously visited their assigned web site. On average, the respondent stayed 16.7 min in the web site, with a standard deviation of 11.9 min. It is important to note that the e-mail list, provided by the portal web site, predominantly consisted of e-mail addresses of business professionals. Furthermore, the coding of the answers to the open question on a respondent's profession corroborated this. Indeed, the typical respondent was found to be a Western-European business professional (e.g., IT consultant, project manager, general manager, in Belgium, The Netherlands, the UK, or France) speaking multiple languages (typically English, Dutch and/or French).

4.2. Data analysis

4.2.1. Data screening

Listwise deletion of missing values on the 60 items resulted in a sample of 719 complete cases. Bollen's [10] model-free multidimensional outlier detection procedure identified four multivariate outliers. However, since no theoretical reason suggested the exclusion of these observations, they were retained. Sensitivity analysis, both in terms of missing data and multivariate outliers, showed no cause for concern (the research results were unaffected when the four multivariate outliers were excluded from the analysis). Furthermore, analysis based on the covariance matrix formed by pairwise deletion of missing data on the 837 usable answers yielded virtually the same parameter estimates and fit as the analyses based on the 719 complete cases [23].

4.2.2. Item purification

In order to purify the pool of 60 items, corrected item-to-total correlations and Cronbach alphas were investigated per dimension. Items with low corrected item-to-total correlations were dropped, while three or four items per dimension were retained for the purposes of confirmatory factor analysis (CFA) [6]. In total, 34 items were retained. The retained items are listed in Appendix A. Cronbach alphas for each of the proposed dimensions of WUS ranged between 0.74 and 0.89 (see Table 2) and clearly exceed the 0.70 cutoff recommended by Nunnally [69] for basic research.

4.2.3. Confirmatory factor analysis

CFA involves the specification, identification, estimation, and evaluation of one or more theoretical models of factor structure, each of which accounts for covariances among a set of observed variables by means of a set of a priori specified latent concepts. Its key feature is that if hypotheses about plausible model structure exist, CFA allows the system to explicitly test these proposed factor structures against sample data. Furthermore, CFA is particularly well-suited to accommodate higher order factor structures [57]. This study used the PROC CALIS procedure in SAS for model description and estimation [75].

4.2.3.1. Specification. The CFA model was specified in accordance with the theoretical structure of Fig. 1.

Table 2
Standardized parameter and reliability estimates ($N = 719$)

Observed variables			First-order factors				Second-order factors		
Item	Factor loading	Reliability	Factor	Structural coefficient	R^2	Reliability	Factor	Structural coefficient	R^2
IR1	0.76[a]	0.58	IR	0.99[a]	0.99	0.82	I	0.73 (13.0)	0.53
IR2	0.81 (22.2)	0.65				0.83	C	0.98 (11.6)	0.96
IR3	0.78 (21.5)	0.61				0.62			
IA1	0.71[a]	0.51	IA	0.88 (16.6)	0.78	0.74			
IA2	0.76 (17.6)	0.57				0.74			
IA3	0.61 (14.6)	0.37				0.49			
ICy1	0.73[a]	0.54	ICy	0.76 (15.6)	0.58	0.83			
ICy2	0.72 (17.9)	0.52				0.83			
ICy3	0.79 (19.4)	0.62				0.55			
ICy4	0.73 (18.2)	0.54							
ICs1	0.87[a]	0.76	ICs	0.86 (19.9)	0.73	0.86			
ICs2	0.84 (27.5)	0.71				0.86			
ICs3	0.76 (23.5)	0.57				0.68			
E1	0.75[a]	0.57	E	0.97 (17.4)	0.94	0.83			
E2	0.84 (22.9)	0.70				0.83			
E3	0.79 (21.4)	0.62				0.63			
EG1	0.90[a]	0.81	EG	0.72[a]	0.53	0.89			
EG2	0.84 (28.7)	0.70				0.89			
EG3	0.82 (27.9)	0.68				0.73			
S1	0.85[a]	0.73	S	0.96 (19.1)	0.92	0.88			
S2	0.83 (27.7)	0.69				0.88			
S3	0.85 (28.8)	0.73				0.71			
H1	0.87[a]	0.76	H	0.83 (17.4)	0.68	0.85			
H2	0.76 (23.2)	0.58				0.86			
H3	0.81 (25.2)	0.65				0.67			
Sp1	0.85[a]	0.72	Sp	0.64 (13.7)	0.41	0.80			
Sp2	0.66 (17.2)	0.43				0.81			
Sp3	0.77 (19.8)	0.60				0.58			
L1	0.87[a]	0.76	L	0.65[a]	0.42	0.85			
L2	0.78 (22.3)	0.60				0.85			
L3	0.77 (22.1)	0.60				0.65			
Lg1	0.84[a]	0.71	Lg	0.24 (5.6)	0.06	0.89			
Lg2	0.89 (27.8)	0.80				0.89			
Lg3	0.84 (26.3)	0.70				0.73			

[a] Indicates a parameter fixed at 1 to set the scale of the construct; *t*-values are indicated in parentheses; individual item reliabilities are reported as estimates of the reliability for the observed variables, while Cronbach alpha, composite reliability, and average variance extracted are reported, respectively, as estimates of the reliability for the first-order factors.

This proposes two factors at the second-order level, information and connection, that combine with the layout and language factors to form the WUS factor at the third level. The possibility of a higher order factor structure in which higher order constructs influence lower order constructs has long been recognized [84]. However, it was not until the advent of structural equation modeling that higher order factor structures could be explicitly tested. Furthermore, the model was specified to have simple structure and uncorrelated

measurement errors. Three or four items per dimensional construct were specified, since, in general, at least three indicators per latent construct are needed for purposes of identification.

4.2.3.2. Model estimation and identification. Maximum likelihood (ML) was used for estimation. The assumption of multivariate normality was supported, given the approximately univariate normal distributions of the retained items (univariate skewness and kurtosis approximately ranged between −1 and +1) and the index of relative multivariate kurtosis of 1.18 [13,64]. Furthermore, the sample size ($N = 719$) was believed to be sufficiently large to justify the use of the ML estimation technique, based on asymptotic distribution theory. The proposed model is theoretically identified, since it is recursive and has at least three indicators per construct. No convergence problems were encountered. Also, no improper solutions (e.g., negative error variances) emerged.

4.2.4. Model evaluation

4.2.4.1. Assessment of overall model fit. Both stand-alone and incremental indices were used to assess overall model fit [11]. The most popular stand-alone index is the chi-square test, which tests the hypothesis that the implied covariance matrix of the model ''exactly'' reproduces the sample covariance matrix. A closely related stand-alone index is the chi-square estimator divided by its degrees of freedom. As to the latter, values ranging from below 2–5 are generally considered evidence of reasonable fit. However, it is important to note that the chi-square statistic is often of limited usefulness because of its overly rigid nature and excessive power.

Given the shortcomings of the chi-square statistic and alternative fit indices based on it (e.g., the goodness of fit index, GFI), the use of fit indices that express model-data fit in terms of noncentrality is recommended [59]. As such, Steiger's [79] root mean squared error of approximation (RMSEA), Bentler's [8] normed comparative fit index (CFI), and the Tucker and Lewis [85] nonnormed fit index (TLI) were considered more appropriate measures. These fit indices acknowledge that specified models can only be approximately true and appear to overcome the problem of dependence on sample size. Many researchers interpret RMSEA values below 0.05 as representing close fit, while values up to 0.08 indicate reasonable fit [14]. For both the CFI and TLI, values of 0.9 or higher are generally considered evidence of good fit.

The postulated model shows good overall model-data fit. The ratio of chi-square to the degrees of freedom was 3.9, while RMSEA, CFI, and TLI, respectively, amounted to 0.06, 0.91, and 0.90. Even though the empirical evidence strongly supports our model, it is of interest to examine plausible alternative factor models to further validate the postulated model. Toward that end, four alternative models of first and higher order factor structure were proposed, showing increasing structure and parsimony. These were directly adapted from Doll et al. who used CFA to investigate the factor structure of the end-user computing satisfaction and the user information satisfaction measures.

Specification of alternative models.

- *Model 1* is a first-order factor model; it stated that one first-order factor, web site user satisfaction, accounts for all the common variance among the 34 scale items. This model was implicit in the general practice of adding the item scores to obtain a total satisfaction score. The dimensionality of the web site user satisfaction construct was not taken into account.
- *Model 2* explicitly assumed 11 orthogonal first-order factors (the 11 dimensions of WUS) without considering WUS as a single factor. Thus, instead of one satisfaction construct, 11 uncorrelated factors were assumed to explain most of the variance. Testing this model also gave us the opportunity to compare the increase in fit if the factors are allowed to correlate (Model 3).
- *Model 3* suggests that the 11 first-order factors are correlated. It rested on the assumption that the scale items have substantial common variance.
- *Model 4* accounts for the common variance in Model 3 by introducing the second-order factor, WUS. This hypothesizes that the first-order correlations were due to the second-order factor.

Model specification, estimation, and identification for each of the alternative models was in line with the procedures used in describing and estimating

the postulated model. A baseline or null model, i.e., a model suggesting that no latent factors underlie the observed variables and that the correlations between observed variables are zero in the population, was included.

4.2.4.2. Model evaluation and comparison. Conforming to our expectations, the null model provided a poor fit to the data. Although Models 1 and 2 show a substantial improvement over the baseline model, neither of them indicated a reasonable fit with the sample data. Model 3 demonstrated close model-data fit, as evidenced by the values for RMSEA (0.05), CFI (0.95), and TLI (0.94), which all fall within the ranges of good fit. Furthermore, based on the construction of confidence intervals around the estimated factor correlations, all factor correlations were found to be significantly different from unity. This supports discriminant validity of the 11 proposed dimensions and thus recognizes the distinct nature of each of the proposed dimensions.

Model 4 provided reasonable model-data fit as indicated by the RMSEA and the chi-square ratio. However, in terms of CFI and TLI, its fit was borderline.

Except for Model 3, all alternative models did not perform as well as the original model. However, it should be noted that the goodness of fit of a higher order factor model can never be better than that of the corresponding first-order model, since the higher order factors are merely trying to explain the covariances among the first-order factors in a more restricted, parsimonious way. In this sense, the first-order model provided a target or optimum fit for the higher order factor model. Therefore, in order to examine the existence of higher order factor structures, the postulated model could be compared to the first-order model by means of the target coefficient. This index reflected the extent to which the higher order factor model accounted for the covariances among the first-order factors and could be interpreted as the percent of variation in the first-order factors that depended on the higher order factor structure.

Using Model 3 as the target model, the target coefficient for the postulated model was 0.66. Thus, as the postulated model accounts for 66% of the variation in the 11 first-order factors in Model 3 as opposed to 58% for Model 4 and shows good model-data fit, it can be seen as a good representation of the underlying structure of the web site user satisfaction construct. Furthermore, because of its solid grounding, the postulated model is of greater theoretical interest, structure, and parsimony than Model 3. In summary, based on the extensive evidence supporting the model postulated in Section 3.4, the latter is entertained as the most plausible account of the data.

4.2.4.3. Assessment of the measurement model. In order to ascertain the quality of construct measurement, the sign and size of the standardized factor loadings and their level of statistical significance were investigated. These standardized factor loadings of observed variables (items) on latent variables (factors) can be interpreted as estimates of the validity of the observed variables. Furthermore, individual item reliabilities (squared correlation between a construct and one of its indicators), composite reliability (squared correlation between a construct and an unweighted composite of its indicators), and average variance extracted (proportion of the total variance in all indicators of a construct accounted for by the construct) were computed to assess reliability [40]. As can be seen in Table 2 the component measures were quite good. All standardized factor loadings were positive as expected, statistically significant, and substantial (>0.6), indicating evidence of good construct validity. All individual item reliabilities were greater than 0.4, except for the third information accuracy item (IA3) whose individual item reliability was marginal (0.37). More important than adequate measurement of a construct by a given item, is the adequate measurement of a construct by all of its indicators jointly. Composite reliabilities and average variance extracted serve this purpose very well. Composite reliabilities for the 11 factors ranged between 0.74 and 0.89, values that are considered adequate, while values for average variance extracted generally exceeded 0.5, which is deemed desirable [3]. In summary, the various indices of component fit used in this study all supported the adequacy of construct measurement.

4.2.4.4. Assessment of the latent variable model. The sign, size, and statistical significance of the standardized structural coefficients were examined to assess

this. These coefficients can be interpreted as indicators of validity of the latent factors as components of the WUS construct. In addition, the proportion of variance (R^2) in the latent variables accounted for by the higher order construct influencing them can be used to estimate the reliability of the latent factors. It can be seen in Table 2 that all the standardized structural coefficients were positive and statistically significant. Also, apart from language customization, all coefficients were substantial (>0.6) and at least 40% of the variance for each latent factor was accounted for by the related higher order construct. In summary, all latent factors appeared to be reliable and valid components of the WUS construct, except for the language customization dimension. A possible reason may have been that, as the WWW is reinforcing the trend toward English as the language of commerce, respondents whose mother tongue is not English still felt comfortable searching a web site in English. Indeed, as our sample mainly consisted of multilingual business people used to speak English in their professional environment this may be a valid explanation. On the other hand, respondents may have chosen not to participate unless the content was provided in their native language.

5. Construction of the WUS measure

The confirmatory factor analysis demonstrated adequate reliability and validity for the postulated model (except for the language customization dimension), therefore this model can be considered appropriate as a basis for constructing a standardized WUS measure: it can be measured by computing the mean for each of the dimensions (except language) and summing these means. Simple unit weighting, rather than weighting in function of the path coefficients, was used for forming composite scores [35].

The means, standard deviations, minimum and maximum values, and decile scores for the measure of WUS, as well as for the measures of its second-order dimensions are shown in Table 3. Cronbach alphas for the WUS measure, the information measure, and the connection measure, respectively, were 0.95, 0.92, and 0.93.

In order to impute meaning to a specific score of the WUS instrument, the raw score should be compared to the total distribution of scores, which is either summarized by its mean and standard deviation or by percentile rank of any particular score. Toward the latter end, Fig. 2 depicts the cumulative relative

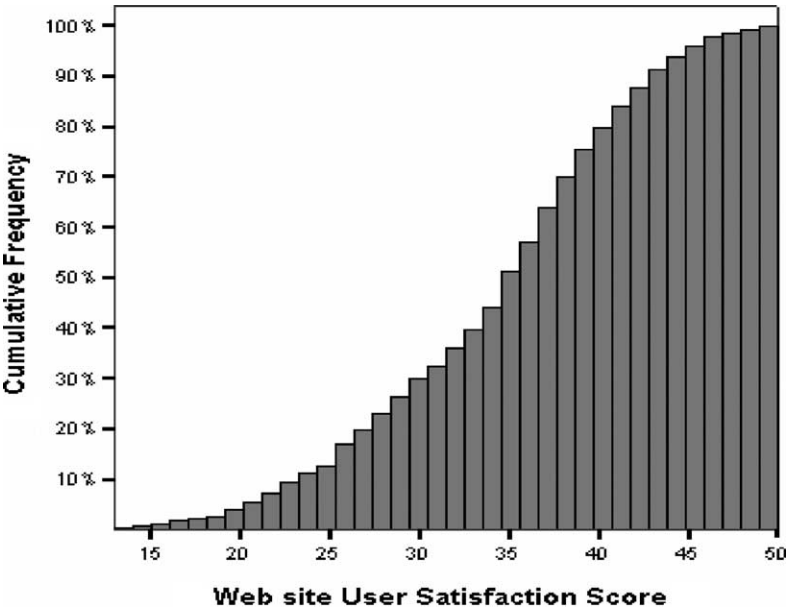

Fig. 2. Cumulative relative frequency of the WUS measure.

Table 3
Descriptive statistics WUS measure

	$\bar{x}$	s	Minimum	P10	P20	P30	P40	P50	P60	P70	P80	P90	Maximum
WUS	34.4	7.4	13	23.9	27.4	30.4	33.6	35.5	37.1	38.8	40.8	43.3	50
Information	13.9	3.2	4.5	9.4	10.9	12.2	13.4	14.3	15.3	16.0	16.6	17.7	20
Connection	17.0	4.3	5.3	10.7	13.0	14.7	16.3	17.7	19.0	19.7	21.0	22.3	25
Layout	3.5	1.0	1.0	2.0	2.7	3.0	3.3	3.7	4.0	4.0	4.3	4.7	5

Computer Manufacturer Web site (N=29)	**Mean**	**Standard deviation**
Web site User Satisfaction	36.6	6.0
Information	15.2	2.4
Information Relevancy	4.1	0.6
Information Accuracy	3.8	0.7
Information Comprehensibility	3.7	0.8
Information Comprehensiveness	3.6	0.8
Connection	18.2	3.5
Ease of Use	3.8	0.7
Entry Guidance	3.9	1.0
Structure	3.5	0.8
Hyperlink Connotation	3.5	0.9
Speed	3.6	0.9
Layout	3.2	0.8
Web site User Satisfaction		
Competition (N=90)	33.2	7.6
Overall (N=719)	34.4	7.4

Fig. 3. Example of WUS scorecard.

frequencies of the WUS scores obtained in our study, while Table 3 includes the decile scores for the WUS measure and the measures of its underlying second-order dimensions.

A practical application consists of developing a WUS scorecard, which shows the score on each of the dimensions, as well as the composite second-order and WUS total score for a particular web site. Fig. 3 provides an instance of such a scorecard for a well-known computer manufacturer's web site. Its WUS total score of 36.6 is positioned between the 50th and 60th percentile in Fig. 2 and clearly indicates there is much room for improvement. When looking at the detailed scores, information comprehensibility, ease of use, structure, hyperlink connotation, and layout seem to be unsatisfactory (below respective median scores). Clearly, as firms have respondents rate competitors' web sites, industry peer benchmarking can be performed.

6. Discussion

6.1. Theoretical contribution

In the field of IS research a well-defined outcome measure is essential. Without a valid measure evaluation remains purely speculative. Notwithstanding the significant advances, existing user measures are being

challenged by new technological developments, especially as the advent of the web seems to redefine the interface.

Our measure of success is a first step toward building a much needed body of knowledge on success in a web context. The web site user satisfaction construct provides researchers with a valid measure to investigate the effectiveness of web applications in supporting users' goal-directed information search processes. By virtue of this construct, traditional antecedents of IS success can be investigated on the web. For instance, the manner in which individual differences (e.g., cognitive style, personality, and demographic/situational variables) influence web site success can be tested using the WUS measure [92]. Furthermore, the WUS measure allows researchers to evaluate user involvement in a web context (e.g., if users have the possibility to customize a web site) [5,52].

6.2. Managerial contribution

The WUS measure can be used by practitioners for assessing the effectiveness, as well as the strengths and weaknesses of their organization's web site in a standardized way. By collecting users' evaluation, practitioners can also assess overall satisfaction with the web site and with specific hypermedia components (i.e., information, connection, and layout). Although practitioners may choose to include additional items to evaluate features specific to their web sites, the WUS measure is generic in nature and thus provides a common framework for comparative analysis.

In addition to the basic practical application of evaluating WUS at a certain point in time, practitioners can track WUS over time. Furthermore, the measure can be used to evaluate a company's web site against competitors' sites. Another important practical application may be the assessment of web site user satisfaction for different user segments (as users may exhibit differing product information needs, brand familiarity levels, bandwidth capacity, computer expertise, and layout preferences).

7. Conclusion

The primary significance of the research presented here is the conceptualization and empirical validation of web site user satisfaction. Through a solid grounding in IS success theory, hypermedia design theory, a qualitative exploratory pilot study and a quantitative online critical incident technique, we have introduced and defined the construct of web site user satisfaction, explored its dimensionality, provided empirical validation of the construct and its underlying dimensionality, developed a standardized instrument with desirable psychometric properties for measuring WUS, and explored the measure's theoretical and practical application.

Acknowledgements

The authors thank a leading international professional services firm and a leading international courier company for their financial support and assistance during data collection. The authors also gratefully thank three anonymous reviewers, the editor, as well as Hans Baumgartner and Bahr Weiss for their advice. The first author gratefully acknowledges financial support from the Intercollegiate Center for Management Science (Brussels).

Appendix A. WUS items (negatively stated items are indicated by an asterisk (∗))

- Information relevancy
  - IR1 The information in the web site is relevant.
  - IR2 The information in the web site is to the point.
  - IR3 The information in the web site does not help me at all (∗).
- Information accuracy
  - IA1 The information in the web site is accurate.
  - IA2 The information in the web site is not precise (∗).
  - IA3 I can rely on the information in the web site.
- Information comprehensibility
  - ICy1 The information in the web site is easy to understand.
  - ICy2 The information in the web site is clear to me.
  - ICy3 The information in the web site is presented clearly.
  - ICy4 The information in the web site is not presented in an adequate way (∗).

Appendix A. (*Continued*)

- Information comprehensiveness
  - ICs1 The web site provides sufficient information.
  - ICs2 The information in the web site is complete.
  - ICs3 The information in the web site is sufficiently detailed.

- Ease of use
  - E1 The web site is easy to use.
  - E2 The web site is user friendly.
  - E3 It is difficult to operate the web site (∗).

- Entry guidance
  - Eg1 The start page leads me easily to the information I need.
  - Eg2 The start page tells me immediately where I can find the information I am looking for.
  - Eg3 The start page does not provide direct access to what I need (∗).

- Web site structure
  - S1 The web site is clearly structured.
  - S2 The web site structure is confusing (∗).
  - S3 The structure of the web site is convenient.

- Hyperlink connotation
  - H1 The hyperlinks tell me clearly what I can find behind them.
  - H2 The hyperlinks are easy to understand.
  - H3 The meaning of some of the hyperlinks is unclear to me (∗).

- Web site speed
  - Sp1 The web site is time consuming (∗).
  - Sp2 It does not take much time to get from one place in the web site to another.
  - Sp3 The web site is fast.

- Layout
  - L1 I like the layout of the web site.
  - L2 The layout of the web site is visually comforting.
  - L3 The layout of the web site is annoying (∗).

- Language customization
  - Lg1 The information is provided in a desired language.

Appendix A. (*Continued*)

- Lg2 The choice of language is adapted to my needs.
- Lg3 I would like the information in another language (∗).

References

[1] I. Ajzen, Attitude structure and behavior, in: A.R. Pratkanis, S.T. Breckler, A.G. Greenwald (Eds.), Attitude Structure and Function, Erlbaum, Hillsdale, NJ, 1988.

[2] R.P. Bagozzi, H. Baumgartner, The evaluation of structural equation models and hypothesis testing, in: R.P. Bagozzi (Ed.), Principles of Marketing Research, Blackwell, Cambridge, MA, 1994, pp. 386–422.

[3] R.P. Bagozzi, Y. Yi, On the evaluation of structural equation models, Journal of the Academy of Marketing Science 16, 1988, pp. 74–94.

[4] J.E. Bailey, S.W. Pearson, Development of a tool for measuring and analyzing computer user satisfaction, Management Science 29 (5), 1983, pp. 530–545.

[5] J.J. Baroudi, M.H. Olson, B. Ives, An empirical study of the impact of user involvement on system usage and information satisfaction, Communications of the ACM 29 (3), 1986, pp. 232–238.

[6] H. Baumgartner, C. Homburg, Applications of structural equation modeling in marketing and consumer research: a review, International Journal of Research in Marketing 13 (2), 1996, pp. 139–161.

[7] J.A. Bellizi, Drawing prospects to e-commerce web sites, Journal of Advertising Research 40 (1–2), 2000, pp. 43–53.

[8] P.M. Bentler, Comparative fit indexes in structural models, Psychological Bulletin 107 (2), 1990, pp. 238–246.

[9] E. Berk, J. Devlin, Hypertext/Hypermedia Handbook, McGraw-Hill Software Engineering Series, New York, 1991.

[10] K.A. Bollen, Structural Equations with Latent Variables, Wiley, New York, 1989.

[11] K.A. Bollen, J.S. Long, Introduction, in: K.A. Bollen, J.S. Long (Eds.), Testing Structural Equation Models, Sage, Newbury Park, CA, 1993, pp. 1–9.

[12] H. Bornman, S.H. von Solms, Hypermedia, multimedia and hypertext—definitions and overview, Electronic Library 11 (4–5), 1993, pp. 259–268.

[13] M.W. Browne, Covariance structures, in: D.M. Hawkins (Ed.), Topics in Applied Multivariate Analysis, Cambridge University Press, Cambridge, 1982, pp. 72–141.

[14] M.W. Browne, R. Cudeck, Alternative ways of assessing model fit, in: K.A. Bollen, J.S. Long (Eds.), Testing Structural Equation Models, Sage, Newbury Park, CA, 1993, pp. 136–162.

[15] V. Bush, As we may think, Atlantic Monthly 176 (1), 1945, pp. 101–108.

[16] Q. Chen, W.D. Wells, Attitude toward the site, Journal of Advertising Research 39 (5), 1999, pp. 27–37.

[17] S.-Y. Choi, D.O. Stahl, A.B. Whinston, The Economics of Electronic Commerce, Macmillan Technical Publishers, Indianapolis, IN, 1997.
[18] CommerceNet. Overview of the 1998 Barriers and Inhibitors Research Project, Report #98-08, April 1998. http://www.commerce.net.
[19] CommerceNet/Nielsen Media Research, Results of Spring 1999 Internet Demographic Survey, Press Release 6/16/99. http://www.commerce.net/news/press/fact0699.html.
[20] J. Conklin, Hypertext: an introduction and survey, IEEE Computer 20 (9), 1987, pp. 17–41.
[21] T.D. Cook, D.T. Campbell, Quasi-Experimentation: Design & Analysis Issues for Field Settings, Houghton Mifflin Company, Boston, 1979.
[22] B. Cotton, R. Oliver, Understanding Hypermedia 2.000: Multimedia Origins, Internet Futures, Phaidon Press Ltd., London, 1997, p. 32.
[23] R. Cudeck, Analysis of correlation matrices using covariance structure models, Psychological Bulletin 105 (2), 1989, pp. 317–327.
[24] J. D'Ambra, R.E. Rice, Emerging factors in user evaluation of the world wide web, Information and Management 38, 2001, pp. 373–384.
[25] F.D. Davis, Perceived usefulness, perceived ease of use, and user acceptance of information technology, MIS Quarterly 13 (3), 1989, pp. 319–340.
[26] F.D. Davis, R.P. Bagozzi, P.R. Warshaw, User acceptance of computer technology: a comparison of two theoretical models, Management Science 35 (8), 1989, pp. 982–1003.
[27] W.H. DeLone, E.R. McLean, Information systems success: the quest for the dependent variable, Information Systems Research 3 (1), 1992, pp. 60–95.
[28] R.F. DeVellis, Scale Development: Theory and Applications, Sage, Newbury Park, CA, 1991.
[29] X. Drèze, F. Zufruyden, Testing web site design and promotional content, Journal of Advertising Research 37 (2), 1997, pp. 77–81.
[30] W.J. Doll, G. Torkzadeh, The measurement of end-user computing satisfaction, MIS Quarterly 12 (2), 1988, pp. 259–274.
[31] W.J. Doll, W. Xia, G. Torkzadeh, A confirmatory factor analysis of the end-user computing satisfaction instrument, MIS Quarterly 18 (4), 1994, pp. 453–461.
[32] W.J. Doll, T.S. Raghunathan, J.-S. Lim, Y.P. Gupta, A confirmatory factor analysis of the user information satisfaction instrument, Information Systems Research 6 (2), 1995, pp. 177–188.
[33] J. Eighmey, Profiling user responses to commercial web sites, Journal of Advertising Research 37 (3), 1997, pp. 59–66.
[34] J. Eighmey, L. McCord, Adding value in the information age: uses and gratifications of sites on the world wide web, Journal of Business Research 41 (3), 1998, pp. 187–194.
[35] H.J. Einhorn, R.M. Hogarth, Unit weighting schemes for decision making, Organizational Behavior and Human Decision Processes 13 (2), 1975, pp. 171–192.
[36] K.A. Ericsson, H.A. Simon, Think-Aloud Protocols as Data, CIP Working Paper, Department of Psychology, Carnegie-Mellon University, 1978.
[37] K.A. Ericsson, H.A. Simon, Verbal reports as data, Psychological Review 87, 1980, pp. 215–251.
[38] K.A. Ericsson, H.A. Simon, Protocol Analysis: Verbal Reports as Data, MIT Press, Cambridge, MA, 1984.
[39] J.C. Flanagan, The critical incident technique, Psychological Bulletin 51 (4), 1954, pp. 327–358.
[40] C. Fornell, D.F. Larcker, Evaluating structural equation models with unobservable variables and measurement errors, Journal of Marketing Research 18 (1), 1981, pp. 39–50.
[41] C.A. Gallagher, Perceptions of the value of a management information system, Academy of Management Journal 17 (1), 1974, pp. 46–55.
[42] F. Garzotto, L. Mainetti, P. Paolini, Hypermedia design, analysis, and evaluation issues, Communications of the ACM 38 (8), 1995, pp. 74–86.
[43] A.W. Gatian, Is user satisfaction a valid measure of system effectiveness? Information and Management 26 (3), 1994, pp. 119–131.
[44] S. Ghoze, W. Dou, Interactive functions and their impacts on the appeal of internet presence sites, Journal of Advertising Research 38 (2), 1998, pp. 29–43.
[45] S. Grabinger, Screen Design, in: P.A.M. Kommers, S. Grabinger, J. Dunlap (Eds.), Hypermedia Learning Environments: Instructional Design and Integration, Lawrence Erlbaum Associates, New Jersey, 1996.
[46] D.L. Hoffman, T.P. Novak, Marketing in hypermedia computer-mediated environments: conceptual foundations, Journal of Marketing 60 (3), 1996, pp. 50–68.
[47] O.R. Holsti, Content analysis, in: G. Lindzey, E. Aronson (Eds.), The Handbook of Social Psychology, vol. 2, Addison-Wesley, Reading, MA, 1968, pp. 596–692.
[48] E.K.R.E. Huizingh, The content and design of web sites: an empirical study, Information and Management 37, 2000, pp. 123–134.
[49] M. Igbaria, M. Tan, The consequences of information technology acceptance on subsequent individual performance, Information and Management 32 (3), 1997, pp. 113–121.
[50] J. Iivari, I. Ervasti, User information satisfaction: IS implementability and effectiveness, Information and Management 27 (4), 1994, pp. 205–220.
[51] T. Isakowitz, E.A. Stohr, P. Balasubramanian, RMM: a methodology for structured hypermedia design, Communications of the ACM 38 (8), 1995, pp. 34–44.
[52] B. Ives, M.H. Olson, User involvement and MIS success: a review of research, Management Science 30 (5), 1984, pp. 586–603.
[53] B. Ives, M.H. Olson, J.J. Baroudi, The measurement of user information satisfaction, Communications of the ACM 26 (10), 1983, pp. 785–793.
[54] R. Kalakota, A.B. Whinston, Frontiers of Electronic Commerce, Addison-Wesley, Reading, MA, 1996, p. 714.
[55] D.F. Larcker, V.P. Lessig, Perceived usefulness of information: a psychometric examination, Decision Sciences 11 (1), 1980, pp. 121–134.

[56] H.C. Lucas, Empirical evidence for a descriptive model of implementation, MIS Quarterly 2 (2), 1978, pp. 27–41.
[57] H.W. Marsh, D. Hovecar, Application of confirmatory factor analysis to the study of self-concept: first- and higher order factor models and their invariance across groups, Psychological Bulletin 97 (3), 1985, pp. 562–582.
[58] C.C. Marshall, F.M. Shipman, Spatial hypertext: designing for change, Communications of the ACM 38 (8), 1995, pp. 88–97.
[59] R.P. McDonald, H.W. Marsh, Choosing a multivariate model: noncentrality and goodness of fit, Psychological Bulletin 107 (2), 1990, pp. 430–445.
[60] W.J. McGuire, The nature of attitudes and attitude change, in: G. Lindzey, E. Aronson (Eds.), The Handbook of Social Psychology, vol. 3, 2nd ed., Addison-Wesley, Reading, MA, 1969, pp. 136–314.
[61] N.P. Melone, A theoretical assessment of the user-satisfaction construct in information systems research, Management Science 36 (1), 1990, pp. 76–91.
[62] M.B. Miles, A.M. Huberman, Qualitative Data Analysis: An Expanded Sourcebook, Sage Publications Ltd., London, 1994, p. 30.
[63] R. Moenaert, W.E. Souder, Context and antecedents of information utility at the R&D/marketing interface, Management Science 42 (11), 1996, pp. 1592–1610.
[64] B. Muthén, D. Kaplan, A comparison of some methodologies for the factor analysis of non-normal Likert variables, British Journal of Mathematical and Statistical Psychology 38, 1985, pp. 171–189.
[65] J. Nanard, M. Nanard, Hypertext design environments and the hypertext design process, Communications of the ACM 38 (8), 1995, pp. 49–56.
[66] National Science Foundation, Research Priorities in Electronic Commerce, NSF Workshop Report, 1999, p. 32. http://cism.bus.utexas.edu/workshop/ecdraft.html.
[67] J. Nielsen, Hypertext and Hypermedia, Academic Press, San Diego, CA, 1990.
[68] R.E. Nisbett, T.D. Wilson, Telling more than we can know: verbal reports on mental processes, Psychological Review 84, 1977, pp. 231–259.
[69] J.C. Nunnally, Psychometric Theory, 2nd ed., McGraw-Hill, New York, 1978, p. 245.
[70] P.C. Palvia, A model and instrument for measuring small business user satisfaction with information technology, Information and Management 31 (3), 1996, pp. 151–163.
[71] M.Q. Patton, Qualitative Evaluation Methods, Sage Publications Ltd., London, 1982.
[72] J.A. Quelch, L.R. Klein, The Internet and international marketing, Sloan Management Review 37 (3), 1996, pp. 60–75.
[73] J.E. Russo, E.J. Johnson, D.L. Stephens, The validity of verbal protocols, Memory and Cognition 17 (6), 1989, pp. 759–769.
[74] T. Saarinen, An expanded instrument for evaluating information systems success, Information and Management 31 (2), 1996, pp. 103–118.
[75] SAS/STAT User's Guide, Version 6, 4th ed., SAS Institute Inc., Cary, NC, 1990.
[76] W. Schuler, J. Hannemann, N. Streitz (Eds.), Designing User Interfaces for Hypermedia, Springer, Heidelberg, 1995.
[77] B. Shneiderman, Designing the User Interface: Strategies for Effective Human–Computer Interaction, Addison-Wesley, Reading, MA, 1987.
[78] B. Shneiderman, G. Kearsly, Hypertext Hands-On! Addison-Wesley, Reading, MA, 1989.
[79] J.H. Steiger, Structural model evaluation and modification: an interval estimation approach, Multivariate Behavioral Research 25, 1990, pp. 173–180.
[80] J. Steuer, Defining virtual reality: dimensions determining telepresence, Journal of Communication 42 (4), 1992, pp. 73–93.
[81] A. Strauss, J. Corbin, Basics of qualitative research, Sage, Newbury Park, CA, 1990.
[82] E.B. Swanson, Management information systems: appreciation and involvement, Management Science 21 (2), 1974, pp. 178–188.
[83] B. Szasjna, Determining information system usage: some issues and examples, Information and Management 25 (3), 1993, pp. 147–154.
[84] L.L. Thurstone, Multiple-Factor Analysis, University of Chicago Press, Chicago, 1947.
[85] L.R. Tucker, C. Lewis, The reliability coefficient for maximum likelihood factor analysis, Psychometrica 38, 1973, pp. 1–10.
[86] A. van Dam, Hypertext'87 keynote address, Communications of the ACM 31 (7), 1988, pp. 887–895.
[87] E.J. Webb, D.T. Campbell, R.D. Schwartz, L. Sechrest, Unobtrusive Measures: Nonreactive Research in the Social Sciences, Rand McNally College Publishing Company, Chicago, 1966.
[88] P. Wright, Message-evoked thoughts: persuasion research using thought verbalizations, Journal of Consumer Research 7 (2), 1980, pp. 151–175.
[89] R.K. Yin, Case Study Research Design and Methods, Sage Publications Inc., Beverly Hills, 1984.
[90] B. Yoo, N. Donthu, Developing a scale to measure the perceived quality of an internet shopping site, Quarterly Journal of Electronic Commerce 2 (1), 2001, pp. 31–46.
[91] R.W. Zmud, Concepts, theories and techniques—an empirical investigation of the dimensionality of the concept of information, Decision Sciences 9 (2), 1978, pp. 187–195.
[92] R.W. Zmud, Individual differences and MIS success: a review of the empirical literature, Management Science 25 (10), 1979, pp. 966–979.

Steve Muylle is an assistant professor of strategic marketing and e-business and director of the European Center for e-business studies at Vlerick Leuven Gent Management School. He is also an assistant professor of marketing research at Solvay Management School, Belgium. His research interests are in the areas of management of electronic commerce, business-to-business electronic commerce,

and web site interface design. His work has been accepted for publication in Communications of the ACM, Decision Support Systems, Journal of Marketing Communications, and IEEE Computer Society Proceedings.

Rudy Moenaert is an associate dean of research and professor of strategic marketing at TIAS Business School (University of Tilburg), as well as professor of marketing at Nyenrode University. His research interests focus on strategic marketing (business road-mapping) and the management and marketing of technological innovation in B2B settings. He has published in journals such as the Journal of Product Innovation Management, Management Science, R&D Management, IEEE Transactions on Engineering Management, and Journal of Management Studies.

Marc Despontin (1951) obtained his PhD in economics at the Vrije Universiteit Brussel (Free University of Brussels), Belgium. He is a professor of statistics and decision making at the same university, where he is also past dean and vice-rector. His research interests are applied quantitative methods in economics and management.