Testing a model of user-experience with news websites

Article *in* Journal of the Association for Information Science and Technology · June 2015

DOI: 10.1002/asi.23462

CITATIONS
12

READS
1,092

2 authors:

Gabor Aranyi
Erste Group
26 PUBLICATIONS **126** CITATIONS
SEE PROFILE

Paul van Schaik
Teesside University
164 PUBLICATIONS **2,476** CITATIONS
SEE PROFILE

Testing a Model of User-Experience with News Websites

Gabor Aranyi
School of Computing, Teesside University, Middlesbrough TS1 3BA, United Kingdom. E-mail: g.aranyi@tees.ac.uk

Paul van Schaik
School of Social Sciences, Business and Law, Teesside University, Middlesbrough, TS1 3BA, United Kingdom. E-mail: p.van-schaik@tees.ac.uk

Although the Internet has become a major source for accessing news, there is little research regarding users' experience with news sites. We conducted an experiment to test a comprehensive model of user experience with news sites that was developed previously by means of an online survey. Level of adoption (novel or adopted site) was controlled with a between-subjects manipulation. We collected participants' answers to psychometric scales at 2 times: after presentation of 5 screenshots of a news site and directly after 10 minutes of hands-on experience with the site. The model was extended with the prediction of users' *satisfaction* with news sites as a high-level design goal. A psychometric measure of *trust* in news providers was developed and added to the model to better predict people's intention to use particular news sites. The model presented in this article represents a theoretically founded, empirically tested basis for evaluating news websites, and it holds theoretical relevance to user-experience research in general. Finally, the findings and the model are applied to provide practical guidance in design prioritization.

Introduction

News media have had a constantly increasing web presence since the middle of the 1990s, when most of the major news providers in the Western world launched online versions of their newspapers. More recently, with the rapid spread of handheld electronic devices, such as tablet computers, smart phones, and e-book readers, online publishing has become a major source of media use. For example, by August 2012, sales of electronic books on Amazon surpassed those of print books by 14%, including print books with no electronic editions and excluding free electronic books (Malik, 2012). With regard to the decline of print news publishing in the United States, Mitchell and Rosenstiel (2012) report that, in 2011, losses in print advertising revenue outweighed gains in online advertising revenue 10 to 1, and that in terms of the combined circulation and advertising revenue, the print-newspaper sector has shrunk by 43% since 2000. The increase in the use of online news, as opposed to print news, is not attributable only to the spread of Internet-enabled handheld devices. Mitchell and Rosenstiel also point out that 8 in 10 people who access news on smart phones or tablets also access news on conventional computers. News sites are becoming the primary source of accessing news and daily information. The success of interactive products, such as news sites, is largely influenced by the extent to which they promote positive experiences in their users (Law & van Schaik, 2010). Indeed, research in user experience (UX) is primarily motivated by the notion that high-quality experiences with interactive products promote the adoption and repeated use of these products (Hassenzahl, 2003).

Model of UX with News Sites

News sites differ from other types of websites in many aspects, which carry important implications to their use, experience of use, evaluation, and quality criteria. News sites are a specific type of information-presenting website, which, although rooted in print journalism, make extensive use of the interactive features of the web as a publication platform (e.g., user-generated content and frequent updates). Because information-presenting websites are focused on the provision of information and information-related services, the quality criteria and models of transaction-based or retail-oriented sites cannot be applied to them directly (see Yang, Cai, Zhou, & Zhou, 2005). For example, whereas the successful, secure, and relatively effortless completion of a checkout procedure may be an important predictor of users' quality perceptions of retail sites, interesting and timely

Received July 14, 2014; revised September 15, 2014; accepted October 22, 2014

 DOI: 10.1002/asi.23462

content and source credibility may be more relevant for news sites. News sites are used differently, and for different reasons, than retail sites (see Aranyi, van Schaik, & Barker, 2012); typically for exploring news content with or without the presence of explicit information goals (e.g., wanting to see what's new or searching for the outcome of a particular sports game). These characteristics necessitate the separate treatment of news sites from other types of websites.

Publishing online news on various hardware platforms raises editorial, technical, and design issues that affect users' experience with these sites. For example, Nebeling, Matulic, and Norrie (2011) demonstrated that news sites are not adapted to large screen sizes that are widespread in homes, offices, and public spaces; in particular, the distribution of content elements does not scale well to large horizontal screens, which leads to a large amount of unused screen space. However, there is a general lack of knowledge as to which aspects of news sites are important, and to what extent, in driving a positive experience of use. Moreover, although there is a wealth of research in information systems (IS) about how users engage with online news (e.g., see O'Brien, 2011; O'Brien & Lebow, 2013; Arapakis, Lalmas, Cambazoglu, Marcos, & Jose, 2013), research in this area typically investigates engagement with news on the news-item level, whereas properties of news websites as interactive products in driving users' experience are not systematically considered. Information Science and human-computer interaction (HCI) models applied to specific domains, such as online news, can aid designers to identify and quantify factors that are important in predicting use, adoption, and positive attitudes toward a particular, or type of, interactive product. To address this need for research, we formulated a comprehensive model of UX with news sites (Aranyi & van Schaik, 2015).

The model of UX with news sites was formulated using data from an online survey that collected people's answers to a UX questionnaire after they used a news site of their choice (see Figure 1). The selection of variables for the model was guided by an exploratory think-aloud study (Aranyi et al., 2012) and a review of models from the HCI literature. The model brings together variables from the technology acceptance model (TAM) (Davis, 1989) and Hassenzahl's UX model (Hassenzahl, 2003), and emotional responses to predict UX outcomes, such as system appraisal and intention to use. Furthermore, the model links UX components to interaction characteristics, such as the perception of artifact properties that can be manipulated by design.

The components of user experience (CUE) model (Thüring & Mahlke, 2007) was used as a framework for developing the model, which integrates the most important aspects of human-technology interaction by incorporating various facets of interaction characteristics, UX, and system appraisal. We identified self-reported categories of experience with a news site in an exploratory think-aloud study (Aranyi et al., 2012), and identified constructs in the existing literature to address these categories, listed as perceived artifact characteristics in Figure 1. We restricted the

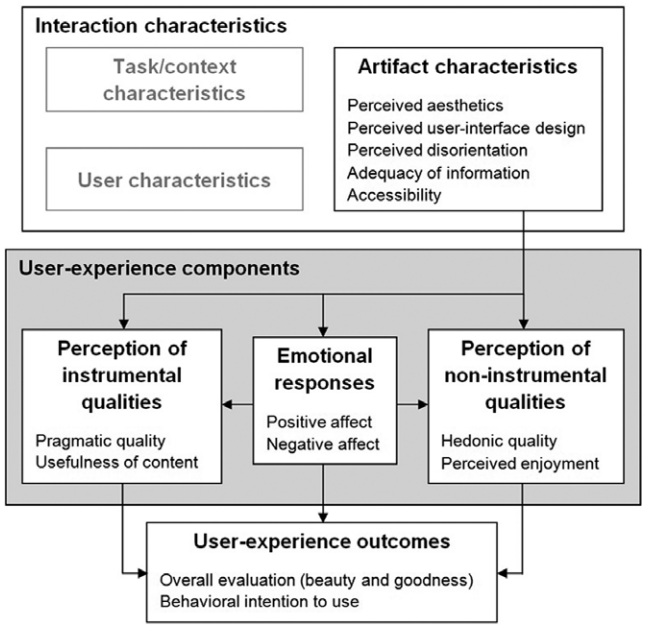

FIG. 1. High-level representation of the model of UX with news sites (based on Thüring & Mahlke, 2007).

measurement of perceived artifact characteristics to those categories of experience that emerged from the protocol analysis of think-aloud recordings collected during users' interactions with a news site. *Perceived aesthetics* (Porat & Tractinsky, 2012) and *perceived user-interface design* (Cho, Cheng, & Lai, 2009) were selected to address users' visual impressions and comments regarding page layout while interacting with news sites, respectively. Furthermore, visual aesthetics play an important role in UX studies (e.g., Hassenzahl, 2004; Zhou & Fu, 2007; van Schaik & Ling, 2009), and research suggests that the visual aesthetics of computer interfaces is a strong determinant of users' pleasure and satisfaction (Lavie & Tractinsky, 2004). We selected *perceived disorientation* (Ahuja & Webster, 2001) to address the issues related to navigating the complex structure of hypertext on a news site.

Furthermore, disorientation, the tendency to lose the sense of location on a website, is one of the most common problems faced by users navigating through hypertext (Nielsen, 2009) and it can lead to frustration, loss of interest, and decline in efficiency (McDonald & Stevenson, 1998; for a different view on disorientation in hypermedia systems, see Landow, 2006). To address participants' comments about news content, we included *adequacy of information* as a measure of perceived information quality, while *accessibility* was included as a measure of service quality, addressing participants' comments regarding the availability of content and loading time (Yang et al., 2005).

According to the CUE framework, UX components in the model include (a) instrumental qualities that concern usefulness and ease of use, and correspond to pragmatic attributes in Hassenzahl's (2003) UX model, (b) noninstrumental

DOI: 10.1002/asi

qualities that correspond to hedonic or "pleasure-producing" attributes in Hassenzahl's UX model, and (c) emotional responses. We used AttrakDiff2 (Hassenzahl, Diefenbach, & Göritz, 2010) to measure *pragmatic quality* and *hedonic quality*, *usefulness of content* (Yang et al., 2005) as a proxy of perceived usefulness,[1] and *perceived enjoyment* (Sun & Zhang, 2008) to measure the perception of instrumental and noninstrumental qualities, and the Positive and Negative Affect Schedule (PANAS) to measure emotional responses (Watson, Clark, & Tellegen, 1988), which has been used successfully in combination with AttrakDiff in previous research (Hassenzahl et al., 2010). Finally, we selected outcome measures both from the UX model (overall evaluative judgments of *beauty* and *goodness*) and TAM (*behavioral intention*). We argue for a joint consideration of UX and technology acceptance, because even if an interactive system is capable of delivering high-quality experiences, it will only be successful if users are willing to employ it (van Schaik & Ling, 2011). An online survey was used to collect data for each of the described psychometric instruments (see Aranyi & van Schaik [2015] for a detailed description of selection of measures and model development).

Interaction characteristics only included the measurement of artifact characteristics in the model. However, according to the CUE model (Thüring & Mahlke, 2007), which served as a basis for the current model, interaction characteristics also include variables related to task/context and person factors. Similarly, the person-artifact-task (PAT) model of flow experience in computer-mediated environments (Finneran & Zhang, 2003) identifies these three principal determinants of UX. The PAT model also proposes that these components do not only exert separate effects on people's experience with technology, but may also interact; therefore, it can also be important to consider the interaction of the effects of person, artifact, and task/context characteristics. The present study expands the model by incorporating a between-subject variable: level of adoption. Because novice users of a news site are likely to have a different quality of experience from experienced users (Aranyi et al., 2012), level of adoption (Magni, Taylor, & Venkatesh, 2010) is manipulated in the current study; half of the participants used a news site that they had been using regularly (experienced users), and the other half used a news site that they had never used before (novice users). Different levels of adoption are included in the current study so as to incorporate a person-factor among variables of interaction characteristics in the model.

The think-aloud study (Aranyi et al., 2012) that formed part of the basis of the model tested in the current study collected real-time accounts of participants' experience with a news site. However, research suggests that retrospective accounts or memories of experience may influence future decisions and behavior to a greater extent than accounts of real-time experience (Kahneman, 2000; Wirtz, Kruger, Scollon, & Diener, 2003). Consequently, collecting retrospective measures of experience can be considered more useful for prediction-based research. Therefore, retrospective measures of experience were applied to predict measures of UX outcomes in the study that provided data for the formulation of the model (Aranyi & van Schaik, 2015); however, the online questionnaire could not ensure that respondents actually used a particular news site just before answering the questions. To control participants' use of an artifact before collecting accounts of experience is important, because retrospective accounts of experience of a particular use episode may differ from cumulative accounts of experience based on a series of past use episodes (Roto, Law, Vermeeren, & Hoonhout, 2011). The laboratory setting applied in the present study ensured that participants used the news site before completing the UX questionnaires.

Additionally, the model is extended with two variables. Trust is added to the model, because it is considered as an important factor in the adoption of online news services (Chen & Corkindale, 2008). Satisfaction is added to the model as an additional outcome variable to enrich the model's predictive value of high-level design goals. Satisfaction differs from pleasure derived from interaction with a particular artifact in an important respect: Satisfaction involves an expectation about the interaction (Demir, Desmet, & Hekkert, 2009; Hassenzahl, 2003). For example, if a particular user expects to find information about a particular topic of his interest on a news site, logs on to the site, and finds a detailed discussion of this topic, it is expected that he will be satisfied. Conversely, if he finds no relevant information regarding his interest, the interaction is likely to result in dissatisfaction. However, he may experience pleasure (i.e., positive affect) despite being dissatisfied about the unfulfilled expectation because of, for example, finding something even more interesting on the news site that fulfills his need for stimulation. Pleasure may also be derived from the intrinsic pleasantness of an artifact (Demir et al., 2009); for example, the impressive graphics design of a news site (i.e., beauty) may be pleasing to the senses, without the possible fulfillment of needs or goals. Furthermore, in the terminology of Norman's (2004) emotional framework for product design, intrinsic pleasantness corresponds to the visceral level of emotional experience, whereas the expectation of behavioral or experiential outcomes and the subsequent satisfaction or dissatisfaction corresponds to the behavioral level of emotional experiences. Therefore, it is useful to include satisfaction as a separate outcome measure. Finally, based on the findings and the model, the present research also includes recommendations for product design and managerial guidance to aid design prioritization. The aims of the study presented will be summarized here.

Aims of the Current Study

Aim 1: test the research model. The primary aim of the study presented in this article is to test the model of UX with news sites in a controlled laboratory setting. As opposed to data collection with an online questionnaire (Aranyi & van Schaik, 2015), data collection through an experiment in a

DOI: 10.1002/asi

laboratory setting can ensure that participants' reports on UX measures are preceded directly by actual hands-on experience with a news site. Additionally, an experiment allows for controlling numerous confounding variables that may affect users' experiences with news sites by decreasing differences stemming from, for example, hardware specification, Internet connection speed, hardware platform (e.g., desktop computer, mobile phone, and tablet) and secondary tasks (e.g., chatting online and listening to music during news-site use).

Aim 2: extend the research model. Satisfaction is added to the model as an additional measure of UX outcome. UX measures are regressed onto satisfaction to identify its significant predictors and evaluate their predictive importance in terms of standardized regression coefficient, variance explained, and effect size. Furthermore, the current study includes the measurement of trust in particular service providers in a nonretail online setting. A measure of trust in news providers is developed and its psychometric properties are reported on.

The current study involves the experimental control of participants' level of adoption, thereby testing the model for a novel artifact, too; that is, a news site that has not been used (or even seen) before by the participants. Additionally, UX measures are administered both after presentation of screenshots of a particular news site and after using the site, for both an adopted and a novel news site, to test for possible differences in the effect of hands-on experience between the two conditions. Participants' reports of experience with an adopted and a novel news site are contrasted, and differences in model parameters between the conditions are tested.

Aim 3: apply the research model. The model was formulated on the basis of people's experiences with news sites of their own choice (Aranyi & van Schaik, 2015). However, models in HCI and IS are generally applied to particular products for various reasons, for example, to guide theoretical development (e.g., Venkatesh & Davis, 2000), for product evaluation and system design (e.g., Lewis, 2002), and for managerial guidance (e.g., Martensen & Grønholdt, 2003). Following model testing and extension, we provide practical implications of the model for use with particular news sites. Although the current study uses news sites as interactive artifacts, it is proposed that the UX model tested here can be adopted for, and applied to, interactive artifacts in general.

Method

Design

An independent-measures experimental design was used with two conditions in a laboratory setting. Participants in one group used the BBC news website, whereas the other group used the *New Zealand Herald* (NZH) news website during the experiment. Each participant recruited for the study was a regular user of the BBC news website (level of adoption: experienced user), whereas no participant used the NZH news website before (level of adoption: novice user). Participants completed the study in groups in a computer laboratory at the authors' institution.

Participants

Adult speakers of English were recruited for the experiment, who used the BBC news site on a regular basis. Every participant in the NZH condition confirmed that they had never used the *New Zealand Herald* before.[2] Participants were recruited from staff, research students, and master's students at the authors' institution. Participants received £8 cash compensation for their time and effort. Eighty-five people participated in the experiment (49 females and 36 males), with a mean age of 33.41 years (standard deviation [*SD*] = 9.96). Mean experience of using the Internet was 12.73 years (*SD* = 3.36). Mean time of Internet use per week was 28.09 hours (*SD* = 15.68). Mean time of using news sites per day was 43.29 minutes (*SD* = 66.59), and mean time of using news sites per visit was 23.08 minutes (*SD* = 36.82). Fifty participants (59%) reported using news sites for between 10 and 20 minutes per visit. Forty-nine participants (58%) reported using news sites daily or more frequently. Random assignment was successful in that no statistically significant between-group differences were found in demographic, Internet use, and news-site use.

Materials and Equipment

Bespoke experimental software was developed and used for presenting the experiment and collecting responses, coded in Visual Studio .NET 2008. The experimental software was run on identical personal computers (CPU: Intel Core2Duo E7500, 2.93 GHz; Memory: 3,072-MB RAM; OS: MS Windows 7 Enterprise 32-bit; Monitor: 17'' TFT, 1280×1024 screen resolution, 75-Hz refresh rate).[3] The BBC news website (http://www.bbc.co.uk) was selected as the adopted news site for the study, because the BBC is one of the most prestigious news portals in the United Kingdom,[4] and in the online questionnaire in Aranyi and van Schaik (2015), two thirds of participants chose to use the BBC when asked to use a news site of their choice. The *New Zealand Herald* (http://www.nzherald.co.nz) was selected for the study as a nonadopted site, because it was an English-language news site that participants in the UK were not likely to use, and it was among the most-visited news sites in New Zealand. Screenshots of the news sites used in the study are presented in Figure 2.

Because the main aim of this study was to test the model of UX with news sites, those measures of UX outcomes and UX components were selected for the measurement model that were used in a nonexperimental study (Aranyi & van Schaik, 2015). The 10-item AttrakDiff2-SF (Hassenzahl et al., 2010) was used to measure perceptions of *pragmatic quality* and *hedonic quality* (four items each), and *beauty*

DOI: 10.1002/asi

FIG. 2. Screenshots of the news sites used in the study. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

and *goodness* evaluations of the sites (one item each). A three-item questionnaire adopted from Sun and Zhang (2008) was used to measure *perceived enjoyment*. *Usefulness of content*, a proxy for *perceived usefulness*, was measured with a three-item scale adopted from Yang et al. (2005). *Behavioral intention* was measured by a two-item scale with appropriately different wording for the two experimental conditions. Affective reactions were measured using the 20-item PANAS (Watson et al., 1988). As measures of perceived artefact characteristics, the seven-item *perceived disorientation* scale (Ahuja & Webster, 2001), the three-item measure of *adequacy of information* (Yang et al., 2005), and the eight-item *perceived aesthetics* scale (Porat & Tractinsky, 2012) were used in the study. Additionally, a four-item measure of *trust* was added to measures of perceptions of instrumental qualities (*usefulness of content* and *pragmatic quality*). The items were worded based on the dimensions of trust in news media (Kohring & Matthes, 2007). A three-item measure of *satisfaction* was added to the indicators of interaction outcomes, adopted from Martensen and Grønholdt (2003). Items of each scale used in the study are presented in Appendix A.[5] The psychometric properties of the scales are discussed in the Measurement Model section and Appendix B.

Procedure

Upon their arrival in the computer lab, participants were randomly assigned to one of the two experimental conditions and completed a consent form. All instructions and material were presented on the computer screen, followed by a self-paced exposure of five screenshots of one of the sites (BBC or NZH, depending on the experimental condition). The screenshots were taken on pages with similar content from the news sites (home page at top, news page at top, news page scrolled down, sports page scrolled down, and entertainment page scrolled down) and were presented in the same order for the two conditions. The screenshots were taken on different pages to provide a cross-section of the sites, rather than just a picture of the home page. Following the screenshots, participants filled in AttrakDiff2-SF (T1–after presentation). Next, participants were instructed to use the news site for 10 minutes "to find out what's on." Attrakdiff2-SF was administered again after the participants finished using the site (T2–after interaction). All remaining measures were also administered after interaction with the site. Participants worked individually and they could not leave the experimental software until they had provided a full response to every question. The full procedure took approximately 30 minutes to complete.

Data Analysis

Partial-least-squares (PLS) path modeling was used for testing the model of UX with news sites, for the following reasons (see Vinzi, Chin, Henseler, & Wang, 2010). PLS is especially suitable for prediction-based research, for the integrated analysis[6] of complex, multistage models, and allows for the integrated analysis of a measurement model[7] and a structural model. PLS is compatible with multiple

DOI: 10.1002/asi

TABLE 1. Descriptive statistics and reliability coefficients.

Construct	No. of items	Average variance extracted	Composite reliability	Mean	*SD*
Adequacy of information	3	0.81	0.93	5.05	1.30
Beauty T1 (before use)	1	1.00	N/A	4.39	1.42
Beauty T2 (after use)	1	1.00	N/A	4.54	1.36
Behavioural intention	2	0.98	0.99	5.16	2.15
Goodness T1	1	1.00	N/A	5.06	1.52
Goodness T2	1	1.00	N/A	5.44	1.52
Hedonic quality T1	4	0.63	0.87	4.61	1.16
Hedonic quality T2	4	0.69	0.90	4.81	1.25
Perceived aesthetics	8	0.58	0.92	4.40	1.19
Perceived disorientation	7	0.57	0.90	2.24	1.12
Perceived enjoyment	3	0.86	0.95	4.96	1.47
Positive affect	10	0.55	0.92	4.22	1.28
Pragmatic quality T1	3	0.69	0.87	4.91	1.21
Pragmatic quality T2	3	0.70	0.87	5.29	1.26
Satisfaction	3	0.84	0.94	4.89	1.34
Trust	4	0.75	0.92	5.37	1.05
Usefulness of content	3	0.62	0.83	5.65	1.00

regression analysis, analysis of variance, and unrelated *t*-tests, which provide special cases of PLS results. Furthermore, as opposed to the technique of calculating the average of manifest variables to obtain latent variable scores, which does not account for measurement error, PLS uses the exact linear combination of manifest variables, thereby providing more-accurate scale values. PLS requires a lower sample size than covariance-based structural equation modeling (CB-SEM) and has less-stringent assumptions regarding the distribution of variables and error terms (see Chin, 2010). Recent research has shown that PLS performs at least as well as and, under various circumstances, is superior to CB-SEM in terms of bias, root mean square error, and mean absolute deviation (Hulland, Ryan, & Rayner, 2010; Vilares, Almeida, & Coelho, 2010). Nevertheless, PLS still requires an adequate sample size. The sample size in the present study ($n = 85$) satisfied the rule-of-ten requirement for robust estimations in PLS path modeling (Henseler, Ringe, & Sinkovics, 2009), which recommends a minimum number of participants 10 times the largest number of structural paths directed at a particular construct in the inner path model. Using the sample-size table provided by Marcoulides and Saunders (2006), we extrapolated a required minimum number of 35 participants to detect an effect with .80 power with normally distributed data (violated; average [av.] skew = –.56 [–1.24; .86]; av. kurtosis = .22 [–0.50; 1.53])[8] and no missing values (true) for the observed average latent-variable correlation of .51 and the observed average indicator loading of .77. Furthermore, in a recent study using Monte Carlo simulation, Goodhue, Lewis, and Thompson (2012) demonstrated across different levels of model complexity that PLS performs well in terms of statistical power and accuracy with a sample size around 90. Therefore, our collected sample size and the properties of the data are appropriate for PLS analysis. Moreover, if the assumption of normality was violated and sample size was low, the only consequence would be a loss of statistical power.[9] All PLS analyses in the current study were conducted using the SmartPLS software version 2.0.M3 (http://www.smartpls.de), with $n = 5{,}000$ bootstrapping samples to test the significance of model parameters (Henseler et al., 2009). The significance of differences between the path estimates for the two experimental conditions was calculated using *t*-tests, based on the differences between β-weights for the groups and the standard errors (SEs) of the path estimates for each group obtained through bootstrapping resampling procedures for calculating the pooled SE (see Chin, 2000).

Results

Measurement Model

Psychometric properties of measurement instruments used in the study were assessed in a measurement model, which specifies the relationships between observed indicators (i.e., scale items) and latent variables (Pedhazur & Pedhazur-Schmelkin, 1991). The measurement model was tested by drawing all possible structural links between each psychometrically measured variable and running the PLS algorithm with a factorial inner-weighting scheme in SmartPLS (Chin, 2010). Descriptive statistics and reliability coefficients of each scale retained for modeling are presented in Table 1. Details of reliability, convergent validity, discriminant validity, and item cross-loadings are presented in Appendix B.

Classical aesthetics and *expressive aesthetics* items of the two-dimensional perceived aesthetics measure (Porat & Tractinsky, 2012) were combined into a one-dimensional (1D) *perceived aesthetics* scale, based on psychometric considerations; cross-loadings between the two aesthetics subscales were high for each item, ranging between .33 and .66,

DOI: 10.1002/asi

and the subscales were highly and significantly correlated ($r = .67$, $p < .001$), indicating a lack of discriminant validity. Conversely, the 1D solution had high loadings (only one item loaded notably below .70), had high internal consistency (composite reliability = .92), and explained 58% variability in the items.

The internal consistency of each scale in the study was satisfactory, with composite reliability values over .80. Item PQ2 (unpredictable-predictable) was removed from the *pragmatic quality* scale, because it had a low loading at both times of measurement. Furthermore, the loading of PQ2 on the *pragmatic quality* scale was nonsignificant at T1. The entire *negative affect* scale was removed from the study, because it only retained an average of 38% variance from its indicators,[10] and 8 of the total 10 items had loadings below .70.[11]

Discriminant validity at the construct level was assessed against the Fornell-Larcker criterion (Fornell & Larcker, 1981), which posits that latent variables should share more variance with their indicators than with other latent variables; in other words, the average variance extracted (AVE) in each scale from their respective indicators should be larger than the squared correlation between the scale and all the other scales. Discriminant validity at the item level was assessed by the examination of item cross-loadings. The Fornell-Larcker criterion was violated in the case of *perceived aesthetics* and *hedonic quality*. Furthermore, *perceived aesthetics* shared an equal amount of variance (58%) with its items and with *goodness* at T2, and *hedonic quality* at T2 shared a similar amount of variance with its items (AVE = 69%) and with *goodness* at T2 ($R^2 = 66\%$). However, the examination of item cross-loadings supported the scales' discriminant validity at the item level, and the violation of the Fornell-Larcker criterion was treated as a result of high scale intercorrelations between *perceived aesthetics*, *hedonic quality* at T2, and *goodness* at T2 (ranging from .76 to .87). Therefore, these three measures were treated as separate in further analyses.

Model Testing

In order to test the model (Aim 1), structural relationships were drawn between measures of (a) artifact characteristics, (b) UX components, and (c) UX outcomes. Hypotheses for model testing were drawn directly from the model of UX with news sites, which was formulated using data from an online study with participants who used a news site of their own choice before completing a UX questionnaire (Aranyi & van Schaik, 2015). Nevertheless, a short rationale for hypotheses is also provided here from the relevant HCI literature to emphasize and illustrate the model's rooting in previous research and theory. The structural model and results of hypothesis tests are presented in Figure 3.

Hypotheses 1–5 pertain to the prediction of UX components from measures of perceived artifact characteristics. According to H1, *perceived disorientation* (Ahuja & Webster, 2001) is a negative antecedent of *pragmatic quality*, which corresponds to the user-perceived usability of an artifact (Hassenzahl, 2004). In other words, the more participants feel lost on the site, the less they judge the site to be usable. H1 was supported: *Perceived disorientation* alone accounted for 50% variability in *pragmatic quality* with a high impact ($\beta = .60$)[12] and large effect size ($f^2 = .80$).[13] According to H2, *adequacy of information*, as a measure of quality of information presented on the site, is positively related to *usefulness of content*, which is a proxy of *perceived usefulness* in the model (Chen & Corkindale, 2008). The hypothesis was supported: *Adequacy of information* had a high impact ($\beta = .44$) on *usefulness of content* with a large effect size ($f^2 = .62$), accounting for 47% of its variance, together with *positive affect*.

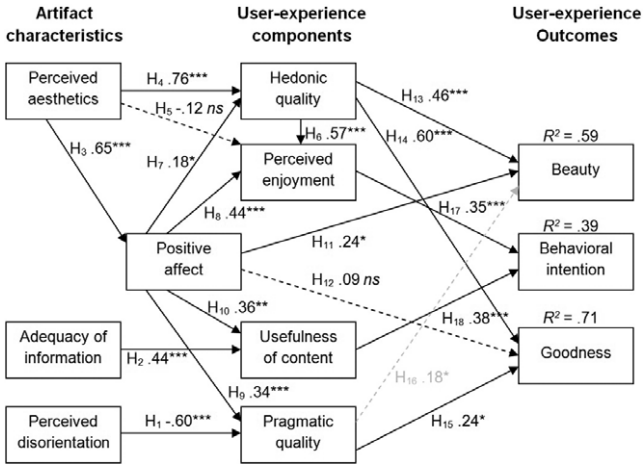

FIG. 3. Hypothesis tests of the model of UX with news sites.
Note. Dotted lines indicate paths that were hypothesized to be nonsignificant. Numbers represent standardized regression coefficients. The only hypothesis not supported is indicated in gray.
*$p < .05$, **$p < .01$, ***$p < .001$.

Visual aesthetics is recognized as an important pleasure-producing product characteristic in the UX literature (Hassenzahl & Tractinsky, 2006; Lavie & Tractinsky, 2004). Based on the findings in Aranyi and van Schaik (2015), we hypothesized that *perceived aesthetics* is a positive predictor of *positive affect* experienced during interaction with a news site (H3), and of the site's perceived ability to support positive experiences, that is, *hedonic quality* (H4). Both hypotheses were supported: *Perceived aesthetics* had a high impact with a large effect size both on *positive affect* (H3; $\beta = .65$, $f^2 = .72$, $R^2 = .42$)[14] and *hedonic quality* ($\beta = .76$, $f^2 = 1.60$, $R^2 = .79$). Additionally, in Aranyi and van Schaik (2015), the effect of *perceived aesthetics* on *perceived enjoyment* was fully mediated through *hedonic quality*; in other words, the extent to which aesthetics promoted enjoyment was mediated through the site's perceived capability to promote high-quality experiences. Accordingly, H5 proposed that *perceived aesthetics* is not a direct antecedent of *perceived enjoyment*, and H6 proposed that *hedonic quality* is an antecedent of *perceived enjoyment*. Both H5 and H6 were supported: The direct effect of *perceived aesthetics* on

DOI: 10.1002/asi

perceived enjoyment ($\beta = .33$, $t = 3.93$, $p < .001$) became nonsignificant with *hedonic quality* as mediator, supporting full mediation, and the direct effect of *hedonic quality* on *perceived enjoyment* was significant ($\beta = .57$, $t = 3.44$, $p < .001$).

Hypotheses 7–12 pertain to the connections between *positive affect* and perceptions of instrumental qualities (*hedonic quality*–H7 and *perceived enjoyment*–H8), perceptions of noninstrumental qualities (*pragmatic quality*–H9 and *usefulness of content*–H10), and overall product evaluations (*beauty*–H11 and *goodness*–H12). In Aranyi and van Schaik (2015), *positive affect* was positively and significantly related to measures of perceptions of instrumental qualities, perceptions of noninstrumental qualities, and *beauty* evaluations, but it was not related to *goodness* judgments. Each of these hypotheses was supported: *Positive affect* was a significant predictor of *hedonic quality* (H7; $\beta = .18$, $f^2 = .09$, $R^2 = .79$), *perceived enjoyment* (H8; $\beta = .44$, $f^2 = .32$, $R^2 = .67$), *pragmatic quality* (H9; $\beta = .34$, $f^2 = .26$, $R^2 = .60$), *usefulness of content* (H10; $\beta = .36$, $f^2 = .20$, $R^2 = .46$), and *beauty* judgments (H11; $\beta = .24$, $f^2 =$, $R^2 = .59$) with medium effect size, but it was not a predictor of *goodness* (H12; $\beta = .09$, not significant [*ns*]).

Hypotheses 13–18 pertain to the prediction of UX outcomes from measures of perceptions of instrumental and noninstrumental characteristics. Based on Hassenzahl's UX model (Hassenzahl, 2003, 2004) and empirical studies that confirmed these in the context of website use (van Schaik & Ling, 2008, 2011), we proposed that *hedonic quality* is a positive determinant of both *beauty* (H13) and *goodness* (H14) evaluations, whereas *pragmatic quality* is a positive determinant of *goodness* (H15), but it is not a determinant of beauty (H16). The relationships between *hedonic quality* and overall evaluations were confirmed: *Hedonic quality* was a significant positive predictor of both *beauty* (H13; $\beta = .46$, $f^2 = .23$, $R^2 = .59$) and *goodness* (H14; $\beta = .60$, $f^2 = .55$, $R^2 = .71$). H15 was also supported: *Pragmatic quality* was a significant determinant of *goodness* (H15; $\beta = .24$, $f^2 = .12$); however, contrary to H16, *pragmatic quality* was found to be a significant predictor of *beauty*, although it had the lowest path coefficient and effect size in the model ($\beta = .18$, $f^2 = .05$). In Hassenzahl's UX model (Hassenzahl, 2004), *pragmatic quality* is not a determinant of *beauty*. Similarly, van Schaik and Ling (2011) found that whereas *goodness* judgments of a web-based encyclopedia were influenced by both *hedonic quality* and *pragmatic quality*, *beauty* judgments were only influenced by ratings of *hedonic quality*. In a separate analysis of the experimental groups, *pragmatic quality* lost its significance on *beauty* in both conditions (BBC: $\beta = .21$, $t = 1.53$, $p = .13$, *ns*; NZH: $\beta = .16$, $t = 1.75$, $p = .08$, *ns*), likely owing to a decrease in sample size (total: $n = 85$; BBC: $n = 43$; NZH: $n = 42$), and the effect size was small in both groups (BBC: $f^2 = .05$; NZH: $f^2 = .06$).

Finally, based on the empirical studies using the technology acceptance model, we hypothesized that *perceived enjoyment* (Davis, 1989; Venkatesh & Davis, 2000) and use*fulness of content* (Cyr, Head, & Ivanov, 2006; Cyr, Hassanein, Head, & Ivanov, 2007; van Schaik & Ling, 2011) are positive determinants of *behavioral intention*. Both hypotheses were supported: *perceived enjoyment* (H17; $\beta = .35$, $f^2 = .15$, $R^2 = .39$) and *usefulness of content* (H18; $\beta = .38$, $f^2 = .18$) were both significant predictors of *behavioral intention* to use news sites.

In summary, test results demonstrated evidence for the model in a controlled setting, except for a significant path from *pragmatic quality* (i.e., user-perceived usability) to *beauty*, which contradicts the assumption of the Hassenzahl UX model; however, this path had a small effect size and path coefficient (for a detailed discussion of the relationship between beauty and perceived usability from another perspective, see Hassenzahl & Monk, 2010). UX components in the model accounted for a moderate-to-substantial proportion of variance in UX outcomes: 39% in *behavioral intention*, 59% in *beauty*, and 71% in *goodness*. Additionally, a moderate-to-substantial proportion of variance was explained in UX components: 42% in *positive affect*, 46% in *usefulness of content*, 60% in *pragmatic quality*, 69% in *perceived enjoyment*, and 79% in *hedonic quality*. Overall, coefficients of path strength, effect size, and explained variance indicate the strength of the model to describe relationships between the included measures and predict endogenous variables.

Model Extension

Consistent with Aim 2, we extended the model of UX with news sites with (a) the measurement of *trust* in news providers to better predict people's intention to use particular news sites and (b) *satisfaction* with news sites as a high-level design goal.

Trust. Based on a review of literature and in-depth interviews with news-media experts, Chen and Corkindale (2008) identified *trust* as one of the main drivers in the adoption of online news services. In their TAM for online news, *trust* serves as a determinant of *perceived usefulness* and a direct determinant of *behavioral intention*. There is empirical support for *trust* as a direct determinant of *behavioral intention* in the TAM applied to the domain of online retail (e.g., Gefen, Karahanna, & Straub, 2003; Pavlou & Gefen, 2004). It is therefore expected that the addition of *trust* to the model would improve the prediction of *behavioral intention*. Additionally, based on Chen and Corkindale (2008), it is proposed that trust is an antecedent of *usefulness of content* (a proxy for *perceived usefulness*). We assume that users may not derive their overall *beauty* judgments of news sites from their perceptions of the sites' trustworthiness as a whole. *Goodness* judgments, on the other hand, would be more plausible to be affected by *trust*. Therefore, the following hypotheses are proposed.

H19: *Trust* is a determinant of *behavioral intention*.

H20: *Trust* is an antecedent of *usefulness of content*.

DOI: 10.1002/asi

H21: *Trust* is not a determinant of *beauty*.

H22: *Trust* is a determinant of *goodness*.

The available measures of online trust in the literature either apply to the contexts of online retail and e-commerce (where trust is typically related to the safe treatment of payment details, which do not apply in the present context), or aim to measure the coverage of particular news events and news topics instead of the perceived trustworthiness of a particular news provider or site; therefore, a trust scale was developed for the study. Consistent with trust in journalistic assessment as a dimension of trust in news media (Kohring & Matthes, 2007), four items were worded to gauge trust as the perceived truthfulness and reliability of information presented on the site, and trust in the competence of journalists providing information to the site. Item 3 (see Appendix A) was adopted from Lavie and Tractinsky (2004), who originally used it to measure service quality. In the measurement model, our trust scale exhibited high internal consistency (composite reliability = .92) and retained 75% of variance from its items (*AVE* = .75). All item loadings were ≥.80, the scale exhibited satisfactory discriminant validity on the construct level, and the cross-loadings showed good discriminant validity at the item level (see Appendix B).

In support of H19, analysis revealed that *trust* was a significant predictor of *behavioral intention* ($\beta = .22$, $t = 2.43$, $p < .05$), but not of *beauty* ($\beta = -.04$, $t = .62$, *ns*; in support of H21) and *goodness* ($\beta = .01$, $t = .08$, *ns*; not in support of H22). H20, proposing that *trust* is an antecedent of *usefulness of content*, was not supported ($\beta = .11$, $t = 1.19$, *ns*). Because *trust* was a significant direct predictor of *behavioral intention* ($sr^2 = .04$, $f^2 = .07$), controlling for *usefulness of content* and *perceived enjoyment*, it was retained in the model among the perception of instrumental qualities, as a determinant of *behavioral intention*.[15]

Satisfaction. A measure of satisfaction was adopted as an additional outcome measure of experience, because it is a frequently used concept in expressing users' attitudes about artifacts among design practitioners (Cho, Park, Han, & Kang, 2011). User satisfaction or consumer satisfaction is a main concept in e-commerce and marketing research (Ha, 2006); it is an important antecedent of loyalty to providers of services (Anderson & Srinivasan, 2003) and it is often used in relation to providing managerial guidance. Research in IS and media studies shows that user satisfaction can be an informative measure of readers' experiences with news websites (e.g., Chung & Nah, 2009) and personalized (online) newspapers (e.g., Sela, Lavie, Inbar, Oppenheim, & Meyer, 2014).

The three-item satisfaction scale by Martensen and Grønholdt (2003) was adopted for the study. In compliance with different measurement practices in customer-satisfaction research (Ryan, Buzas, & Ramaswamy, 1995), the questionnaire contains an item each regarding (a) the overall level of satisfaction, (b) a judgment related to the extent to which a particular artifact meets one's expectations, and (c) how close the artifact is to one's ideal (see Appendix A). Furthermore, the original scale was successfully applied by Martensen and Grønholdt in their study using PLS path modeling. In our measurement model, the *satisfaction* scale exhibited high internal consistency (composite reliability = .94) and the scale retained 84% of variance from its three indicators. Discriminant validity was satisfactory, both at the construct and item level, and although several items showed cross-loadings on other scales, the loadings exceeded the cross-loadings in every case (see Appendix B).

TABLE 2. Tests of hypotheses related to predicting *satisfaction*.

	Predictor variable	β	*t*[a]	Supported
H23	Pragmatic quality	.23	1.96*	Yes
H24	Hedonic quality	.19	1.69	No
H25	Usefulness of content	.28	2.70**	Yes
H26	Perceived enjoyment	.34	2.71**	Yes
H27	Positive affect	−.08	0.86	No
H28	Trust	.19	2.96**	Yes

Note. [a]Bootstrap, $n = 5{,}000$.
$^*p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Each UX component was proposed as a predictor of satisfaction (H23–28). These hypotheses were guided by the assumption that higher levels of perceptions of instrumental and noninstrumental qualities, and positive affect experienced during the interaction, lead to higher levels of *satisfaction* in users of news sites. For example, Flavián, Guinalíu, and Guerra (2006) found that the usability of a website influences *satisfaction*, which, in turn, leads to greater loyalty to the site, and higher levels of *trust* were associated with higher levels of *satisfaction*. The hypotheses related to *satisfaction* and the corresponding test results are presented in Table 2.

In accord with the findings of Flavián et al. (2006), *pragmatic quality* ($f^2 = .11$) and *trust* ($f^2 = .10$) were positively and significantly related to *satisfaction*. *Usefulness of content* ($f^2 = .24$) and *perceived enjoyment* ($f^2 = .43$) were also significant predictors of *satisfaction*. The effects of *hedonic quality* and *positive affect* were not significant; however, both predictors serve as antecedents of perceived enjoyment in the model (see Figure 3). Mediation analysis confirmed the full mediation of the direct effect of *hedonic quality* on *satisfaction* ($\beta = .31$, $t = 2.49$, $p < .05$) through *perceived enjoyment*. The direct effect of *positive affect* on *satisfaction* ($\beta = .26$, $t = 2.76$, $p < .01$) was fully mediated through *perceived enjoyment* and *hedonic quality*. The significant predictors accounted for 74% of variability in *satisfaction*, which indicates that UX components are strong predictors of users' overall satisfaction with news sites.

Between-Group Differences

Differences in experience. In a think-aloud study exploring users' experience with a news website, Aranyi et al. (2012)

DOI: 10.1002/asi

TABLE 3. Differences in UX between the experimental conditions.

Variable	BBC mean (*SE*)	NZH mean (*SE*)	Mean diff.	*t*[a]	*df*[b]	*p*	Effect size *(r)*
Pragmatic quality (T1)[c]	5.09 (.17)	4.74 (.20)	0.35	1.36	83	.196	0.15
Hedonic quality (T1)	5.07 (.15)	4.14 (.17)	0.93	3.99	83	.000***	0.40
Beauty (T1)	4.84 (.20)	3.93 (.22)	0.91	3.11	82.04	.003**	0.32
Goodness (T1)	5.56 (.20)	4.55 (.24)	1.01	3.23	83	.002**	0.33
Pragmatic quality (T2)	5.53 (.17)	5.03 (.21)	0.50	1.86	83	.067	0.20
Hedonic quality (T2)	5.17 (.17)	4.44 (.20)	0.73	2.80	83	.006**	0.29
Beauty (T2)	4.74 (.16)	4.33 (.24)	0.41	1.40	71.72	.167	0.16
Goodness (T2)	6.00 (.15)	4.86 (.27)	1.14	3.71	62.73	.000***	0.42
Perceived disorientation	2.00 (.15)	2.49 (.19)	–0.49	–2.04	83	.044*	0.22
Perceived enjoyment	5.35 (.18)	4.57 (.25)	0.78	2.50	83	.014*	0.26
Positive affect	4.46 (.20)	3.98 (.19)	0.48	1.76	83	.083	0.19
Perceived aesthetics	4.75 (.17)	4.05 (.18)	0.70	2.82	83	.006**	0.30
Usefulness of content	5.98 (.13)	5.30 (.16)	0.68	3.29	83	.001***	0.34
Adequacy of information	5.28 (.19)	4.82 (.20)	0.46	1.65	83	.103	0.18
Trust	5.71 (.14)	5.02 (.16)	0.69	3.14	83	.002**	0.33
Behavioural intention	6.41 (.13)	3.88 (.36)	2.53	6.60	51.94	.000***	0.68
Satisfaction	5.39 (.15)	4.37 (.23)	1.02	3.76	70.61	.000***	0.41

Note. All variables were measured on a 7-point scale.
[a]Bootstrap, $n = 5{,}000$.
[b]Fractional degrees of freedom are presented where Levene's test for equality of variances was significant.
[c]T1, before use; T2, after use.
*$p < .05$, **$p < .01$, ***$p < .001$.

found that regular users of a particular news sites scored significantly and consistently higher on psychometric scales measuring UX than those who used the site for the first time. Similarly, the BBC and NZH groups in the current study differed in level of adoption (experienced and novice users, respectively). Independent samples *t*-tests were conducted on each UX measure to explore differences between the BBC and NZH groups (see Table 3).

With regard to outcome variables, the BBC received significantly higher ratings on *behavioral intention*, *satisfaction*, *goodness* (both before and after use), and before-use *beauty*. The BBC was not rated significantly higher on *beauty* than NZH after use (T2); however, the BBC was rated significantly more aesthetic than NZH. With regard to measures of UX components, the BBC received significantly higher ratings of *perceived enjoyment*, *usefulness of content*, *trust*, and *hedonic quality* (both before and after use). However, ratings of *pragmatic quality* did not differ significantly in the two conditions, neither before nor after use. Participants in both conditions were experienced users of news sites; therefore, they were familiar with their general layout. The results suggest that participants did not perceive the NZH, which they encountered for the first time, less usable than the BBC. Furthermore, ratings of *adequacy of information* and *positive affect* did not differ significantly between the experimental conditions.

There was no significant difference between the number of pages visited by BBC users (mean [*M*] = 22.51, *SE* = 1.66) and NZH users (*M* = 22.50, *SE* = 1.63), $t(83) = 0.01$, $p = .10$. Similarly, the number of home-page revisits for the BBC group (*M* = 3.09, *SE* = 0.43) and the NZH group (*M* = 3.26, *SE* = 0.53) was not significantly different, $t(83) = -0.30$, $p = .76$. Although the NZH group reported a higher level of disorientation than the BBC group, *perceived disorientation* was not correlated with the number of home-page revisits (BBC: $r = -.04$, $p = .79$; NZH: $r = -.02$, $p = .90$) and the number of pages visited (BBC: $r = .18$, $p = .24$; NZH: $r = -.18$, $p = .25$).

Although there was a significant difference in before-use *beauty* judgments in favor of the BBC, it was not rated significantly more beautiful than the NZH after use, which suggests a change in *beauty* judgments between before and after use. Previous research found judgments of *beauty* and perceptions of *hedonic quality* to be stable over time, whereas perceived usability and ratings of *goodness* were affected by experience (Hassenzahl, 2004). Participants completed the AttrakDiff2 questionnaire twice during the experiment: after a self-paced exposure of five screenshots (before use) and after 10 minutes use of the news site (after use). To assess the effect of interaction with the news sites on perceptions of product attributes and quality judgments, paired-samples *t*-tests were conducted between before and after use (see Table 4).

For the BBC as the adopted news site, *hedonic quality* and *beauty* judgments proved to be stable. However, *pragmatic quality* and *goodness* ratings were significantly higher after use. These findings are in line with those of Hassenzahl (2004), who argues that *beauty* judgments are not strongly affected by experience and *hedonic attributes* can be derived from the appearance of products, whereas use experience may provide further information about usability. However, the pattern of results shows the opposite for NZH, the non-adopted news site: *Pragmatic quality* and *goodness* proved to be stable, whereas *beauty* ratings were significantly

DOI: 10.1002/asi

TABLE 4. Paired-samples *t*-tests for AttrakDiff2 in the two experimental conditions.

Group	Variable	T1 Mean (*SE*)	T2 Mean (*SE*)	Mean diff.	*t*	*df*	*p*	Effect size (*r*)
BBC	Pragmatic quality	5.09 (.17)	5.53 (.17)	–0.44	–2.95	42	.005**	0.41
	Hedonic quality	5.07 (.15)	5.17 (.17)	–0.10	–0.87	42	.390	0.13
	Beauty	4.84 (.20)	4.74 (.16)	0.10	0.52	42	.605	0.08
	Goodness	5.56 (.20)	6.00 (.15)	–0.44	–2.50	42	.017*	0.36
NZH	Pragmatic quality	4.74 (.20)	5.03 (.21)	–0.29	–1.33	41	.190	0.20
	Hedonic quality	4.14 (.17)	4.44 (.20)	–0.30	–1.87	41	.069	0.28
	Beauty	3.93 (.22)	4.33 (.25)	–0.40	–2.07	41	.045*	0.31
	Goodness	4.55 (.24)	4.86 (.27)	–0.31	–1.65	41	.108	0.25

Note. *$p < .05$, **$p < .01$.

higher after use, and the increase of *hedonic quality* ratings after use approached significance. In an experiment using an information-oriented website, van Schaik and Ling (2009) found that ratings of *perceived aesthetics* increased after the use of the site. Although the present study did not measure *perceived aesthetics* before use, the current findings indicate that use of the site can provide further information about the appearance of a news site as a whole when the site is new to users.

After-use *hedonic quality* and *goodness* were significantly different between the experimental groups. Two analyses of covariance (ANCOVA) were conducted to compare the BBC and NZH groups on after-use *hedonic quality* and *goodness*, with before-use *hedonic quality* and *goodness* as covariates, respectively. The ANCOVA testing the effect of the experimental group on after-use *hedonic quality* with before-use *hedonic quality* as a covariate found the effect of the experimental group nonsignificant, $F(1, 81) = .004$, $p = .95$, $\varepsilon^2 = -.01$. However, the effect of before-use *hedonic quality* as a covariate was significant, $F(1, 81) = 68.52$, $p < .001$, $\varepsilon^2 = .33$, which suggests that the difference between the BBC and NZH groups in after-use *hedonic quality* was attributable to differences in before-use *hedonic quality* between the groups. In turn, *hedonic quality* was strongly correlated to *perceived aesthetics* ($r = .63$ before use and $r = .87$ after use, $p < .01$), which supports the notion that aesthetic perceptions can form very early (Lindgaard, Fernandes, Dudek, & Brown, 2006; see also Tuch, Presslaber, Stöcklin, Opwis, & Bargas-Avila, 2012) and may remain stable even after several minutes of hands-on experience. The ANCOVA testing the effect of experimental group on after-use *goodness* as a dependent variable with before-use *goodness* as a covariate yielded a similar result. Although both the effect of experimental group ($F[1, 81] = 11.50$, $p = .001$, $\varepsilon^2 = .06$) and the covariate ($F[1, 81] = 58.78$, $p < .001$, $\varepsilon^2 = .34$) were significant, the markedly higher effect size of the covariate suggests that differences in after-use *goodness* between the groups were predominantly attributable to differences in before-use *goodness* ratings.

Between-group differences in model parameters. Separate analyses for the BBC and NZH groups were conducted to explore the differences between the experimental conditions and test the robustness of the model across different levels of adoption and different artifacts. The results of between-group comparisons of path coefficients predicting UX outcomes are presented in Figure 4.

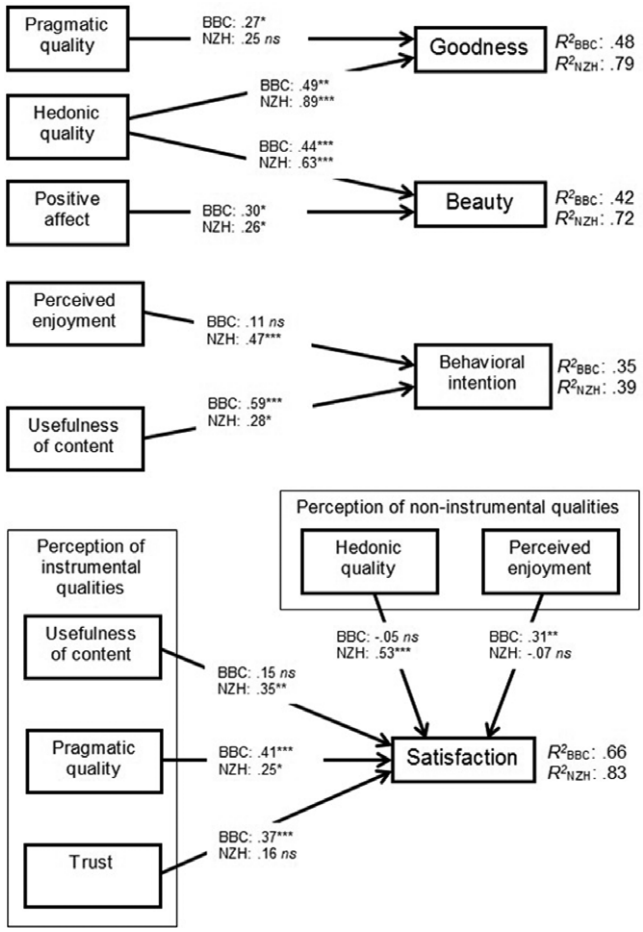

FIG. 4. Prediction of UX outcomes in the two experimental conditions. *Note*. The prediction of outcome measures is presented separately and variables of perceived artifact characteristics are excluded to avoid clutter; all variables are still part of the same single-path model.

The prediction of UX components from perceived artifact characteristics was robust across the conditions; therefore, these tests are not presented in Figure 4. The relationship between UX components differed slightly across conditions:

DOI: 10.1002/asi

The path from *hedonic quality* to *perceived enjoyment* ($\beta = .24$, *ns*) and the path from *positive affect* to *hedonic quality* ($\beta = .09$, *ns*) lost significance in the BBC condition, whereas both paths remained significant in the NZH condition ($\beta = .57$, $p < .001$; $\beta = .29$, $p < .01$, respectively). Moreover, the path from *hedonic quality* to *perceived enjoyment* was significantly different between the conditions ($\Delta\beta = .38$, $t = 2.25$, $p < .05$); *hedonic quality* was a high-impact and significant predictor of *perceived enjoyment* in the NZH condition, but its impact was smaller and nonsignificant in the BBC condition.

Although the drop in sample size resulting from separate analysis in the experimental conditions rendered some paths with smaller effect sizes nonsignificant, other significant paths indicate effects that are robust across the conditions. Differences of path coefficients between the experimental conditions allow for the examination of the model across two artifacts at different stages of adoption.

In general, prediction accuracy of outcome variables in the NZH condition was higher, regardless of the number of significant predictors of any given variable. In the model based on the pooled sample (see Table 2), the significant predictors of *satisfaction* were *pragmatic quality*, *usefulness of content*, *trust*, and *perceived enjoyment*. *Usefulness of content* was not a significant predictor of *satisfaction* in the BBC condition ($\beta = .15$, $t = 1.47$, *ns*, $f^2 = .04$), but it was significant and had a large effect size in the NZH condition ($\beta = .35$, $t = 2.61$, $p < .01$, $f^2 = .57$). The effect of *trust* on *satisfaction*, on the other hand, was significant in the BBC condition and represented a large effect size ($f^2 = .36$), but it was not significant in the NZH condition ($\beta = .07$, $t = 0.84$, *ns*, $f^2 = .03$). *Pragmatic quality* was the strongest predictor of *satisfaction* in the BBC condition with a medium effect size ($f^2 = .28$), whereas it was the least strong predictor in the NZH condition with a medium effect size ($f^2 = .19$). Nevertheless, user-perceived usability was a significant predictor of people's satisfaction with a news site in both conditions. Overall, these findings indicate that the model is applicable to news sites in general, but differences in the prediction of UX outcomes from UX components can be expected across different artifacts at different levels of adoption.

Discussion

Summary of Findings

The model of UX with news sites was tested in a laboratory setting with two experimental groups; one group used the BBC, the other the NZH. All participants were experienced users of the BBC, but no participant had used the NZH before the experiment. UX measures were collected following a self-paced presentation of screenshots and after 10 minutes of interaction with the sites.

Hypothesis tests provided further evidence for the research model (Aim 1); however, contrary to expectation, user-perceived usability was a significant predictor of *beauty*, but with small effect size. UX components predicted interaction outcomes to a moderate-to-substantial degree. To summarize users' general attitude toward news sites, a measure of satisfaction was included among the variables of interaction outcomes. *Satisfaction* was the most accurately predicted outcome variable, with substantial values of coefficients of determination in both experimental conditions. *Satisfaction* was predicted from a different set of UX variables in the two conditions, which indicates differences in how a judgment of satisfaction is constructed for novel and familiar artifacts.

Aim 2 was addressed by extending the research model with the constructs of trust and satisfaction. A psychometric measure of trust was constructed to assess the extent to which people perceive the provider of a particular news service to be trustworthy. The developed *trust* measure showed adequate psychometric properties (reliability, construct validity, and discriminant validity) and was a significant predictor of *behavioral intention* in the context of news-site use. Furthermore, *trust* was a significant predictor with a large effect size of *satisfaction* in the BBC group, whereas its impact was markedly lower and nonsignificant in the NZH condition. These findings provide evidence for the idea that the reputation of a news provider can be instrumental in driving satisfaction with news sites. Additionally, differences in the effect of *trust* on *satisfaction* between an adopted and nonadopted news service provide evidence for the sensitivity of the developed trust measure. Chen and Corkindale's (2008) assumption that *trust* is an antecedent of *perceived usefulness* in the context of news-site use was not supported.

Hassenzahl (2004) found that perceptions of *beauty* and *hedonic quality* are stable over time, whereas perceptions of *pragmatic quality* and *goodness* judgments of interactive artifacts are more likely to be affected by hands-on experience. Hassenzahl's findings were supported in the analysis of the stability of perceptions of product attributes and overall quality judgments in the BBC group. However, comparison of after-presentation and after-interaction ratings in the NZH group revealed that though there were no significant differences in the levels of *pragmatic quality* and *goodness*, *beauty* was rated significantly higher after use and the difference represented a medium effect size. This finding suggests that the effects of hands-on experience on product judgments can differ across levels of adoption, and that hands-on experience can provide more additional information about the appearance of a news site than about its usability when users browse a site without set tasks.

Practical Implications

Consistent with Aim 3, we provide practical implications based on the model and illustrate the application of the model to particular news sites. The current study supported the importance of content in driving the use and adoption of a news site (Chen & Corkindale, 2008); quality of information presented on news sites, conceptualized as *usefulness of*

DOI: 10.1002/asi

content, was a significant predictor of use intention. In turn, *adequacy of information*, as a measure of the completeness and comprehensiveness of content presented on the site, was a significant predictor of *usefulness of content*. Although *trust* significantly correlated with *adequacy of information* and *usefulness of content*, it was an independent determinant of use intentions and *satisfaction*, which suggests that the perceived credibility and competence of a particular news provider (i.e., reputation) may not be exclusively derived from perceptions of information quality. For example, *design aesthetics* was a direct predictor of *trust* in an experiment using retail websites as artifacts while controlling for *usefulness*, *ease of use*, and *customization* (Li & Yeh, 2010). In summary, the model suggests that quality of information as perceived by readers can be promoted by increased adequacy of information. Information quality is also positively associated with *trust* in the news provider as a determinant and it is an important driver of intention to use and satisfaction with news sites.

Information architecture in the context of news-site use is related to issues with navigating complex, highly interconnected structures of hyperlinks on frequently updated websites containing hundreds or thousands of pages. Users' level of disorientation on news sites was strongly and negatively related to the perceived usability of the sites, which, in turn, was a predictor with medium effect size of *goodness* and *satisfaction*. Therefore, it is important to address information architecture to improve the users' perception of a news site's usability. Good information architecture is expected to result in users experiencing less disorientation. The information architecture of news sites affects the efficiency of users finding the news they are looking for (Li, 2002). Moreover, according to Nielsen (2009), information architecture is still the most important factor causing usability problems for users of websites.

Information architecture[16] can be improved by addressing the labeling, organization, and navigation systems of websites (Rosenfeld & Morville, 2006). In particular, the labeling system can be improved by increasing the information scent of links and headings on a particular site, the extent to which readers can predict the content they will find by following hyperlinks (Nielsen, 2004); in this sense, a link label acts as an "advanced organizer." The probability of users feeling lost on a website can also be decreased by improving the organization of information and by facilitating the ease of navigation between pages presenting related content. Issues with information architecture can be identified with expert-based usability tests, such as "cognitive jogthrough" (Rowley & Rhoades, 1992), and user tests, such as card sorting and rating (Bernard, 2000), but these methods can only be applied to small and medium sites (less than approximately 100 pages). News sites are typically large websites that may contain thousands of pages. However, cognitive computational modeling techniques, such as the comprehension-based linked model of deliberate search (Kitajima, Blackmon, & Polson, 2000), have been proposed to offer an automated solution to analyzing the information architecture of large websites. Although good information architecture may not be essential for finding particular pieces of information on the web using search engines, news-site use typically involves looking for news on a particular site; in other words, information goals are not (always) defined in advance. Therefore, we argue that good information architecture is important for news sites, even when an embedded search function is provided.

Overall, the findings of the current study support the importance of noninstrumental aspects of interaction with news sites. The differences in the strength of path coefficients between novice and experienced users imply that noninstrumental aspects of interaction, such as enjoyment and perceptions of hedonic quality, may play an even greater role in new users' experiences.

Although prediction-focused models drawn from laboratory studies may be theory driven, these models can be flexible in terms of measurement specifications and prediction goals. For example, measured aspects of experience and interaction characteristics can be tailored to a specific news site, based on managerial and design considerations. With regard to the flexibility of prediction goals, aspects of experience and interaction characteristics can be used to predict dimensions of interaction of specific interest. For example, a prediction target can be the satisfaction of users with a particular news site, another target can be the users' intention to revisit the site, and yet another target may be the users' positive attitudes about the news provider. Laboratory studies and surveys involving the collection of larger data sets can be designed according to specific prediction goals. A utility of the model of UX with news sites presented here lies in providing a theoretically founded, empirically tested basis for the selection of measures and their structural relationships. Experiments and surveys can be used to test different designs of the same online news medium simultaneously (within-medium testing) and test different versions of the same medium before and after redesign (design-improvement testing). Practical guidance can be expressed in the form of impact-performance matrices (Martensen & Grønholdt, 2003) to guide managerial and design decisions.

In impact-performance matrices, the performance scores[17] of variables predicting each outcome measure are presented in relation to their impact scores. An example of an impact-performance matrix based on the NZH group ($n = 42$) with *satisfaction* as the outcome variable is presented in Figure 5. Parameters of the analysis are presented in Table 5.

The x-axis represents the impact scores (standardized regression coefficients) of the predictors of *satisfaction*. The vertical line represents the average impact of the three predictors. The y-axis represents the mean performance scores of participants on the predictors, which are latent variable scores acquired from running the PLS procedure.[18] The horizontal line represents the average performance of the predictors. The matrix can be used to assess which areas of UX characteristics need to be addressed in design. The vertical

DOI: 10.1002/asi

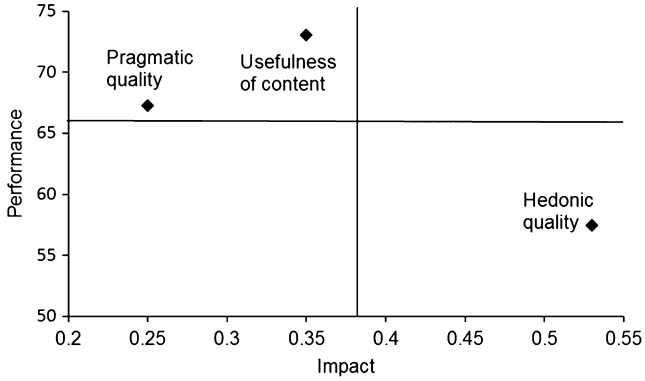

FIG. 5. Impact-performance matrix of satisfaction for NZH.

TABLE 5. Parameters of the impact-performance analysis of the NZH group.

Variable	Mean performance (*SD*)	Impact (β)
Hedonic quality (predictor)	57.43 (21.66)	.53
Pragmatic quality (predictor)	67.27 (22.72)	.25
Usefulness of content (predictor)	72.98 (17.29)	.35
Satisfaction (criterion)	56.23 (24.52)	
Average impact of predictors (*SD*):	.38 (.14)	
Average performance of predictors (*SD*):	65.89 (7.87)	
Variance explained in satisfaction:	83%	

(average impact) and horizontal (average performance) lines in impact-performance matrices divide the matrix into four cells. Each predictor variable can be placed in one of the quadrants. The lower-left quadrant contains variables that have a low performance, and have a low impact on the level of the dependent variable at the same time, relative to other predictors in the matrix. Low-performing significant determinants of interaction quality can be addressed in design. However, because the low-performing determinants in the bottom-left quadrant also have relatively low impact on interaction quality, it is not important to assign high priority to these aspects of interaction in system (re)design. No variable falls into this quadrant in the current example.

The upper-left quadrant contains relatively high-performing, but low-impact, variables. Variables that fall into this quadrant need not be prioritized for design intervention, because they are already performing relatively high and their impact is relatively low. It is suggested that high performance on variables in the upper-left quadrant, although not crucial, is beneficial for maintaining the high quality of service. *Pragmatic quality* and *usefulness of content* fall into this quadrant in the current example. Both predictors are rated relatively high, compared to *hedonic quality*; however, the performance scores indicate that there is space for improvement for both variables, especially *pragmatic quality*.

The lower-right quadrant contains relatively low-performing, but high-impact, variables. Variables that fall into this area are the most important targets for design intervention. Variables in the lower-right quadrant are important predictors, but are not performing high relative to other predictors. *Hedonic quality* falls into this quadrant in the current example. *Hedonic quality* has the highest impact in predicting *satisfaction*, but it performs at a relatively low level, which indicates room for improvement. The impact-performance analysis suggests that satisfaction of novice users of the NZH site could be improved predominantly by addressing the site's aesthetic appeal in design (e.g., see Hekkert, Snelders, & van Wieringen, 2003; Hekkert, 2006; Tuch et al., 2012 for practical guidance). Furthermore, according to Hassenzahl and Roto (2007), hedonic-quality perceptions play an important part in *emotional product attachment*, which suggests that people may be more likely to continually use and describe themselves as users of a particular artifact if they perceive the site to be highly conducive to having high-quality experiences. Finally, the upper-right quadrant contains variables that are important predictors and perform high at the same time. Predictors that fall into this quadrant represent strong areas of design that should be kept at a high level. No variable falls into this quadrant in the current example.

In summary, quantitative UX data can be used to construct impact-performance matrices, which can be used to identify areas of improvement for design. Impact-performance matrices are flexible in terms of prediction goals and can be drawn for any criterion variable with two or more significant predictors. It is important for the predictors in impact-performance matrices to be interpretable in terms of design. In the current examples, predictors were all interaction-characteristic variables that were connected to artifact characteristics. Although the matrix presented here was drawn on the basis of the model of UX with news sites, any plausible set of predictors can be measured to predict criterion variables of particular interest for any particular artifact. However, the current model can be used to identify relevant sets of predictors of UX outcomes and describe their relationships for the evaluation of (versions of) news sites.

Limitations and Future Work

A limitation of the current study is that level of adoption and news site were not independently manipulated. Consequently, this study cannot distinguish between differences arising from level of adoption and from differences in artifact characteristics. The design of the current study, however, was constrained by practical considerations, but this limitation did not hinder the aims of the study: (Aim 1) testing the model in a controlled setting; (Aim 2) model extension; and (Aim 3) model application to particular news sites. The experiment demonstrated the model's ability to account for a large amount of variance in endogenous variables related to UX and UX outcomes, the measures

DOI: 10.1002/asi

demonstrated sensitivity to distinguish between different conditions, and the model was generally supported in a controlled setting, using both an adopted and a nonadopted artifact.

Because existing sites were used in the current study, it did not include the systematic manipulation of artifact characteristics in order to examine the effects of manipulations on users' experiences. However, the current model is predominantly based on the components of the UX model (Thüring & Mahlke, 2007), the formulation of which was based on systematic manipulation of design characteristics of interactive artifacts (mobile telephone and portable audio player layouts). Furthermore, connections between variables in the current model that pertain to technology-acceptance and UX constructs were tested separately in previous HCI research involving other interactive artifacts (see Aranyi, 2012).

The model fit, in terms of prediction accuracy of outcome variables, was markedly higher in the NZH condition, which suggests that the model could be applied more accurately to certain news sites or subtypes of sites (e.g., subscription-based vs. advertisement-based business model). Model testing would benefit from the use of multiple news sites, and recruiting a larger number of experienced and novice users to each of the sites. The model could be used in future work to evaluate existing news sites (with using experienced and novice users as participants) and to evaluate different versions of news sites (design-improvement testing). Finally, the current study used personal computers to display the news sites. Future work may involve testing a news site across multiple platforms, such as mobile phones and tablet computers.

Endnotes

1. A proxy was used because the usefulness of content scale describes usefulness more accurately in the present context (i.e., news-site use with no set tasks) than a generic perceived-usefulness scale (which concerns a particular artifact's capability of promoting the attainment of behavioral goals).

2. Note that artifact characteristics and level of adoption (novice/experienced users of the site used during the study) were confounded (not controlled independently); implications are discussed in the Limitations and Future Work section.

3. The experiment software used the standard settings for the browser (Internet Explorer) that had already been defined on the machines that were used (identical on all machines).

4. According to Alexa web analytics (www.alexa.com), the BBC is the highest-ranking news site in the UK and the 11th-highest ranking news site in the world (as of July 2014).

5. Note that we did not include *perceived user-interface design* (PUID) and *accessibility* among measures of perceived artifact characteristics in the current study: PUID was excluded for parsimony (it shared a large amount of variance with aesthetics), and *accessibility* was not relevant in the present laboratory setting.

6. In this context, "integrated" means that all relationships between the variables are estimated simultaneously, as opposed to the technique of conducting a separate (multiple) regression analysis for each criterion variable.

7. Note that chi-square statistics and various goodness-of-fit measures applied in CB-SEM are not applicable in PLS. Nevertheless, PLS measurement models can still be adequately evaluated. In any case, CB-SEM fit indices indicate the extent to which parameter estimates are able to match sample covariances (measurement model), but do *not* indicate the goodness of the (structural) model in terms of predictive power or theoretical sense (see Chin, 1998a, 2010).

8. Goodhue et al. (2012) note that PLS, as well as CB-SEM, is remarkably robust against moderate departure from normality (up to skew = 1.1 and kurtosis = 1.6).

9. Furthermore, the probability of false positives (type II errors) is lower in small samples when PLS is used than when CB-SEM is used (Goodhue et al., 2012).

10. According to Henseler et al. (2009), average variance extracted from the indicators should be greater than 50% for each scale.

11. Generally, loadings should be above .70 in a measurement model, which indicates that at least 50% of variance is shared by a particular indicator and the latent variable it loads on (Henseler et al., 2009).

12. Note that, according to Chin (1998a), standardized path coefficients (β; also referred to as *impact*) should be around .20 (and ideally above .30) to be considered meaningful.

13. Cohen's (1988) conventions for effect size in multiple regression were used: 0.02–0.14 for small, 0.15–0.34 for medium, and 0.35 and above for large.

14. R^2 values can be interpreted as the total variance explained in a criterion by its predictors (e.g., $R^2 = 0.42$ stands for 42% variance explained). According to Chin (1998b), R^2 values for endogenous variables are characterized as: .67—substantial, .33—moderate, and .19—weak.

15. A figure displaying extensions to the model is not presented here separately to avoid duplication. See Figure 4 in the Between-group differences in model parameters section for path models including *trust* and *satisfaction*.

16. Although no single accepted definition exists, the main aim of information architecture is to provide effective access to relevant online information resources (usually delivered through a website). Information architecture focuses on systems for the organization, labeling, and navigation of information (Rosenfeld & Morville, 2006) for the benefit of end users who need to find and use information. The role of an information architecture is to provide the design of information for building a website, but does not address information presentation in a site in terms of, for example, layout and the use of color and graphics.

17. Note that "performance" in this context refers to the average magnitude of scores on a particular measure (i.e., how well an artifact performs on a measure), and not to the performance of participants.

18. For the purpose of this analysis, the performance scores have been transformed to a common 0–100 scale to promote the interpretation of scores in terms of percentage achieved from the maximum score.

References

Ahuja, J.S., & Webster, J. (2001). Perceived disorientation: An examination of a new measure to assess web design effectiveness. Interacting With Computers, 14(1), 15–29.

Anderson, R.E., & Srinivasan, S.S. (2003). E-satisfaction and e-loyalty: A contingency framework. Psychology & Marketing, 20(2), 123–138.

Aranyi, G. (2012). Developing a psychological model of end-users' experience with news web sites. (Unpublished doctoral dissertation). Teesside University, Middlesbrough, UK. Retrieved from http://tees.openrepository.com/tees/handle/10149/236331

Aranyi, G., & van Schaik, P. (2015). Modeling user-experience with news websites. Journal of the Association for Information Science and Technology. Retrieved from http://sss-studnet.tees.ac.uk/psychology/staff/Paul_vs/papers/Aranyi-van-Schaik-JASIST-2014.pdf

Aranyi, G., van Schaik, P., & Barker, P. (2012). Using think-aloud and psychometrics to explore users' experience with a news Web site. Interacting With Computers, 24(2), 69–77.

Arapakis, I., Lalmas, M., Cambazoglu, B.B., Marcos, M.C., & Jose, J.M. (2013). User engagement in online news: Under the scope of sentiment, interest, affect, and gaze. Journal of the Association for Information Science and Technology, 65(10), 1988–2005. doi:10.1002/asi.23096

Bernard, M. (2000). Constructing user-centered websites: The early design phases of small to medium sites. Usability News, 2(1), Article 8. Retrieved from http://www.surl.org/usabilitynews/21/webdesign.asp

Chen, Y.H., & Corkindale, D. (2008). Towards an understanding of the behavioral intention to use online news services: An exploratory study. Internet Research, 18(3), 286–312.

Chin, W.W. (1998a). Commentary: Issues and opinion on structural equation modeling. MIS Quarterly, 22(1), vii–xvi.

Chin, W.W. (1998b). The partial least squares approach for structural equation modelling. In G.A. Marcoulides (Ed.), Modern methods for business research (pp. 295–336). Mahwah, NJ: Erlbaum.

Chin, W.W., (2000). Frequently asked questions—Partial least squares & PLS-graph. Home page. Retrieved from http://disc-nt.cba.uh.edu/chin/plsfaq.htm

Chin, W.W. (2010). How to write up and report PLS analyses. In V.E. Vinzi, W. Chin, J. Henseler, & H. Wang (Eds.), Handbook of partial least squares: Concepts, methods and applications in marketing and related fields (pp. 655–690). Berlin: Springer.

Cho, V., Cheng, T.C.E., & Lai, W.M.J. (2009). The role of perceived user-interface design in continued usage intention of self-paced e-learning tools. Computers & Education, 53(2), 216–227.

Cho, Y., Park, J., Han, S.H., & Kang, S. (2011). Development of a web-based survey system for evaluating affective satisfaction. International Journal of Industrial Ergonomics, 41(3), 247–254.

Chung, D.S., & Nah, S. (2009). The effects of interactive news presentation on perceived user satisfaction of online community newspapers. Journal of Computer-Mediated Communication, 14(4), 855–874.

Cohen, J. (1988). Statistical power analysis for the behavioural sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

Cyr, D., Head, M., & Ivanov, A. (2006). Design aesthetics leading to m-loyalty in mobile commerce. Information & Management, 43(8), 950–963.

Cyr, D., Hassanein, H., Head, M., & Ivanov, A. (2007). The role of social presence in establishing loyalty in e-service environments. Interacting With Computers, 19(1), 43–56.

Davis, F.D. (1989). Perceived usefulness, perceived ease of use and user acceptance of information technologies. MIS Quarterly, 13(3), 319–340.

Demir, E., Desmet, P.M.A., & Hekkert, P. (2009). Appraisal patterns of emotions in human-product interaction. International Journal of Design, 3(2), 41–51.

Finneran, C.M., & Zhang, P. (2003). A person-artefact-task (PAT) model of flow antecedents in computer-mediated environments. International Journal of Human-Computer Studies, 59(4), 475–496.

Flavián, C., Guinalíu, M., & Guerra, R. (2006). The role played by perceived usability, satisfaction and consumer trust on website loyalty. Information & Management, 43, 1–14.

Fornell, C., & Larcker, D.F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. Journal of Marketing Research, 18(3), 328–388.

Gefen, D., Karahanna, E., & Straub, D.W. (2003). Trust and TAM in online shopping: An integrated model. MIS Quarterly, 27(1), 51–90.

Goodhue, D.L., Lewis, W., & Thompson, R. (2012). Does PLS have advantages for small sample size or non-normal data? MIS Quarterly, 36(3), 981–1001.

Ha, H.-Y. (2006). An integrative model of consumer satisfaction in the context of e-services. International Journal of Consumer Studies, 30(2), 137–140.

Hassenzahl, M. (2003). The thing and I: Understanding the relationship between user and product. In M. Blythe, C. Overbeeke, A. Monk, & P. Wright (Eds.), Funology: From usability to enjoyment (pp. 31–42). Dordrecht, The Netherlands: Kluwer.

Hassenzahl, M. (2004). The interplay of beauty, goodness, and usability in interactive products. Human-Computer Interaction, 19(4), 319–349.

Hassenzahl, M., & Monk, A. (2010). The inference of perceived usability from beauty. Human-Computer Interaction, 25(3), 235–260.

Hassenzahl, M., & Roto, V. (2007). Being and doing: A perspective on user experience and its measurement. Interfaces, 72, 10–12.

Hassenzahl, M., & Tractinsky, N. (2006). User experience—A research agenda. Behaviour and Information Technology, 25(2), 91–97.

Hassenzahl, M., Diefenbach, S., & Göritz, A. (2010). Needs, affect, interactive products—Facets of user experience. Interacting With Computers, 22(5), 353–362.

Hekkert, P. (2006). Design aesthetics: Principles of pleasure in product design. Psychology Science, 48(2), 157–172.

Hekkert, P., Snelders, D., & van Wieringen, P. (2003). Most advanced, yet acceptable: Typicality and novelty as joint predictors of aesthetic preference in industrial design. British Journal of Psychology, 94(1), 111–124.

Henseler, J., Ringe, C., & Sinkovics, R. (2009). The use of partial least squares path modelling in international marketing. In T. Cavusgil, R. Sinkovics, & P. Ghauri (Eds.), New challenges to international marketing. Advances in international marketing (Vol. 20, pp. 277–319). London: Emerald.

Hulland, J., Ryan, M.J., & Rayner, R.K. (2010). Modelling customer satisfaction: A comparative performance evaluation of covariance structure analysis versus partial least squares. In V.E. Vinzi, W. Chin, J. Henseler, & H. Wang (Eds.), Handbook of partial least squares: Concepts, methods and applications in marketing and related fields (pp. 307–325). Berlin: Springer.

Kahneman, D. (2000). New challenges to the rationality assumption. In D. Kahneman & A. Tversky (Eds.), Choices, values and frames (pp. 758–774). New York: Russell Sage Foundation.

Kitajima, M., Blackmon, M.H., & Polson, P.G. (2000). A comprehension-based model of web navigation and its application to web usability analysis. In S. McDonald, Y. Waern, & G. Cockton (Eds.), People and computers XIV—Usability or else! (Proceedings of HCI 2000) (pp. 357–373). Berlin: Springer.

Kohring, M., & Matthes, J. (2007). Trust in news media: Validation of a multidimensional scale. Communication Research, 34(2), 231–252.

Landow, G.P. (2006). Hypertext 3.0: Critical theory and new media in an era of globalization. Baltimore: JHU.

Lavie, T., & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. International Journal of Human-Computer Studies, 60(3), 269–298.

Law, E., & van Schaik, P. (2010). Modelling user experience—An agenda for research and practice. Interacting With Computers, 22(5), 313–322.

Lewis, J.R. (2002). Psychometric evaluation of the PSSUQ using data from five years of usability studies. International Journal of Human-Computer Interaction, 14(3–4), 463–488.

Li, X. (2002). Web page design affects news retrieval efficiency. Newspaper Research Journal, 23(1), 38–49.

Li, Y.-M., & Yeh, Y.-S. (2010). Increasing trust in mobile commerce through design aesthetics. Computers in Human Behaviour, 26(4), 673–684.

Lindgaard, G., Fernandes, G., Dudek, C., & Brown, J. (2006). Attention web designers: You have 50 milliseconds to make a good first impression! Behaviour & Information Technology, 25(2), 115–126.

Magni, M., Taylor, M.S., & Venkatesh, V. (2010). "To play or not to play": A cross-temporal investigation using hedonic and instrumental perspectives to explain user intentions to explore a technology. International Journal of Human-Computer Studies, 68(9), 575–588.

Malik, S. (2012). Kindle ebook sales have overtaken Amazon print sales, says book seller. The Guardian. Retrieved from http://www.guardian.co.uk/books/2012/aug/06/amazon-kindle-ebook-sales-overtake-print?newsfeed=true

Marcoulides, G.A., & Saunders, C. (2006). Editor's comments: PLS: A silver bullet? Mis Quarterly, 30(2), iii–ix.

Martensen, A., & Grønholdt, L. (2003). Improving library users' perceived quality, satisfaction and loyalty: An integrated measurement and management system. Journal of Academic Librarianship, 29(3), 140–147.

DOI: 10.1002/asi

McDonald, S., & Stevenson, R.J. (1998). Effects of text structure and prior knowledge of the learner on navigation in hypertext. Human Factors, 40(1), 18–27.
Mitchell, A., & Rosenstiel, T. (2012). The state of the news media, 2012: An annual report on American journalism. Retrieved from http://stateofthemedia.org
Nebeling, M., Matulic, F., & Norrie, M.C. (2011). Metrics for the evaluation of news site content layout in large-screen contexts. CHI 2011 Proceedings, 1511–1520.
Nielsen, J. (2004). Deceivingly strong information scent costs sales. Retrieved from http://www.useit.com/alertbox/20040802.html
Nielsen, J. (2009). IA task failures remain costly. Retrieved from http://www.useit.com/alertbox/ia-failures.html
Norman, D.A. (2004). Emotional design: Why we love (or hate) everyday things. New York: Basic Books.
O'Brien, H.L. (2011). Exploring user engagement in online news interactions. Proceedings of the American Society for Information Science and Technology, 48(1), 1–10.
O'Brien, H.L., & Lebow, M. (2013). Mixed-methods approach to measuring user experience in online news interactions. Journal of the American Society for Information Science and Technology, 64(8), 1543–1556.
Pavlou, P.A., & Gefen, D. (2004). Building effective online marketplaces with institution-based trust. Information Systems Research, 15(1), 37–59.
Pedhazur, E.J., & Pedhazur-Schmelkin, L. (1991). Measurement, design, and analysis: An integrated approach. London: Erlbaum.
Porat, T., & Tractinsky, N. (2012). It's a pleasure buying here: The effects of web-store design on consumer's emotions and attitudes. Human-Computer Interaction, 27(3), 235–276.
Rosenfeld, L., & Morville, P. (2006). Information architecture for the World Wide Web: Designing large-scale web sites (3rd ed.). Sebastopol, CA: O'Reilly.
Roto, V., Law, E., Vermeeren, A., & Hoonhout, J. (Eds..). (2011). User experience white paper. Results from the Dagstuhl seminar on demarcating user experience, September 15–18, 2010. Retrieved from http://www.allaboutux.org/files/UX-WhitePaper.pdf
Rowley, D., & Rhoades, D. (1992). The cognitive jogthrough: A fast-paced user interface evaluation procedure. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 11, 389–395. Retrieved from http://dl.acm.org/citation.cfm?id=142869&bnc=1
Ryan, M.J., Buzas, T., & Ramaswamy, V. (1995). Making customer satisfaction measurement a power tool. Marketing Research, 7(3), 11–16.
Sela, M., Lavie, T., Inbar, O., Oppenheim, I., & Meyer, J. (2014). Personalizing news content: An experimental study. Journal of the Association for Information Science and Technology, http://onlinelibrary.wiley.com/doi/10.1002/asi.23167/abstract
Sun, H., & Zhang, P. (2008). An exploration of affect factors and their role in user technology acceptance: Mediation and causality. Journal of the American Society for Information Science and Technology, 59(8), 1–12.
Thüring, M., & Mahlke, S. (2007). Usability, aesthetics and emotions in human-technology interaction. International Journal of Psychology, 42(4), 253–264.
Tuch, A.N., Presslaber, E.E., Stöcklin, M., Opwis, K., & Bargas-Avila, J.A. (2012). The role of visual complexity and prototypicality regarding first impression of websites: Working towards understanding aesthetic judgments. International Journal of Human-Computer Studies, 70(11), 794–811.
van Schaik, P., & Ling, J. (2008). Modelling user experience with web sites: Usability, hedonic value, beauty and goodness. Interacting With Computers, 20(3), 419–432.
van Schaik, P., & Ling, J. (2009). The role of context in perceptions of the aesthetics of web pages over time. International Journal of Human-Computer Studies, 67(1), 79–89.
van Schaik, P., & Ling, J. (2011). An integrated model of interaction experience for information retrieval in a Web-based encyclopaedia. Interacting With Computers, 23(1), 18–32.
Venkatesh, V., & Davis, F.D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. Management Science, 46(2), 186–204.
Vilares, M.J., Almeida, M.H., & Coelho, P.S. (2010). Comparison of likelihood and PLS estimators for structural equation modelling: A simulation with customer satisfaction data. In V.E. Vinzi, W. Chin, J. Henseler, & H. Wang (Eds.), Handbook of partial least squares: Concepts, methods and applications in marketing and related fields (pp. 289–305). Berlin: Springer.
Vinzi, V.E., Chin, W.W., Henseler, J., & Wang, H. (Eds.). (2010). Handbook of partial least squares: Concepts, methods and applications in marketing and related fields. Berlin: Springer.
Watson, D., Clark, L.A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. Journal of Personality and Social Psychology, 54(6), 1063–1070.
Wirtz, D., Kruger, J., Scollon, C.N., & Diener, E. (2003). What to do on spring break? The role of predicted, on-line, and remembered experience in future choice. Psychological Science, 14(5), 520–524.
Yang, Z., Cai, S., Zhou, Z., & Zhou, N. (2005). Development and validation of an instrument to measure user perceived service quality of information presenting Web portals. Information & Management, 42(4), 575–589.
Zhou, H., & Fu, X. (2007). Understanding, measuring, and designing user experience: The causal relationship between the aesthetic quality of products and user affect. In J.A. Jacko (Ed.), Human-computer interaction. Interaction design and usability (pp. 340–349). Berlin: Springer.

Appendix A

Psychometric Scales Used in the Study

*AttrakDiff2-SF*I judge the website to be . . .

PQ1	Confusing	—	Structured
PQ2	Unpredictable	—	Predictable
PQ3	Impractical	—	Practical
PQ4	Complicated	—	Simple
HQ1	Dull	—	Captivating
HQ2	Tacky	—	Stylish
HQ3	Cheap	—	Premium
HQ4	Unimaginative	—	Creative
BEAUTY	Ugly	—	Beautiful
GOODNESS	Bad	—	Good

Note. Response format: 7-point semantic differential.

Positive and Negative Affect Schedule (PANAS)

During the use of the site I felt . . .

POS1	Interested	NEG6	Irritable
NEG1	Distressed	POS6	Alert
POS2	Excited	NEG7	Ashamed
NEG2	Upset	POS7	Inspired
POS3	Strong	NEG8	Nervous
NEG3	Guilty	POS8	Determined
NEG4	Scared	POS9	Attentive
NEG5	Hostile	NEG9	Jittery
POS4	Enthusiastic	POS10	Active
POS5	Proud	NEG10	Afraid

Note. Items presented mixed.

Perceived Enjoyment (PE)

PE1	I find using this news site to be enjoyable.
PE2	The actual process of using this news site is pleasant.
PE3	I have fun using this news site.

DOI: 10.1002/asi

Perceived Disorientation (PD)

During the use of the site . . .

PD1	I felt lost.
PD2	I felt I was going around in circles.
PD3	It was difficult to find a page I had previously viewed.
PD4	Navigating between the pages was a problem.
PD5	I didn't know how to get to my desired location.
PD6	I felt disoriented.
PD7	After browsing for a while I had no idea where to go next.

Note. Anchor points: *never* to the left and *always* to the right.

Perceived Aesthetics (AES)

CA1	Clean
CA2	Pleasant
CA3	Symmetrical
CA4	Aesthetic
EA1	Original
EA2	Sophisticated
EA3	Spectacular
EA4	Creative

Note. Anchor points *strongly disagree* and *strongly agree*. CA, classical aesthetics; EA, expressive aesthetics. Items presented mixed.

Usefulness of Content (UC) and Adequacy of Information (AI)

UC1	The site provides relevant information.
UC2	The site provides up-to-date information.
UC3	The site provides unique content.
AI1	The site provides comprehensive information.
AI2	The site provides complete content.
AI3	The site provides sufficient information.

Note. Anchor points *strongly disagree*, *neutral*, and *strongly agree*.

Trust

TRU1	I believe the site provides truthful information.
TRU2	I trust the information presented on the site.
TRU3	The site provides reliable information.
TRU4	I trust the competence of journalists working for the site.

Note. Anchor points *strongly disagree*, *neutral*, and *strongly agree*.

Behavioral Intention (BI)

BI1	I intend to use the site in the future.
BI2	I predict that I will use the site in the future.

Note. Anchor points *strongly disagree* and *strongly agree*.

Satisfaction

SAT1	Considering all your experience with the site, how satisfied are you in general?
SAT2	To what degree do you consider that the site fulfils your expectations?
SAT3	Imagine a news site that is perfect in all aspects. How close to this ideal do you consider this site to be?

Note. Anchor points: SAT1 *completely dissatisfied—completely satisfied*; SAT2 *much less than expected—much more than expected*; SAT3 *very far away—very close*.

One item was presented at a time from each scale during the study with instructions presented above each item. Except for AttrakDiff2-SF, each scale was measured using a 7-point Likert scale.

Appendix B

Details of Analysis of Measurement Model

TABLE B1. Coefficients of reliability and convergent validity.

Construct/indicator	Average variance extracted	Composite reliability	Loading	Standard error	t^a
Perceived aesthetics	0.58	0.92			
—CA1			*0.69	0.10	7.13
—CA2			0.83	0.03	27.50
—CA3			*0.47	0.12	4.04
—CA4			0.79	0.05	15.57
—EA1			0.78	0.06	13.83
—EA2			0.81	0.05	16.57
—EA3			0.82	0.04	21.24
—EA4			0.84	0.04	22.53
Perceived disorientation	0.57	0.90			
—PD1			0.81	0.06	14.08
—PD2			0.73	0.09	8.49
—PD3			*0.67	0.09	7.25
—PD4			0.71	0.08	9.12
—PD5			0.82	0.04	20.38
—PD6			0.82	0.05	15.48
—PD7			0.72	0.06	12.98
Adequacy of information	0.81	0.93			
—AI1			0.90	0.03	34.72
—AI2			0.90	0.02	38.41
—AI3			0.89	0.03	26.02

DOI: 10.1002/asi

Construct/indicator	Average variance extracted	Composite reliability	Loading	Standard error	t^a
Usefulness of content	0.62	0.83			
—UC1			0.72	0.11	6.68
—UC2			0.87	0.05	16.23
—UC3			0.78	0.04	18.53
Trust	0.75	0.92			
—TRU1			0.80	0.08	9.52
—TRU2			0.91	0.05	19.57
—TRU3			0.91	0.03	35.88
—TRU4			0.84	0.05	18.55
Perceived enjoyment	0.86	0.95			
—PE1			0.95	0.01	84.07
—PE2			0.96	0.01	110.33
—PE3			0.88	0.05	16.19
Behavioral intention	0.98	0.99			
—BI1			0.99	0.00	239.93
—BI2			0.99	0.00	209.25
Satisfaction	0.84	0.94			
—SAT1			0.94	0.01	65.25
—SAT2			0.90	0.03	32.47
—SAT3			0.91	0.02	45.97
Pragmatic quality T1	0.53	0.80			
—PQ1_1			0.82	0.04	18.33
—PQ2_1			*0.23	0.22	1.07
—PQ3_1			0.86	0.06	15.01
—PQ4_1			0.81	0.05	17.81
Pragmatic quality T2	0.58	0.84			
—PQ1_2			0.88	0.03	30.67
—PQ2_2			*0.50	0.15	3.29
—PQ3_2			0.82	0.09	9.28
—PQ4_2			0.79	0.06	14.19
Construct/indicator	Average variance extracted	Composite reliability	Loading	Standard error	t^a
Hedonic qualityT1	0.63	0.87			
—HQ1_1			0.80	0.05	16.90
—HQ2_1			0.77	0.06	12.71
—HQ3_1			0.74	0.10	7.35
—HQ4_1			0.85	0.03	27.28
Hedonic qualityT2	0.69	0.90			
—HQ1_2			0.84	0.04	19.24
—HQ2_2			0.86	0.03	28.64
—HQ3_2			0.71	0.11	6.50
—HQ4_2			0.91	0.02	41.73
Positive effect	0.55	0.92			
—POS1			*0.69	0.05	13.33
—POS2			0.86	0.03	33.55
—POS3			0.74	0.05	13.53
—POS4			0.78	0.07	10.86
—POS5			0.78	0.05	14.83
—POS6			*0.57	0.09	6.52
—POS7			0.79	0.05	16.31
—POS8			0.78	0.06	13.14
—POS9			*0.63	0.09	7.07
—POS10			0.75	0.06	13.33
Negative effect	0.38	0.85			
—NEG1			*0.50	0.15	3.26
—NEG2			*0.52	0.18	2.95
—NEG3			*0.47	0.19	2.50
—NEG4			*0.69	0.19	3.66
—NEG5			*0.63	0.18	3.59
—NEG6			0.73	0.20	3.65
—NEG7			0.73	0.22	3.39
—NEG8			*0.69	0.18	3.76
—NEG9			*0.42	0.24	1.77
—NEG10			*0.65	0.21	3.06

Note. *Loading <.70.

[a]Bootstrap, $n = 5{,}000$.

DOI: 10.1002/asi

TABLE B2. Coefficients of discriminant validity of retained scales.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1 Perceived aesthetics	0.76																
2 Adequacy of information	**.59	0.90															
3 Beauty T1 (before use)	**.63	**.32	1.00														
4 Beauty T2 (after use)	**.68	**.34	**.60	1.00													
5 Behavioral intention	**.52	**.35	**.38	**.47	0.99												
6 Goodness T1	**.70	**.53	**.68	**.54	**.52	1.00											
7 Goodness T2	**.76	**.49	**.54	**.68	**.69	**.70	1.00										
8 Hedonic quality T1	**.63	**.50	**.75	**.50	**.45	**.72	**.59	0.79									
9 Hedonic quality T2	**.87	**.51	**.63	**.73	**.54	**.65	**.81	**.72	0.83								
10 Perceived disorientation	**–.51	**–.38	**–.41	**–.49	**–.39	**–.48	**–.52	**–.32	**–.42	0.76							
11 Perceived enjoyment	**.67	**.48	**.53	**.62	**.53	**.65	**.80	**.64	**.76	**–.53	0.93						
12 Positive affect	**.64	**.46	**.45	**.64	**.48	**.49	**.63	**.52	**.67	**–.33	**.75	0.74					
13 Pragmatic quality T1	**.46	**.50	**.40	*.26	**.31	**.63	**.46	**.43	**.34	**–.35	**.37	.22	0.83				
14 Pragmatic quality T2	**.65	**.45	**.49	**.59	**.46	**.62	**.66	**.48	**.63	**–.70	**.71	**.54	**.50	0.84			
15 Satisfaction	**.73	**.64	**.52	**.57	**.67	**.76	**.79	**.62	**.71	**–.58	**.74	**.61	**.52	**.69	0.92		
16 Trust	**.22	**.53	.17	.05	**.33	*.24	.10	**.30	.12	–.02	.08	.13	.19	.07	**.34	0.87	
17 Usefulness of content	**.56	**.60	**.35	**.37	**.55	**.48	**.59	**.44	**.53	**–.34	**.51	**.58	**.33	**.44	**.67	**.34	0.79

Note. Diagonal elements are square root of average variance extracted. Off-diagonal elements are correlation coefficients.
*$p < .05$, **$p < .01$.

TABLE B3. Loadings and cross-loadings of indicators on latent variables.

Item	AES	AI	B1	B2	BI	G1	G2	HQ1	HQ2	NEG	PD	PE	POS	PQ1	PQ2	SAT	TRU	UC
AI1	0.47	0.90	0.25	0.27	0.29	0.49	0.39	0.47	0.44	–0.08	–0.27	0.43	0.43	0.47	0.37	0.55	0.49	0.60
AI2	0.60	0.90	0.33	0.33	0.35	0.52	0.46	0.46	0.47	–0.17	–0.41	0.44	0.41	0.44	0.50	0.63	0.49	0.49
AI3	0.51	0.89	0.28	0.31	0.30	0.41	0.47	0.42	0.45	–0.15	–0.33	0.43	0.39	0.44	0.29	0.54	0.43	0.52
B1	0.63	0.32	1.00	0.60	0.38	0.68	0.54	0.75	0.63	–0.28	–0.41	0.53	0.45	0.40	0.48	0.52	0.16	0.35
B2	0.68	0.34	0.60	1.00	0.47	0.54	0.68	0.50	0.73	–0.33	–0.49	0.62	0.64	0.25	0.57	0.57	0.05	0.37
BI1	0.53	0.36	0.38	0.48	0.99	0.51	0.69	0.44	0.55	–0.30	–0.39	0.54	0.49	0.29	0.46	0.66	0.33	0.56
BI2	0.49	0.34	0.36	0.46	0.99	0.52	0.68	0.44	0.53	–0.31	–0.38	0.52	0.45	0.32	0.43	0.66	0.34	0.53
CA1	0.69	0.39	0.40	0.36	0.39	0.58	0.53	0.37	0.51	–0.18	–0.30	0.42	0.36	0.33	0.48	0.55	0.15	0.38
CA2	0.83	0.48	0.52	0.52	0.58	0.67	0.79	0.54	0.75	–0.31	–0.54	0.69	0.59	0.43	0.66	0.76	0.14	0.60
CA3	**0.47**	0.23	0.23	0.21	0.17	0.31	0.17	0.22	0.28	0.00	–0.17	0.08	0.21	0.21	0.30	0.30	0.14	0.34
CA4	0.79	0.54	0.48	0.50	0.40	0.48	0.61	0.43	0.69	–0.28	–0.47	0.48	0.38	0.38	0.54	0.54	0.15	0.41
DIS1	–0.45	–0.25	–0.46	–0.45	–0.31	–0.45	–0.49	–0.25	–0.37	0.25	0.81	–0.49	–0.32	–0.38	–0.67	–0.46	0.05	–0.27
DIS2	–0.29	–0.23	–0.15	–0.22	–0.33	–0.34	–0.39	–0.18	–0.28	0.14	0.73	–0.44	–0.24	–0.22	–0.43	–0.44	0.05	–0.24
DIS3	–0.27	–0.14	–0.25	–0.25	–0.14	–0.33	–0.23	–0.19	–0.17	0.14	0.67	–0.29	–0.12	–0.23	–0.41	–0.30	–0.07	–0.08
DIS4	–0.38	–0.22	–0.31	–0.20	–0.12	–0.31	–0.22	–0.21	–0.24	0.23	0.71	–0.26	–0.07	–0.16	–0.44	–0.26	–0.01	–0.11
DIS5	–0.42	–0.35	–0.41	–0.45	–0.38	–0.41	–0.45	–0.40	–0.35	0.34	0.82	–0.42	–0.33	–0.30	–0.59	–0.54	–0.12	–0.37
DIS6	–0.45	–0.31	–0.24	–0.47	–0.40	–0.30	–0.46	–0.17	–0.38	0.37	0.82	–0.44	–0.31	–0.20	–0.60	–0.52	0.05	–0.37
DIS7	–0.41	–0.42	–0.30	–0.42	–0.29	–0.37	–0.40	–0.24	–0.35	0.30	0.72	–0.37	–0.25	–0.33	–0.53	–0.45	–0.03	–0.24
EA1	0.78	0.47	0.52	0.57	0.40	0.57	0.56	0.53	0.66	–0.05	–0.43	0.53	0.58	0.41	0.44	0.54	0.21	0.37
EA2	0.81	0.53	0.48	0.55	0.38	0.49	0.60	0.50	0.73	–0.22	–0.35	0.44	0.45	0.37	0.41	0.54	0.22	0.45
EA3	0.82	0.49	0.53	0.62	0.29	0.51	0.52	0.57	0.74	–0.06	–0.30	0.55	0.56	0.30	0.49	0.52	0.15	0.36
EA4	0.84	0.41	0.60	0.71	0.45	0.56	0.64	0.60	0.82	–0.22	–0.47	0.64	0.64	0.31	0.52	0.62	0.16	0.44
G1	0.69	0.53	0.68	0.54	0.52	1.00	0.70	0.72	0.65	–0.19	–0.48	0.65	0.49	0.63	0.60	0.76	0.24	0.47
G2	0.76	0.49	0.54	0.68	0.69	0.70	1.00	0.59	0.81	–0.36	–0.52	0.80	0.63	0.46	0.66	0.79	0.10	0.59
HQ1_1	0.42	0.41	0.51	0.40	0.41	0.53	0.47	0.80	0.56	–0.21	–0.21	0.48	0.44	0.31	0.29	0.47	0.20	0.46
HQ1_2	0.72	0.44	0.50	0.55	0.48	0.49	0.74	0.54	0.84	–0.21	–0.44	0.74	0.58	0.29	0.58	0.61	0.03	0.45
HQ2_1	0.40	0.36	0.48	0.35	0.29	0.54	0.40	0.77	0.50	–0.26	–0.17	0.51	0.35	0.31	0.42	0.37	0.23	0.27
HQ2_2	0.74	0.41	0.61	0.59	0.48	0.65	0.67	0.70	0.86	–0.33	–0.27	0.57	0.52	0.35	0.52	0.57	0.15	0.48
HQ3_1	0.47	0.39	0.52	0.26	0.32	0.52	0.44	0.74	0.54	–0.22	–0.26	0.44	0.30	0.29	0.35	0.54	0.28	0.30
HQ3_2	0.65	0.46	0.29	0.54	0.34	0.40	0.51	0.43	0.71	–0.14	–0.24	0.47	0.47	0.22	0.37	0.53	0.13	0.36
HQ4_1	0.67	0.42	0.82	0.54	0.39	0.68	0.55	0.85	0.66	–0.16	–0.34	0.59	0.51	0.44	0.40	0.56	0.23	0.36
HQ4_2	0.79	0.39	0.66	0.74	0.50	0.59	0.75	0.70	0.91	–0.28	–0.43	0.73	0.66	0.27	0.56	0.64	0.08	0.44

DOI: 10.1002/asi

Item	AES	AI	B1	B2	BI	G1	G2	HQ1	HQ2	NEG	PD	PE	POS	PQ1	PQ2	SAT	TRU	UC
NEG1	–0.16	–0.16	–0.06	–0.05	–0.07	0.01	–0.06	–0.05	–0.07	0.50	0.26	–0.07	0.00	–0.03	–0.23	–0.11	–0.06	–0.13
NEG10	–0.09	–0.03	–0.06	–0.14	–0.17	–0.12	–0.17	–0.06	–0.15	0.65	0.14	–0.13	–0.03	–0.19	–0.26	–0.16	0.01	–0.14
NEG2	–0.21	0.00	–0.19	–0.27	–0.25	–0.12	–0.26	–0.08	–0.21	0.52	0.10	–0.09	0.00	–0.01	–0.10	–0.18	–0.05	–0.08
NEG3	–0.05	–0.06	–0.03	–0.16	–0.04	–0.04	0.01	–0.08	–0.04	**0.47**	0.11	0.09	0.17	–0.12	–0.04	0.00	–0.05	–0.03
NEG4	–0.08	0.05	–0.14	–0.17	–0.11	–0.01	–0.14	–0.13	–0.11	0.69	0.08	0.02	0.10	–0.01	0.01	–0.02	–0.01	–0.08
NEG5	–0.11	–0.17	–0.02	–0.14	–0.23	–0.10	–0.28	–0.14	–0.13	0.63	0.12	–0.11	0.09	–0.13	–0.08	–0.11	–0.05	–0.09
NEG6	–0.30	–0.21	–0.42	–0.41	–0.32	–0.26	–0.40	–0.40	–0.40	0.73	0.42	–0.39	–0.16	–0.13	–0.29	–0.34	–0.07	–0.30
NEG7	–0.01	–0.02	–0.07	–0.15	–0.15	–0.07	–0.18	–0.07	–0.09	0.73	0.17	–0.07	0.05	–0.18	–0.16	–0.10	0.09	–0.25
NEG8	–0.12	–0.06	–0.20	–0.22	–0.15	–0.11	–0.22	–0.15	–0.16	0.69	0.27	–0.15	–0.06	–0.13	–0.21	–0.15	–0.05	–0.11
NEG9	0.08	–0.08	0.00	0.21	–0.01	–0.11	–0.03	–0.03	0.07	**0.42**	0.14	–0.02	0.23	–0.22	–0.08	–0.07	–0.09	–0.06
PE1	0.69	0.50	0.57	0.65	0.60	0.66	0.81	0.67	0.78	–0.30	–0.53	0.95	0.71	0.40	0.70	0.77	0.13	0.49
PE2	0.65	0.49	0.51	0.56	0.46	0.64	0.76	0.61	0.71	–0.23	–0.54	0.96	0.70	0.39	0.73	0.71	0.06	0.49
PE3	0.49	0.33	0.38	0.50	0.40	0.48	0.62	0.48	0.62	–0.13	–0.36	0.88	0.68	0.22	0.47	0.57	0.00	0.41
POS1	0.56	0.42	0.50	0.54	0.60	0.54	**0.73**	0.57	0.67	–0.26	–0.47	**0.81**	0.69	0.33	0.60	0.67	0.12	0.60
POS10	0.42	0.19	0.27	0.49	0.34	0.29	0.41	0.19	0.34	0.08	–0.26	0.41	0.75	0.13	0.30	0.43	0.02	0.38
POS2	0.57	0.33	0.52	0.61	0.42	0.48	0.54	0.58	0.62	–0.05	–0.20	0.69	0.86	0.16	0.45	0.49	0.10	0.41
POS3	0.46	0.18	0.39	0.57	0.26	0.29	0.39	0.38	0.47	0.15	–0.15	0.49	0.74	0.05	0.30	0.32	–0.01	0.23
POS4	0.40	0.41	0.27	0.45	0.33	0.38	0.46	0.42	0.51	–0.11	–0.31	0.66	0.78	0.11	0.42	0.48	0.10	0.48
POS5	0.57	0.36	0.35	0.48	0.31	0.37	0.46	0.36	0.52	0.13	–0.20	0.49	0.78	0.20	0.35	0.42	0.12	0.39
POS6	0.23	0.38	0.12	0.21	0.18	0.27	0.27	0.23	0.25	0.17	–0.18	0.38	0.57	0.23	0.24	0.29	0.08	0.30
POS7	0.55	0.37	0.26	0.53	0.32	0.31	0.44	0.35	0.56	–0.06	–0.23	0.53	0.79	0.13	0.34	0.47	0.13	0.45
POS8	0.46	0.23	0.22	0.43	0.27	0.18	0.34	0.23	0.44	0.14	–0.11	0.43	0.78	–0.03	0.25	0.32	0.01	0.42
POS9	0.40	0.44	0.24	0.30	0.28	0.32	0.37	0.26	0.36	–0.09	–0.19	0.42	0.63	0.25	0.34	0.45	0.25	0.44
PQ1_1	0.37	0.37	0.41	0.19	0.27	0.56	0.38	0.44	0.27	–0.20	–0.30	0.25	0.13	0.82	0.42	0.44	0.14	0.32
PQ1_2	0.56	0.38	0.53	0.53	0.42	0.60	0.62	0.51	0.58	–0.23	–0.62	0.65	0.50	0.46	0.88	0.62	0.06	0.42
PQ2_1	0.05	0.02	–0.01	–0.09	0.13	0.05	–0.02	–0.12	–0.10	–0.05	–0.14	–0.09	–0.06	**0.23**	0.08	0.03	0.10	–0.03
PQ2_2	0.25	0.07	0.14	0.11	0.11	0.11	0.24	0.08	0.19	–0.21	–0.35	0.19	0.06	0.21	0.50	0.21	–0.04	0.08
PQ3_1	0.38	0.47	0.27	0.27	0.30	0.50	0.42	0.28	0.30	–0.17	–0.30	0.32	0.22	0.86	0.39	0.49	0.19	0.31
PQ3_2	0.51	0.41	0.30	0.53	0.47	0.37	0.57	0.30	0.51	–0.41	–0.63	0.57	0.45	0.35	0.82	0.58	0.14	0.42
PQ4_1	0.39	0.41	0.32	0.18	0.19	0.52	0.34	0.36	0.28	–0.07	–0.27	0.37	0.20	0.81	0.43	0.36	0.14	0.19
PQ4_2	0.56	0.32	0.40	0.42	0.26	0.58	0.47	0.38	0.48	–0.02	–0.52	0.56	0.39	0.45	0.79	0.52	–0.03	0.25
SAT1	0.67	0.57	0.50	0.53	0.70	0.73	0.81	0.58	0.65	–0.26	–0.55	0.72	0.58	0.49	0.62	0.94	0.30	0.70
SAT2	0.66	0.47	0.45	0.55	0.57	0.66	0.69	0.51	0.64	–0.24	–0.50	0.66	0.59	0.46	0.63	0.90	0.17	0.56
SAT3	0.69	0.70	0.47	0.49	0.57	0.68	0.67	0.61	0.65	–0.24	–0.54	0.67	0.52	0.47	0.61	0.91	0.46	0.57
TRU1	0.18	0.48	0.19	–0.02	0.24	0.19	0.07	0.25	0.08	–0.03	–0.08	0.01	0.08	0.10	0.09	0.26	0.80	0.32
TRU2	0.12	0.40	0.05	–0.04	0.33	0.21	0.06	0.17	0.00	–0.05	–0.01	–0.02	0.07	0.18	–0.01	0.29	0.91	0.30
TRU3	0.13	0.43	0.08	0.06	0.32	0.15	0.04	0.22	0.06	–0.04	0.02	0.05	0.11	0.06	–0.01	0.27	0.91	0.29
TRU4	0.29	0.50	0.23	0.14	0.27	0.27	0.14	0.35	0.21	–0.06	0.02	0.19	0.17	0.29	0.12	0.34	0.84	0.27
UC1	0.32	0.43	0.19	0.17	0.42	0.31	0.31	0.26	0.27	–0.22	–0.28	0.21	0.24	0.22	0.25	0.53	0.33	0.72
UC2	0.39	0.47	0.22	0.23	0.46	0.28	0.42	0.28	0.36	–0.24	–0.28	0.30	0.37	0.23	0.30	0.50	0.36	0.87
UC3	0.55	0.50	0.38	0.42	0.42	0.49	0.59	0.46	0.55	–0.14	–0.26	0.60	0.65	0.31	0.43	0.55	0.16	0.78

Note. Gray background indicates the loadings of items belonging to a particular scale. Bold numbers indicate loadings smaller than .50 and cross-loadings exceeding loadings. AES, aesthetics; AI, adequacy of information; B1, beauty before use; B2, beauty after use; BI, behavioral intention; G1, goodness before use; G2, goodness after use; HQ1, hedonic quality before use; HQ2, hedonic quality after use; NEG, negative affect; PD, perceived disorientation; PE, perceived enjoyment; POS, positive affect; PQ1, pragmatic quality before use; PQ2, pragmatic quality after use; SAT, satisfaction; TRU, trust; UC, usefulness of content; CA, classical aesthetics; EA, expressive aesthetics.

DOI: 10.1002/asi