2004

A Confirmatory Factor Analysis of Two Web Site Usability Instruments

David T. Green
Southern Illinois University Carbondale, dtgreen@siu.edu

J Michael Pearson
Southern Illinois University Carbondale, jpearson@cba.siu.edu

A Confirmatory Factor Analysis of Two Web Site Usability Instruments

David T. Green
Southern Illinois University Carbondale
dtgreen@siu.edu

J. Michael Pearson
Southern Illinois University Carbondale
jpearson@cba.siu.edu

ABSTRACT

Many perspectives of user acceptance of Web sites have been examined, yet information systems research often overlooks the human-computer interaction aspects, particularly in the area of Web site usability. Web site usability has recently gained greater acceptance in information literature through the development of instruments by Palmer (2002) and Agarwal and Venkatesh (2002). This study conducted a confirmatory factor analysis of both instruments in an attempt to validate the two instruments. Our results found that the Palmer instrument exhibited satisfactory measurement properties, although allowing room for further refinement. The Agarwal and Venkatesh instrument, although useful as a practical metric, displayed poor validity for the underlying constructs that compose Web site usability. Validation of these instruments furthers their scope and potential use by researchers and practitioners in helping them better understand the capabilities of their Web sites, while providing a foundation for further refinement of the Web site usability construct.

Keywords

Web site usability, confirmatory factor analysis

INTRODUCTION

Developing a standardized instrument is important for research to accumulate in a topic area as direct comparisons become possible between individuals, time periods, industries, cultures, or geographic regions (Cook and Campbell, 1979). The research cycle for developing a standardized instrument involves two steps: (1) exploratory studies that put forward hypothesized measurement model(s) via the analysis of empirical data from a referent population; and (2) confirmatory studies that test the hypothesized measurement model(s) against new data gathered from the same referent population. Both Palmer (2002) and Agarwal and Venkatesh (2002) developed instruments to measure Web site usability. Their work completed the first exploratory step in the instrument development cycle. The current research completes the second step by gathering new data to test the validity and reliability of both the Palmer and Agarwal and Venkatesh instruments.

This paper continues the discussion of the Web site usability literature and instruments. The following sections detail two recent attempts at measuring Web site usability and define the research model. The paper concludes by detailing the study's methodology and results of the confirmatory factor analyses, while section six provides discussion of the findings, implications for future research and conclusions.

LITERATURE REVIEW

Web Site Usability

Although the HCI literature has examined a few areas of Web site usability, it has only been recently that information systems literature addressed the need to examine Web site usability in understanding success in e-commerce. The development of metrics to measure Web site usability allows managers to make comparisons, benchmark performance, and plan improvements. Usability is increasingly important as organizations compete in the always changing, hypercompetitive online markets (Bogner and Barr, 2000, Zohar and Morgan, 1996). Therefore, having valid and reliable instruments to measure Web site usability is important for both business and research.

RESEARCH MODEL

The current study tested Web site usability underlying first-order factor structure as set forth by Palmer (2002) and Agarwal and Venkatesh (2002).

Palmer Instrument

Palmer (2002) developed a measure of Web site usability that found significant associations between Web site design and Web site performance indicating that the constructs demonstrate good nomological validity. Palmer's study found that his findings provide a set of measures with acceptable validity and reliability. Web site success was found to be significantly associated with Web site download delay (speed of access and display rate within the Web site), navigation (organization, arrangement, layout, and sequencing), content (amount and variety of product information), interactivity (customization and interactivity), and responsiveness (feedback options and FAQs). As underlying constructs

that are hypothesized as measures of Web site usability, a correlated first-order factor structure is proposed in accordance with the original instrument (Figure 1- *see full paper*).

Five factors are included in the Palmer instrument. Download delay is the initial request for access to the page and then each subsequent request for changing pages with the site (Rose & Straub, 1998). Web site designers can choose not to include slow loading elements, thus length of wait is important. Organization and navigation is important to outcomes (Nielsen, 2000). Navigability is a measure that is defined as the sequencing of pages, well organized layout, and consistency of navigation protocols (Palmer, 2002). Content includes the amount and variety of content as well as the use of text, graphics, and multimedia. The manipulation and utilization of information provided by Web sites to the user is strongly influenced by interface design. Interactivity includes the ability to customize the site's look, feel, and content as well as provide interaction with the user (Palmer, 2002). Berry (1999) suggests that some of the most valued metrics revolve around customer reaction to the site. Responsiveness is defined as the presence of feedback to users and the availability of response from the site managers (Palmer, 2002).

Agarwal and Venkatesh Instrument

Agarwal & Venkatesh (2002) presented categories and subcategories comprising the Microsoft Usability guidelines, while developing an instrument that operationalizes Web site usability. Their findings suggested that the evaluation procedure, the instrument, as well as the usability metric exhibit good properties. A correlated four-factor model of Web site usability is hypothesized (Figure 2- *see full paper*).

Ease of use is the first of the first of four factor included in the study. Ease of use describes the cognitive effort required in using a Web site. MIS research has widely accepted this measure as an antecedent to various technology acceptance outcomes (Davis, Bagozzi, & Warshaw, 1989). Ease of use is conceptualized through the subcategories of goals, structure, and feedback (Agarwal & Venkatesh, 2002). Made-for-the-medium relates to tailoring a Web site to fit a particular user's needs. Peppers and Rogers (1999) suggest contemporary marketing strategies require Web sites with dynamic content tailored to specific user needs. The three subcategories of made-for-the-medium are community, personalization, and refinement (Agarwal & Venkatesh, 2002).

Emotion taps into affective reactions invoked by a Web site. Agarwal and Venkatesh (2002) note that affective responses have shown to be important in computer use situations (Agarwal & Karahanna, 2000; Vanketesh, 2000; Venkatesh & Speier, 2000; Webster & Ho, 1997). Four subcategories compose the emotion construct: challenge, plot, character strength, and pace. Content assesses the informational and transactional capabilities of a Web site (Agarwal & Venkatesh). The content construct is similar to the technology acceptance constructs of perceived usefulness construct (Davis et al., 1989) and relative advantage (Agarwal & Prasad, 1997; Moore & Benbasat, 1991). The four subcategories (questions) of content are relevance, media use, depth and breadth, and current and timely information.

RESEARCH METHODS

The research instruments of interest have followed the scientific research cycle as proposed by Mackenzie and House (1979) and McGrath (1979), by conducting exploratory research. Confirmatory research and conceptual refinements are the next step in the research cycle.

CONFIRMATORY SAMPLE

Although Web site performance is important to several target populations, the current study focused on the online consumer in accordance with the Palmer (2002) and Agarwal and Venkatesh (2002) studies. The study's participants were 159 undergraduate business students from a large mid-western university. University students are considered an appropriate surrogate in consumer research that examines products the students are likely to purchase on a regular basis (Lynch, Calder, Philips, & Tybout, 1982), thus music and bookstore sites were chosen due to their ability to provide appropriate external validity.

In the study, participants were randomly assigned to one of eight potential scenarios. The scenario involved the process of searching, selecting, and inquiring about a product of their choice that was available from the selected Web retailer. The potential Web retailers included four bookstores and four music stores. The book and music sectors of e-commerce were chosen due to their applicability for college students, providing them with a realistic buying scenario (Figure 3- *see full paper)*.

DATA ANALYSIS

Confirmatory Factor Analysis

LISREL 8.52 was used to conduct a confirmatory factor analysis of the Web site usability intstruments. Confirmatory factor analysis (CFA) is appropriate in this study due to the Web site usability instruments' empirical foundation, allowing the researcher to specify the exact factor model in advance as opposed to exploratory factor analysis which seeks an undetermined structure for a set of variables (Stevens, 2002). CFA has high demands for sample size. The 159 participants is a satisfactory sample size when conducting structural equation models (Anderson & Gerbing, 1988). Doll, Xia, and Torkzadeh (1994) note the lack of one universally accepted index of model adequacy, therefore there was an expectation that convergence upon multiple indexes would define an appropriate model. In other words, a good instrument

should exhibit adequate numbers in the results of multiple indexes, not just one or a few, for the model to be considered an adequate one.

Validity and Reliability

Factor loadings are viewed as regression coefficients in confirmatory factor analysis. The larger the factor loadings or coefficients as compared with their standard errors and expressed by the corresponding t values, the stronger the evidence that the measured variables represent the underlying constructs (Bollen, 1989; Mueller, 1994). The reliability of items, factors, and the overall instrument can also be evaluated with confirmatory factor analysis. The proportion of variance in the observed variables that is accounted for by the latent variables that influence them is used to estimate the reliability of the items (Bollen, 1989; Mueller, 1994). The overall reliability of the instrument is found in the total coefficient of determination for observed variables.

RESULTS

The LISREL estimates for the measurement model were examined and found not to have any offending estimates (a loading greater than 1.0). The overall model was first assessed for fitness, which is the degree to which the specified indicators represent the hypothesized constructs. Three types of overall model fit measures were examined in the LISREL output: absolute, incremental, and parsimonious fit measures (*Table 1*).

Fit measure	Recommended Value	Palmer (2002) Result	Agarwal & Venkatesh (2002) Result
χ^2	p>0.05	0.00011	0.00000
χ^2/df	<5.00	.000001	0
GFI	>0.90	0.90	0.75
NFI	>0.90	0.96	0.91
AGFI	>0.90	0.85	0.67
RMSR	<0.05	0.047	0.098

Table 1. CFA Results

Palmer CFA Results

The likelihood ratio chi-square statistic for the Palmer (2002) instrument (χ^2= 118.21, 67 d.f.) has a statistical significance level of p < .00011, which is above the minimum level of .05 but not above the more conservative levels of .10 or .20, showing some support for believing that the difference of the predicted and actual matrices are significant, indicating unacceptable fit. Hair, Tatham, Anderson, and Black (2000) suggest that the potential exists for the χ^2 with sample sizes of 100 or less to denote no differences even when the model has no significant relationships, thus additional measures of fit should be employed. For the Palmer instrument, the goodness-of-fit index (GFI) has a value of .90, which is high but not adjusted for model parsimony. Another measure, the root mean square residual (RMSR), indicates that the average residual correlation is .047, which is acceptable given the fairly strong correlations in the original correlation matrix (Byrne, 1989; Joreskog and Sorbom, 1984).

With absolute measures within acceptable levels, the incremental fit and parsimonious fit indices are needed to help make sure the model is acceptable from other perspectives. The next goodness-of-fit measure assesses the incremental fit of the model compared to a null model. The Normed Fit Index (NFI) yields .96 for the Palmer instrument surpassing the recommended level of .90, further supporting the proposed model.

Parsimonious fit measures evaluate the fit of the model versus the number of estimated coefficients needed to achieve that level of fit. The adjusted goodness-of-fit index (AGFI) is .85 for the Palmer instrument and is fairly close to the recommended level of .90, thus marginal acceptance can be given on this measure. The normed chi-square (χ^2/df) for the Palmer instrument has a value of 1.76 , falling within the recommended levels of 1.0 to 2.0. Both parsimonious fit measures provide support to be given for the Palmer instrument's model parsimony.

Agarwal and Venkatesh CFA Results

The Agarwal and Venkatesh (2002) instrument (χ^2= 499.92, 146 d.f) results also demonstrate a lack of fit through the significant differences between the predicted and actual matrices. In contrast to the Palmer model, the Agarwal and Venkatesh instrument has a GFI of .75 and RMSR of .098, a weaker demonstration of fit as compared to the Palmer instrument.

The Normed Fit Index (NFI) yields .96 for the Palmer instrument and .91 for the Agarwal and Venkatesh instrument, both surpassing the recommended level of .90, further supporting the proposed model. The Normed Fit Index (NFI) was .91 for the Agarwal and Venkatesh instrument, also surpassing the recommended level of .90, supporting the proposed model. In contrast, the Agarwal and Venkatesh instrument has an AGFI of .67, far below the recommended level of .90 for parsimonious fit. while the Agarwal and Venkatesh instrument is 3.42, outside the recommended region for acceptable parsimonious fit. The parsimonious fit measures provide a lack of support for the Agarwal and Venkatesh instrument's model parsimony.

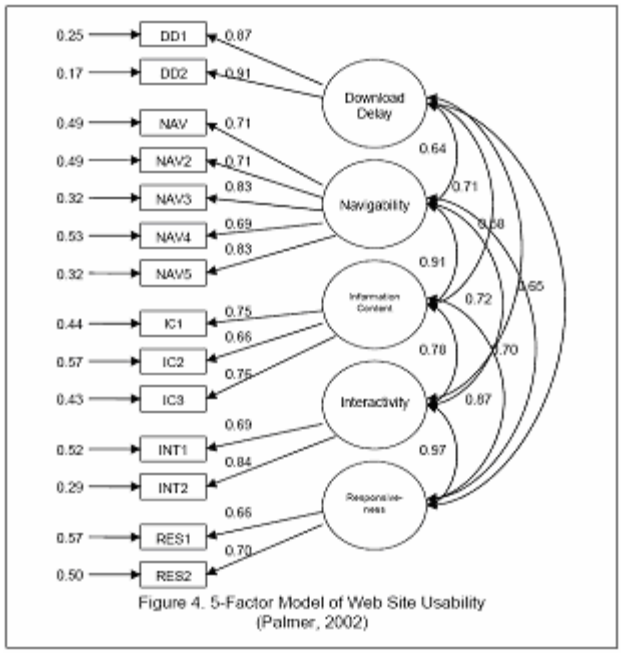

Figure 4. 5-Factor Model of Web Site Usability (Palmer, 2002)

Figure 4. 5-Factor Model of Web Site Usability (Palmer, 2002)

Discussion

The Agarwal and Venkatesh instrument demonstrated weaknesses at multiple absolute, incremental, and parsimonious fit levels, signifying the superiority of the Palmer instrument in its various measures of overall model goodness-of-fit for the Palmer instrument lending sufficient support to deeming the results an acceptable representation of the hypothesized constructs.

With the acceptance of the overall Palmer model (*Figure 4*), each of the constructs in that instrument alone, can now be evaluated separately. The constructs were evaluated based on the statistical significance of the indicator loadings and the construct's reliability and variance extracted. Each variable's t-values associated with each of the loadings exceed the critical values for the .05 significance level, thus showing that all variables are significantly related to their specified constructs, verifying the proposed relationships between the exogenous indicator variables and the exogenous constructs (*Table 2*). These significant relationships support the hypothesized relationships of the observed variables with the overall latent constructs.

CONCLUSIONS

This research attempted to validate two proposed metrics/instruments for use in e-commerce research. Although both the Palmer (2002) and the Agarwal and Venkatesh (2002) Web site usability instruments displayed nomological validity in their respective validation studies it remained unclear up until this point if the instruments were useful as standardized research instruments for understanding the underlying constructs of the Web site usability construct. Although the underlying theoretical constructs that compose both instruments remains open to argument and will see future debate, it appears that of the two, the Palmer (2002) Web site usability instrument contains great potential for its utility in research examining Web site use, particularly in an e-commerce setting. Researchers must continue to explore the underlying constructs of usability in both the setting of Web sites as well as software and information technologies.

Observed Variables			Latent Variables
	Factor Loading	R-Square (Reliability)	
DD1	0.87	0.76	Download
D2	0.91	0.83	Delay
NAV1	0.71	0.50	Navigability
NAV2	0.71	0.50	
NAV3	0.83	0.69	
NAV4	0.69	0.48	
NAV5	0.83	0.69	
IC1	0.75	0.56	Information
IC2	0.66	0.44	Content
IC3	0.76	0.58	
INT1	0.69	0.58	Interactivity
INT2	0.84	0.71	
RES1	0.66	0.44	Responsiveness
RES2	0.70	0.49	

Table 2. Standardized Parameter Estmates and t-values for the Palmer Instrument (2002) (n=159)

(Supporting Figures and Tables are part of the full version of the paper and may be requested from the authors.)

REFERENCES

1. Agarwal, R., & Karahanna, E. (2000) Time Flies When You're Having Fun: Cognitive Absorption and Beliefs about Information Technology. *MIS Quarterly,* 24, 4, 665-694.

2. Agarwal, R., & Prasad, J. (1997) The Role of Innovation Characteristics and Perceived Voluntariness in the Acceptance of Information Technologies. *Decision Sciences,* 28, 3, 557-582.

3. Agarwal, R., & Venkatesh, V. (2002) Assessing a Firm's Web Presence: A Heuristic Evaluation Procedure for the Measurement of Usability. *Information Systems Research,* 13, 2, 168-186.

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

5. Bentler, P. M., & Chou, C. (1987) Practical Issues in Structural Modelling. *Sociological Methods and Research,* 16, 78-117.
6. Berry, J. (1999, October 4) The World According to e-Biz Metrics. *Internetweek,* 38.
7. Bogner, W. C., & Barr, P. S. (2000) Making Sense in Hypercompetitive Environments: A Cognitive Explanation of High Velocity Competition. *Organization Science,* 11*,* 2, 212-226.
8. Bollen, K. A. (1989) *Structural Equations with Latent Variables*. New York: Wiley.
9. Boudreau, M., Gefen, D., & Straub, D. W. (2001) Validation in IS Research: A State-of-the-Art Assessment. *MIS Quarterly,* 25, 1, 1-24.
10. Byrne, B. M. (1989) *A Primer of LISREL: Basic Applications and Programming for Confirmatory Factor Analytic Models*. New York: Springer-Verlag.
11. Cook, T. D., & Campbell, D. T. (1979) *Quasi Experimentation: Design and Analytical Issues for Field Settings*. Chicago: Rand McNally.
12. Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989) User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science,* 35*,* 8, 982-1003.
13. Doll, W. J., Xia, W., & Torkzadeh, G. (1994) A Confirmatory Factor Analysis of the End-User Computing Satisfaction Instrument. *MIS Quarterly,* 18*,* 4, 453-461.
14. Hair, J. F., Tatham, R. L., Anderson, R. E., & Black, W. (1998) *Multivariate Data Analysis*. New York: Pearson Education.
15. Joreskog, K. G., & Sorbom, D. (1989) *LISREL 7 User's Reference Guide*. Chicago: Scientific Software Inc.
16. Lynch, J. G., Jr., Calder, B. J., Phillips, L. W., & Tybout, A. M. (1982) On the External Validity of Experiments in Consumer Research: The Concept of External Validity. *Journal of Consumer Research,* 9*,* 3, 225-244.
17. Mackenzie, K. D., & House, R. (1979) Paradigm Development in the Social Sciences. In R. T. M. a. R. M. Steers (Ed.), *Research in Organizations: Issues and Controversies*. Santa Monica, CA: Goodyear Publishing.
18. McGrath, J. (1979) Toward a 'Theory of Method' for Research on Organizations. In R. T. M. a. R. M. Steers (Ed.), *Research in Organizations: Issues and Controversies*. Santa Monica, CA: Goodyear Publishing.
19. Moore, G., & Benbasat, I. (1991) Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation. *Information Systems Research*, 2, 192-222.
20. Mueller, R. O. (1994) *Basic Principles of Structural Equation Modeling. An Introduction to LISREL and EQS*. New York: Springer-Verlag.
21. Nielsen, J. (2000) *Designing Web Usability: The Practice of Simplicity*. Indianapolis: New Riders Publishing.
22. Palmer, J. (2002) Web Site Usability, Design, and Performance Metrics. *Information Systems Research,* 13, 2, 151-167.
23. Peppers, D., & Rogers, M. (1999) *Enterprise One to One: Tools for Competing in the Interactive Age*. New York: Doubleday.
24. Rose, G., & Straub, D. W. (1998) Predicintg General IT Use: Applying TAM to the Arabic World. *Journal of Global Information Management,* 6, 3, 39-46.
25. Stevens, J. P. (2002) *Applied Multivariate Statistics for the Social Sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.
26. Straub, D. W. (1989) Validating Instruments in MIS Research. *MIS Quarterly,* 13, 2, 147-169.
27. Straub, D. W., Hoffman, D. J., Weber, B. W., & Steinfield, C. (2002) Measuring e-Commerce in Net-Enabled Organizations: An Introduction to the Special Issue. *Information Systems Research,* 13, 2, 115-124.
28. Tanaka, J. S., & Huba, G. J. (1984) Confirmatory Hierarchical Factor Analysis of Psychological Distress Measures. *Journal of Personality and Social Psychology,* 46, 3, 621-635.
29. Venkatesh, V. (2000) Determinants of Perceived Ease of Use: Integrating Perceived Behavioral Control, Computer Anxiety and Enjoyment into the Technology Acceptance Model. *Information Systems Research,* 11, 342-365.
30. Venkatesh, V., & Speier, C. (2000) Creating an Effective Training Environment for Enhancing Telework. *International Journal of Human-Computer Studies,* 52, 991-1005.
31. Weber, T. (1998, April 16) Who, What, Where: Putting the Internet in Perspective*. The Wall Street Journal,* pp. B12.
32. Webster, J., & Ho, H. (1997) Audience Engagement in Multimedia Presentations. *Data Base for the Advancement of Information Systems,* 28, 63-77.