

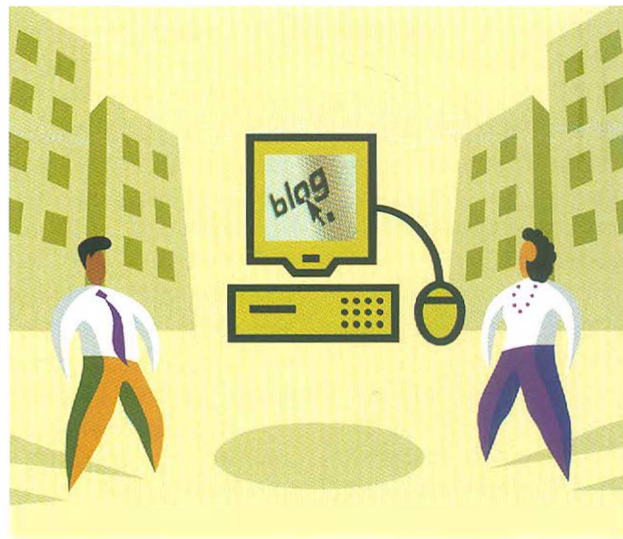
ABSTRACT

Predictions are that blogs will be the promotional tool of the future, if not now. The question remains: how credible are blogs as a source of consumer information? How do consumers perceive the credibility of blogs from different sources when making purchase decisions and product evaluations? The primary objective of the present study is to assess validity and reliability of a set of Semantic Differential scales purported to measure perceived credibility of different types of blogs. Confirmatory factor analysis via LISREL8.5 package was used to obtain appropriate statistics for convergent validity, discriminant validity, and composite reliability. In general, the scales used to measure the two hypothesized factors (source credibility and content credibility) for each type of blog achieved a satisfactory level of construct validity.



Measures Of Perceived Credibility Of Blogs: Construct Validation By Lisrel

Dr.Uday S. Tate, Dr. Elizabeth Alexander,
Dr. Suneel K. Maheshwari



INTRODUCTION

Predictions are that blogs will have a major impact on the operations of corporations world-wide. Blogs, short for weblogs, are defined as interactive websites or online journals. Essentially, a blog allows an individual to read online postings, which are frequently updated. Blogs also allow individuals to post their own comments or responses.

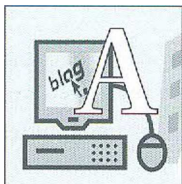
Blogs may contain text, media, images and data, all arranged in a chronological fashion. Some blogs also utilize links to other sites and advertising (Dearstynne 2005; Quible 2005).

Blogs are a form of communication widely used by individuals for various purposes: for virtual family reunions, to express an individual's point of view, to allow others to keep up to date with one's activities (Quible 2005). Organizations can use blogs internally and externally. An internal blog can be utilized to disseminate information quickly and more thoroughly between areas of an organization. Externally, the blog may be used to communicate to prospects and customers, providing a two-way flow of information that allows deeper insight into customer needs and wants.

Although blogging is still not utilized by many corporations, it is being hailed as a new marketing medium. A recent study indicated that less than 5% of the Fortune 1000 was using blogs, with varying degrees of effectiveness (Baker 2006; DeFelice 2006). Simultaneously, business publications, like *Business Week* (May 2, 2006) have featured cover stories on blogs and their impact on corporate America.

Blogs by corporations, if used effectively, may play a role in the consumer decision process. Potential buyers may use corporate blogs as a source of information and decision aid. In the post-purchase stage, the blog can be used as a channel for either complaint or praise. Corporate blogs can also provide information of additional product benefits or answers to utilization questions. Of course, consumers may also choose to use other blog types for similar tasks. A blog maintained by an individual could be accessed for product information before and after purchase. Blogs of all types may provide motivation to seek additional product information.

Blogs are here to stay as a widely used source of consumer information, it is, therefore, important that one must investigate credibility of blogs as perceived by consumers. After reviewing several studies on blogs, Johnson and Kaye (2004) concluded that while there is no dearth of research on credibility of traditional sources of consumer information, majority of the reviewed studies focused on the traditional media such as radio, television, and newspapers. These authors have called for research to examine the credibility of blogs as sources of information. However, before any such attempt is made, it is imperative that researchers focus on the psychometric properties (i.e. validity and reliability) of scales they use to measure source and content credibility with respect to blogs. The present study attempts to assess the construct validity (convergent validity, discriminant validity, and reliability) of scales used to measure perceived credibility of blogs developed by different sources, namely, members of the media, individuals, and corporations.



TWO-STEP VALIDATION APPROACH

Bagozzi (1980) suggests a two-step approach in assessing validity and reliability of measurement scales. First, one must

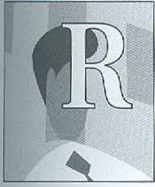
determine if a hypothesized factor structure is supported by empirical evidence gathered via measurement scales. In this regard, a confirmatory factor analysis should be used to examine a theoretical soundness of the proposed factor structure. Second, after testing the viability of hypothesized factor structures, one then must assess psychometric properties of the scales used to measure the hypothesized factors.

Based on an extensive literature review, the present research hypothesized two factors underlying the credibility with regard to blogs, namely, source credibility and content credibility. In addition, we used an exploratory factor analysis (principal component analysis) with a holdout sample of 80 out of the original sample of 418. The principal component analysis identified two factors labeled as: source credibility and content credibility.

The second step deals with the assessment of construct validity of measures of hypothesized constructs. According to Trochim (2006), "construct validity refers to the degree to which inferences can legitimately be made from the operationalizations in one's study to the theoretical constructs on which those operationalizations were based." In other words, construct validity is the extent of agreement between a theoretical concept and its empirical evidence (in terms of specific measurement scales). Bagozzi (1980), James, Mulaik, and Brett (1982), Joreskog and Sorbom (1983), Fornell and Larcker (1981), and Netermeyer, Johnston, and Burton (1990) propose three basic criteria of construct validity, namely, convergent validity, discriminant validity, and reliability. Specifically, convergent validity is defined as "the degree to which the operationalization is similar to (converges on) other operationalizations that it theoretically should not be similar to" (Trochim 2006). On the other hand, discriminant validity is "the degree to which the operationalization is not similar to (diverges from) other operationalizations that it theoretically should not be similar to" (Trochim 2006). Finally, in the present research, reliability was defined in terms of internal consistency of the scales used to measure the hypothesized constructs (source credibility and content credibility). Churchill (1979) defines the internal consistency reliability as the degree to which the individual items or indicators of a construct are intercorrelated; the higher the intercorrelations among the items, the higher the internal consistency reliability. The current research used the three criteria to establish the construct validity of the scales reported in Table 1 for each source.

To assess convergent validity, discriminant validity, and internal consistency of the scales used to measure source and content credibility, the present research used the correlation matrix of the nine scales of perceived credibility with regard to each blog source (members of media, individuals, and corporations). LISREL 8.5 version (Jorskog and Sorbom 2003

was used to obtain appropriate indices and statistics to assess the construct validity and reliability of the measures of source and content credibility of blogs. Due to space limitations, the correlation matrices are not displayed; interested readers or reviewers can obtain them from the contact author.



RESEARCH QUESTIONS AND METHODOLOGY

The present research examines the psychometric properties of scales used to measure perceptions of credibility with regard to blogs authored by the media, individuals, and corporations. Results of our study will allow for an initial understanding of the psychometric properties of the scales used to measure consumer perceptions with regard to source and content credibility of blogs. In addition, it will provide a foundation to further refine the existing scales.

In order to collect data for our study, a questionnaire based on a comprehensive literature review, a focus group session, and faculty input was used. The questionnaire included a series of scales measuring the experience of respondents with blogs and their perceptions of blogs as a source of information. Respondents' perceptions were measured on nine 5-point, Semantic Differential scales (for example: trustworthy-untrustworthy, honest-dishonest and so on). The same set of nine items was used in the case of blogs by the members of media, individuals, and corporations. The present study used the measurement scales developed by Beltramini (1988). However, one of the original ten items was omitted due to its irrelevance to the topic under study.

The study used convenience sampling conducted by trained research assistants who distributed the questionnaire to student and non-student groups in Huntington, West Virginia, USA, and the surrounding communities. Analysis confirms a representative sample in terms of age, gender, education, and income groups.

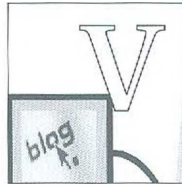


SURVEY ANALYSIS

The survey resulted in 418 usable questionnaires. The sample consisted on 52 percent male and 48 percent female respondents with the average of 27.6 years.

The majority of the respondents were college students (56 percent). Of the total sample, 65 (15%) reported not having access to a computer at their residence, and 43 (9%) reported not using the Internet. When reporting blog experience, 99 (23%) reported some level of experience with blogs and 169 (39%) reported knowledge of but no experience with a blog. Of those reporting at least knowledge of blogs, only 166 or 2/3rd were able to correctly identify the definition of a blog.

Respondents with at least some blog knowledge were asked to evaluate the credibility of information contained in a blog operated and owned by members of the media, an individual, and a corporation. A previously developed scale was adjusted and utilized for this measure (Beltramini, 1988; Bruner and Hensel 1998). Nine items were rated on a 5-point Semantic Differential scale. The means and standard deviations of the scales for each type of blog are reported in Table 1.



VALIDATION RESULTS AND DISCUSSION

The main purpose of the present study was to assess the construct validity of scales used to measure perceived credibility of blogs developed by different sources, namely, by members of the media, individuals, and corporations. As suggested previously, the first step was to test if the two hypothesized factors (source credibility and content credibility) were supported by the measurement scales. For this purpose, we compared three measurement models: no-factor (null) model, one-factor model, and the hypothesized two-factor model. For each type of blog, the correlation matrix was submitted to LISREL 8.5 to obtain appropriate statistics and indices.

As can be seen in Table 2, across the three blogs, the worst fit is for the null model followed by one-factor model. On the other hand, the best fit is for the correlated two-factor model across the three types of blog, as evidence by non-normed fit index, (NNFI), parsimony normed fit index (PNFI), adjusted goodness of fit index (AGFI), and root mean square residuals (RMSR). According to Bentler and Bonnett (1980, a value of 0.90 and above for NNFI and AGFI indicates an "acceptable" fit for the model. From Table 2, one can see that both the estimated AGFI (except for blogs by corporations) and NNFI satisfy the Bentler-Bonnett condition. PNFI is useful when comparing a series of nested models. In the present case, the PNFI index increased from zero for the null model to 0.67, 0.69, and 0.65 for the correlated two-factor model in each case (blogs by members of media, individuals, and corporations, respectively). Finally, according to Medsker, Williams, and Holahan (1994), RMSR indicates the mean of the differences between of the theoretical covariance matrix and the observed covariance matrix. According to Joreskog and Sorbom (1989), the value of RMSR closer to zero indicates a better fit of the model. Table 2 shows that RMSR for the three sources of blogs decreases from the null model to its counterpart of correlated two-factor model (for example, the RMSR value for the null model for the blogs by members of media is 0.39 and that of the correlated two-factor model is 0.06). In sum, the two-factor model for each of the source of blogs had achieved an acceptable goodness-of-fit, as indicated by various indices shown in Table 2.

It should be noted that a mere empirical justification of the hypothesized factors is not sufficient. As a next step, one must assess the psychometric properties of the scales used to measure the hypothesized factors. As stated previously, the present research used three criteria for construct validity: convergent validity, discriminant validity, and internal consistency reliability. Again, we used LISREL 8.5 to obtain appropriate statistics to assess the construct validity of the measurement scales. Tables 3 through 5 show the necessary statistics for the validation purpose.

To assess convergent validity of the measures, Joreskog and Sorbom (1989) suggest that a factor loading of 0.70 and above is an acceptable value to assess the convergence between a measure and its theoretical construct (factor). As can be seen in Tables 3-5, the factor loadings for each of the factors (source credibility and content credibility) are at or above 0.70, with an exception of the factor loading for the item labeled as "unquestionable." The variance-extracted index for factor is 0.56 and 0.57, for source credibility and content credibility, respectively. A value of 0.50 or above is acceptable (Netermeyer, Johnston, and Burton 1990), indicating a satisfactory evidence for convergent validity of the scales used to measure perceived source and content credibility across three types of blogs.

Discriminant validity is the extent to which two factors (theoretical constructs) are independent from each other, measured in terms of their respective scales. Fornell and Larcker (1981) suggest that discriminant validity is evident when the squared value of the correlation between two factors is less than the variance extracted from each factor separately. Discriminant validity of the measures of the two hypothesized factors was assessed by comparing the squared value of the correlation between the two factors with their respective variance extracted. As can be seen in Table 6, the squared correlation value between the two factors for each type of blog was less than the variance extracted for each factor separately, implying that the two factors have achieved an acceptable level of discriminant validity. Another test for discriminant validity was performed by comparing two models (correlated two-factor models versus perfectly correlated two-factor models). Table 2 shows that NNFI, PNFI, AGFI, and RMSR show poor fit of the models for all perfectly correlated two-factor models, implying that the two factors, source credibility and content credibility have achieved discriminant validity.

Finally, reliability can be determined by assessing the internal consistency of the measures. The composite reliability of the measures of the two hypothesized factors was calculated using the formula provided by Fornell and Larcker (1981, p. 45). An internal consistency of 0.70 is considered as satisfactory (Joreskog and Sorbom 1989). The composite reliability indices (internal consistency) of the measures of the two hypothesized factors are shown in Table 6. In each case, the

reliability index of the measures of the two hypothesized factors was high, ranging from 0.84 to 0.88, indicating a strong evidence for internal consistency reliability for the measures of the two hypothesized factors.



CONCLUSION

It is imperative that researchers must evaluate the construct validity and reliability of the scales used to measure theoretical constructs before any predictive or explanatory conclusions can be drawn in their study. Specifically, if we are to examine how consumers use blogs as a source of information, what specific information do they use in arriving at their purchase decisions, and when do they use such information, first we must assess how credible do consumers perceive blogs as sources of information. The primary focus of the present study was to assess the convergent validity, discriminant validity, and composite reliability of the scales used to measure two factors of credibility of blogs, namely, source credibility and content credibility. For this purpose we used a rigorous methodology known as structural equations methodology to obtain appropriate statistics to assess the construct validity. The preliminary results indicate that the scales have achieved satisfactory level of construct validity.

Future research effort should examine the underlying causes and consequences of the factors of perceived credibility with regard to blogs used by members of media, individuals, and corporations. Such an effort can be helpful in developing blogs and their contents that can be perceived as credible by consumers. In addition, researchers need to examine the use of blogs as a source of information by consumers under various buying situations and at various stages in consumer decision processes. Finally, a segmentation study can be performed to identify groups among blogs users and to develop their demographic and psychographic profiles. Such research effort can be useful in targeting blogs to specific groups of interest.

As with any research design, there are several caveats in the present study. One, the sample was a convenient sample. Two, the study was a cross-sectional and not longitudinal. Three, the limited sample size did not allow to take a holdout sample for the purpose of cross validation of the results. Four, the present research used only one method (Semantic Differential Scales) to measure the credibility constructs. Hence there is a possibility of methods bias in the results. Ideally, one must use a multi-trait, multi-method approach to assess construct validity (Bagozzi 1980). Due to these caveats, the results of the present study cannot be generalized to other populations.

Table 1
Means and Standard Deviations of Scales Measuring Credibility Perception

Scale	Blogs by Members of Media		Blogs by Individuals		Blogs by Corporations	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Unbelievable-Believable	2.96	.96	2.90	1.08	3.20	1.12
Trustworthy-Untrustworthy	2.95	.91	2.90	1.04	3.08	1.04
Not convincing-Convincing	3.10	.94	2.98	1.00	3.34	.94
Unreasonable-Reasonable	3.07	.87	2.96	.95	3.31	.93
Credible-Not credible	2.88	.93	2.97	.99	3.08	.99
Honest-Dishonest	2.96	.86	2.98	1.01	3.07	1.00
Unquestionable-Questionable	2.73	.90	2.79	.96	2.91	.93
Conclusive-Inconclusive	2.96	.87	2.88	.97	3.14	.97
Not authentic-Authentic	3.02	.87	3.00	1.06	3.26	.98

Note: 1 represents the "negative anchor" of the scale, whereas 5 reflects the positive anchor.

Table 2
Nested Model Comparison: Three Sources of Blogs

Model	X ²	df	NNFI	PNFI	AGFI	RMSR
Blogs by Members of Media						
Null Model	1033.58	36	.00	.00	.24	.39
One-Factor Model	320.14	27	.61	.52	.55	.13
Correlated Two-Factor Model	79.69	26	.93	.67	.90	.06
Perfectly Correlated Two-Factor Model	138.78	27	.85	.65	.82	.49
Blogs by Individuals						
Null Model	967.05	36	.00	.00	.31	.35
One-Factor Model	467.16	27	.37	.39	.41	.22
Correlated Two-Factor Model	37.12	26	.98	.69	.94	.04
Perfectly Correlated Two-Factor Model	136.82	27	.84	.64	.85	.65
Blogs by Corporations						
Null Model	1319.03	36	.00	.00	.18	.43
One-Factor Model	468.24	27	.54	.48	.41	.15
Correlated Two-Factor Model	123.88	26	.89	.65	.84	.05
Perfectly Correlated Two-Factor Model	178.21	27	.84	.65	.79	.51

Note: df=degrees of freedom; NNFI=non-normed fit index; PNFI=parsimony fit index; AGFI=adjusted goodness of fit index; RMSR=root mean square residuals.

Table 3
Measurement Properties for Blogs by Members of Media

Factor and Indicators	Standardized loading	Error variance	Variance extracted estimate	t-values ^b
Source Credibility ^d	---	.88 ^a	.56	
Trust	.70	.52	.49	---
Credibility	.82	.33	.67	11.26
Honest	.82	.33	.67	11.30
Unquestionable	.67	.55	.45	9.38
Conclusive	.75	.44	.56	10.43
Content Credibility ^d	---	.84 ^a	.57	
Believable	.80	.36	.64	---
Authentic	.62	.62	.38	9.37
Reasonable	.77	.40	.60	11.99
Convincing	.80	.36	.64	12.38

Note: ^adenotes composite reliability using the formula by Fornell and Lacker (1981).
^bt-values significant at p-level < .01.
^cThe indicator fixed to the value of 1 as a reference variable.
^dThe estimated correlation between the two factors = .51.

Table 4
Measurement Properties for Blogs by Individuals

Factor and Indicators	Standardized loading	Error variance	Variance extracted estimate	t-values ^b
Source Credibility ^d	---	.85 ^a	.54	
Trust	.79	.38	.62	---
Credibility	.70	.51	.49	10.64
Honest	.75	.44	.56	11.44
Unquestionable	.62	.62	.38	9.34
Conclusive	.80	.36	.64	12.17
Content Credibility ^d	---	.87 ^a	.62	
Believable	.71	.50	.50	---
Authentic	.81	.34	.66	11.28
Reasonable	.81	.34	.66	11.23
Convincing	.82	.33	.67	11.39

Note: ^adenotes composite reliability using the formula by Fornell and Lacker (1981).
^bt-values significant at p-level < .01.
^cThe indicator fixed to the value of 1 as a reference variable.
^dThe estimated correlation between the two factors = .45.

Table 5
Measurement Properties for Blogs by Corporations

Factor and Indicators	Standardized loading	Error variance	Variance extracted estimate	t-values ^b
Source Credibility ^d	---	.88 ^a	.55	
Trust	.81	.34	.66	---
Credibility	.90	.20	.81	16.00
Honest	.80	.36	.64	13.83
Unquestionable	.60	.64	.36	9.61
Conclusive	.75	.44	.56	12.63
Content Credibility ^d	---	.88 ^a	.66	
Believable	.75	.43	.57	---
Authentic	.88	.23	.77	13.51
Reasonable	.81	.34	.66	12.50
Convincing	.79	.38	.62	12.11

Note: ^adenotes composite reliability using the formula by Fornell and Lacker (1981).
^bt-values significant at p-level < .01.
^cThe indicator fixed to the value of 1 as a reference variable.
^dThe estimated correlation between the two factors = .52.

Table 6
Evidence of Discriminant Validity of the Two Factors

Factors	Blogs by Members of Media		Blogs by Individual		Blogs by Corporations	
	Source Credib.	Content Credib.	Source Credib.	Content Credib.	Source Credib.	Content Credib.
Variance Extracted	.56	.57	.54	.62	.55	.66
Squared Correlation Between two factors		.26		.20		.27
Composite reliability	.88	.84	.85	.87	.88	.88

REFERENCES

1. Bagozzi, R. P. (1980). *Causal modeling in marketing*. New York, NY: Wiley.
2. Baker, S. (2006). The inside story on company blogs. Accessed March 6, 2006, [available at <http://www.businessweek.com/>].
3. Beltramini, R. F. (1988). Perceived believability of warning label information presented in cigarette advertising. *Journal of Advertising*, 17 (1): 26-32.
4. Bentler, P. M. & Bonnett, D. C. (1980). Significant tests and goodness fit in the analysis of covariance structures. *Psychological Bulletin*, 80: 588-606.
5. Blood, R.. "Weblogs: A history and perspective", *Rebecca's Pocket*. 07 September 2000. Accessed on August 30, 2006 at www.rebeccablood.net/essays/weblog_history.html.
6. Bruner, G. C. & Hensel, P. J. (1998). *Marketing Scales Handbook: A compilation of Multi-Item Measures, Vol. II*. Chicago, IL: American Marketing Association.
7. BusinessWeek, May 2, 2006.
8. Churchill, G. A. A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, 16 (February): 64-73
9. Dearstyne, B. W. (2005). Blogs: the new information revolution?. *The Information Management Journal*, September/October: 38-44.
10. DeFelice, A. (2006). A new marketing medium. *Customer Relationship Management*, December: 32-35.
11. Fornell, C. & Lacker, D. F. (1981). Evaluating structural equation models with observed variables and measurement errors. *Journal of Marketing Research*, 18: 39-50.
12. James, L. R., Mulaik, S. A., & Brett, J. (1982). *Causal analysis: Models, assumptions, and data*. Beverly Hills, CA: Sage.
13. Johnson, T. J. & Kaye, B. K. (2004). Wag the blog: How reliance on traditional media and the internet influence credibility perceptions of weblogs among blog users. *J&MC Quarterly*, (81 (3): 622-642.
14. Joreskog, R. G. & Sorbom, D. (1983, 1989). LISREL: Analysis of linear structural relations by the method of maximum likelihood (Version VII). Chicago: National Education Resources.
15. Medsker, G. J., Williams, L. J., & Holahan, P. J. (1994). A review of current practices for evaluating causal models in organizational behavior and human resources management research. *Journal of Management*, 20 (2): 439-464.
16. Netermeyer, R. G., Johnston, M. W., & Burton, S. (1990). Analysis of role conflict and role ambiguity in a structural equations framework. *Journal of Applied Psychology*, 75 (2): 148-157.
17. Quible, Z. K (2005). Blogs and aritten business communication courses: A perfect union. *Journal of Education for Business*, July/August: 327-332.
18. Trochim, M. K. (2006). Convergent and discriminant validity. Accessed on November 20, 2006, at www.cosialresearchmethods.net/kbconvdisc.htm.