A Word-of-Mouth Scale in a Services Context

Jillian C. Sweeney, Geoffrey N. Soutar and Tim Mazzarol, University of Western Australia

Abstract

Word of mouth (WOM) marketing has become a key focus for many organisations, given the decreasing trust customers are placing in organisations and advertising. Against this background, academic research into WOM has failed to keep pace with managers' escalating interest in the topic. Questions such as what is WOM, how is it measured and how can it be managed have not been fully answered by prior research. In particular, much of the literature views WOM as the simple delivery of information and does not take account of the construct's richness. The present study addressed this gap by developing a parsimonious and practical three-factor scale (emotional intensity of the message, the cognitive quality of the message and the credibility of the message) with good measurement properties. The scale is appropriate for both WOM delivery (i.e. when WOM is modelled as an outcome of a service process) as well as WOM receipt (i.e. examining outcomes following WOM). No previous WOM measures have been usable in both contexts.

Introduction

Organizations are continually striving for new ways to achieve and retain a competitive edge while, simultaneously, customers are increasingly discerning and demanding and have more choice. Further, recent research has shown consumers are less attentive to traditional advertising (McDonnell, 2005), leading many organisations to reduce advertising expenditure. In this context, word-of-mouth (WOM) provides a new approach to marketing communication (First International Conference on Word-of-Mouth, 2005). WOM is not a new idea. Indeed, in the early 1970's, studies suggested WOM was nine times as effective as advertising in converting unfavourable or neutral predispositions into positive attitudes (Day, 1971). Despite this, there has been very little research into the meaning of the construct, although the present study is an attempt to fill this gap.

WOM has been defined as informal communication between private parties about goods or services (Anderson, 1998). In line with this definition, several researchers measured WOM in terms of the number of people to whom WOM was given and the frequency of mention (e.g. Westbrook, 1987). However, such a definition fails to address WOM's power and scope; certainly it should not be reduced to a number of utterances. In a similar vein, others conceived WOM as the likelihood people would engage in WOM and used a single item to measure this (e.g., Hartline and Jones, 1996; Parasuraman *et al.*, 1988). Anderson (1998) is one of the few researchers to suggest WOM can vary in terms of its vividness, pleasantness and novelty and the extent to which it conveys a sender's experiences. However, Anderson did not consider these aspects in his measure, instead adopting a similar measure to that used by Westbrook and others. Harrison-Walker (2001), recognising weaknesses in previous WOM measures and consistent with the present discussion, developed a two-dimensional WOM measure (WOM activity, in line with previous research, and WOM praise that addressed positive or negative valence and the extent of such valence). While this was a first step in the development of a more sensitive WOM measure, the present study sought to explore WOM in more detail.

Harrison-Walker, and many previous researchers, examined WOM as it related to a particular organisation over time. In contrast, we believe it is important to address a particular incident in which WOM is given or received as these WOM incidents are the basic building block of WOM activity. The WOM incident was the unit of analysis used in the present study; this is discussed in the next section. The present study concentrated on WOM from a giver's perspective in a services context. A services context was chosen as services are intangible, difficult to evaluate before purchase, not covered by guarantee and non-standardised and, hence, are often perceived as high risk (e.g. Murray, 1991), which means WOM communication is likely to be more important as it acts as a risk reliever in such situations (Ennew, Bannerjee and Li, 2000; File, Cermak and Prince, 1994).

The Present Study

A total of 358 consumers were contacted by phone using a random sample obtained from the data base of a financial services provider. Respondents were asked to recall a situation in which they gave WOM about any service provider and to describe the incident in terms of the nature of the message and the way in which it was delivered. Sixty descriptors were used for this purpose that had been identified from a previous exploratory phase that included focus groups and critical incident interviews. Respondents mentioned a variety of services, including financial services, restaurants, health services and property services.

The data were examined in a variety of ways to find a parsimonious word of mouth scale for use in marketing studies The descriptive statistics suggested that a number of items had high skew (greater than 6) or low variance (lower than 0.60 on the 5-point scales used). These items were removed, since items that are skewed create problems in analysis, while items that have very little variance provide little information. As there was no indication of dimensionality other than the earlier qualitative analysis, an initial exploratory factor analysis (EFA) of the remaining 28 items was undertaken. The Measure of Sampling Adequacy (MSA) in this case was 0.94, which is meritorious (Kaiser, 1974) and suggests that an EFA is likely to provide useful information. The MSAs for each of the twenty eight items were also computed and ranged from 0.90 to 0.97, which suggests all should be included in the analysis (Hair et al., 2006). A principal axis factor analysis found four factors with eigenvalues greater than one. However, the scree diagram strongly suggested three factors, which explained 54% of the variance in the item set, should be retained. An oblimin rotation was used to obtain simple structure as there was no reason to assume the factors were not correlated. Indeed, the resulting factor correlations ranged from 0.41 to 0.46, which suggested there were moderate positive correlations between the retained factors.

The rotated factor loadings are shown in Table 1. As can be seen from the Table, most of the items loaded onto one and only one factor beyond the 0.40 level. However, three items (explicit, reinforcing and detailed) had cross loadings greater than 0.50, while two items (interesting and lifelike) had loadings that were less than 0.40. Following Verbeke and Bagozzi's (2002) suggestion, these five items were removed, leaving 23 items. The items that loaded most strongly onto the first factor included items such as using strong words, powerfully, in a strong way and intense, suggesting an emotional intensity to the message delivery. The second factor was more cognitive as it related to the transfer of the WOM information. The items that loaded most strongly onto the third factor included influential, helpful and effective and the factor was therefore termed the credibility of the message.

The correlations between the items and the total scale were computed and items that had relatively low correlations were excluded (Churchill, 1979). This resulted in two items being removed from the first factor, three items being removed from the second factor and one item being removed from the third factor. The remaining seventeen items were used to more fully examine the sub-scales' measurement properties by undertaking a series of confirmatory factor analyses that initially looked at the three factors separately. As the objective was to find a parsimonious set of word of mouth sub-scales, further items were removed if they had standardised coefficients that were less than 0.60 (Bagozzi and Foxall, 1996), or if they had errors that were correlated with items that had larger loadings as such items often cause problems if used in structural equation modelling (Byrne, 2001). Items were removed iteratively, starting with those that had low loadings, before correlations between error terms were examined.

Table 1: Rotated Factor Loadings – Word of Mouth Delivery

Questionnaire Item	**Emotional intensity of message delivery**	**Cognitive quality of message delivery**	**Credibility of the message**
Using strong words	0.72		
Powerfully	0.71		
In a strong way	0.70		
Intense	0.69		
Powerful.	0.67		
In a way that was explicit	0.66	0.62	
Elaborate	0.65		
In a strong tone of voice	0.64		
In a reinforcing way	0.63		
Explicit	0.62		
Reinforcing	0.62		0.60
Persuasively	0.62		
Persuasive	0.62		
Vivid	0.58		
Interesting	0.33		
Sincerely		0.78	
Informatively		0.78	
In a reliable manner		0.76	
Honestly		0.72	
As advice		0.68	
In a detailed way	0.55	0.61	
Detailed		0.59	
Objectively		0.56	
Influential			0.75
Helpful			0.71
Effective			0.69
Credible			0.67
Lifelike			0.28

After this procedure had been undertaken for the first factor, four items were retained, as can be seen in Table 2. The correlation between the summated scale obtained from all of the

items and the summated scale obtained from the four items was 0.89, which suggests little information was lost by using the smaller number of items (Thomas, Soutar and Ryan, 2000). The data fitted the model well in this case, as the chi square statistic was not significant (1.01, df=2) and all of the other goodness of fit measures were acceptable. The same procedure was used to examine the second and third factors. A total of five items and four items were retained for these two scales respectively (based on correlations of 0.94 and 0.95 with the original scales and chi-square values of 12.88, df=5 and 5.93, df=2 respectively). The chi-square values were not significant in either case and all of the other goodness of fit measures were acceptable.

Table 2: Confirmatory Factor Analyses – Word of Mouth Sub-Scales

Sub-Scales	**Standardised Loadings**	**Average Variance Extracted**	**Construct Reliability**	**CFA Chi-square**
Factor 1		0.51	0.81	1.01(2df)
Using strong words	0.73			
Powerfully	0.79			
In a strong way	0.66			
In a reinforcing way	0.65			
Factor 2		0.52	0.84	12.88(5df)
Sincerely	0.82			
Informatively	0.75			
In a reliable manner	0.77			
Honestly	0.75			
As advice	0.70			
Factor 3		0.54	0.82	5.93(2df)
Influential	0.73			
Helpful	0.72			
Effective	0.77			
Credible	0.71			

Table 2 also shows the construct reliability and average variance extracted (AVE) for each of the sub-scales. As can be seen from the Table, all of the AVE values exceeded 0.50 and all of the construct reliability values exceeded 0.70, which suggests the three scales are reliable and have convergent validity (Fornell and Larcker, 1981). However, further analysis is needed to determine whether the sub-scales have discriminant validity (i.e. whether they measure different WOM aspects). Fornell and Larcker (1981) have suggested that discriminant validity is present when the square of the correlation between two constructs is less than the minimum AVE of both constructs. As the minimum AVE was 0.51 and the maximum squared correlation was 0.48 (0.69^2), there is discriminant validity between the three scales.

The analysis suggested there is a parsimonious set of items that can be used to measure three distinct aspects of word of mouth (the emotional intensity of the message delivery, the cognitive quality of the message delivery and the credibility of the message) in such a way that the sub-scales have good measurement properties. The means for the three sub-scales were 3.36, 4.40 and 4.16 respectively on the five-point scale used. It seems that the senders surveyed felt they provided useful information through their word of mouth activity and that it was relatively influential. However, they did not believe they delivered the message very

"powerfully". This was consistent with the focus groups results from the prior exploratory stage.

As a final measure of the sub-scales' usefulness, correlations were computed between the sub-scales and a number of constructs that previous research has suggested should be related to people's willingness to engage in word of mouth activity. The antecedent constructs included in the survey are shown in Table 3 and, as can be seen from the Table, all of the correlations were positive, as was expected. The correlations varied from 0.17 to 0.42 but, in all cases, the correlations were significant beyond the one percent level. It seems that the sub-scales also have criterion related validity as they behaved "as expected with measures of other constructs external to the scale itself" (Sweeney and Soutar, 2001, p. 214).

Table 3: Correlations between Selected constructs and the Word of Mouth Sub-Scales

Antecedent Construct	Emotional intensity of message delivery	Cognitive quality of message delivery	Credibility of the message
The provider exceeded my expectations	0.23	0.30	0.23
The provider gave consistent quality service	0.19	0.42	0.37
The provider gave me an enjoyable interaction	0.19	0.26	0.33
Staff who were interested in me as a person	0.18	0.25	0.34
The provider gave me good value for money	0.17	0.34	0.32
The provider had high quality service	0.19	0.42	0.39
I was satisfied with my interaction	0.18	0.38	0.42
I usually try to use this provider	0.27	0.29	0.24
I am proud to be associated with this provider	0.32	0.29	0.35

Conclusions

The present study developed a short and practical WOM scale that, while built on previous research, recognised that WOM is not a simple functional activity of passing on a comment about a good or service to a private party and, hence, cannot be measured by simply counting the number of occasions when this occurs. The researchers feel the incident in which WOM is given or received should be the basic unit of analysis and explored the WOM construct in this context. The study was able to investigate integral aspects of WOM, such as the richness of the message and the way it was delivered. Three WOM dimensions were identified, which were termed the emotional intensity of the message delivery, the cognitive quality of the message delivery and the credibility of the message. While the latter is consistent with Gilly et al.'s (1998) and Bansal and Voyer's (2000) measure relating to WOM influence, the former two reflect the richness of the delivery process, an aspect new to the literature. The resulting 13-item scale, which was reliable and valid, should be helpful to researchers examining WOM and to managers who are interested in measuring customers' WOM. The scale may also be appropriate when interviewing consumers who have received WOM (i.e. describing the message that someone gave them in a given incident), although this is a topic for future research.

References

Anderson, E. W., 1998. Customer satisfaction and word of mouth. Journal of Service Research 1 (1), 5-17.

Bagozzi, R. P., Foxall, G. R., 1996. Construct validation of a measure of adaptive-innovative cognitive styles in consumption. International Journal of Research in Marketing 13 (3), 201-213.

Bansal, H. S., Voyer, P. A., 2000. Word-of-mouth processes within a service purchase decision context. Journal of Service Research 3 (2), 166-177.

Byrne, B. M., 2001. Structural Equation Modeling with Amos: Basic Concepts Applications, and Programming, Mahwah, NJ, Lawrence Erlbaum Associates.

Churchill, G. A., Jr., 1979. A paradigm for developing better measures of marketing constructs. Journal of Marketing Research 16 (1), 64-73.

Day, G. S., 1971. Attitude change, media and word of mouth. Journal of Advertising Research 11 (6), 31-40.

Ennew, C. T., Banerjee, A. K., Li, D., 2000. Managing word of mouth communication: Empirical evidence from India. The International Journal of Bank Marketing 18 (2), 75-83.

File, K. M., Cermak, D. S. P., Prince, R., Alan, 1994. Word-of-mouth effects in professional services buyer behaviour. The Service Industries Journal 14 (3), 301-314.

Fornell, C., Larcker, D. F., 1981. Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research 18 (2), 39-50.

Gilly, M. C., Graham, J. L., Wolfinbarger, M. F., Yale, L. J., 1998. A dyadic study of interpersonal information search. Journal of the Academy of Marketing Science 26 (2), 83-100.

Hair, J. F. J., Anderson, R. E., Tatham, R. L., Black , W. C., 2006. Multivariate Data Analysis (Sixth ed.), Upper Saddle River, New Jersey, Prentice Hall.

Harrison-Walker, L. J., 2001. The measurement of word-of-mouth communication and investigation of service quality and customer commitment as potential antecedents. Journal of Service Research 4 (1), 60-75.

Hartline, M., Jones, K. C., 1996. Employee performance cues in a hotel environment: Influences on perceived service quality, value and word of mouth intentions. Journal of Business Research 35 (3), 207-215.

Kaiser, H. F., 1974. An index of factorial simplicity. Psychometrika 39, 31-36.

McDonnell, F. (2005, 6-7 October). *Integrating word of mouth marketing.* Paper presented at the International Word-of-mouth Marketing Conference, Hamburg, Germany.

Murray, K., 1991. A test of services marketing theory: Consumer information acquisition activities. Journal of Marketing, 55 (January), 10-25.

Parasuraman, A., Zeithaml, V. A., Berry, L. L., 1988. SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. Journal of Retailing 64 (Spring), 12-40.

Sweeney, J. C., Soutar, G. N., 2001. Consumer perceived value: The development of a multiple item scale. Journal of Retailing 77 (2), 203-220.

Thomas, R. W., Ryan, M. M., Soutar, G. N., 2001. The selling orientation-customer orientation (S.O.C.O.) scale: A proposed short form. Journal of Personal Selling & Sales Management 21 (1), 63-69.

Verbeke, W., Bagozzi, R. P., 2002. A situational analysis on how salespeople experience and cope with shame and embarrassment,. Psychology and Marketing 19 (9), 713-741.

Westbrook, R. A., 1987. Product/consumption-based affective responses and postpurchase processes. Journal of Marketing Research 24 (3), 258-270.