

**Measuring word-of-mouth activity with  
recommendation items in service research:  
What is captured and what is lost?**

Magnus Söderlund, Stockholm School of Economics, Sweden  
Magnus.Soderlund@hhs.se

Jan Mattsson, Roskilde University, Denmark  
mattsson@ruc.dk

**Abstract**

Word-of-mouth activity (WOM) is a common variable in service research, particularly in research on customers' reactions to service encounters. Often, however, WOM is conceptualized and operationalized in terms of recommendations, which is only one among several aspects of WOM. Therefore, we argue that a focus on recommendations is likely to misrepresent what a customer says to other customers after a service encounter. Our empirical study found support for this argument: we found that telling what had happened in a service encounter, with no explicit recommendations, was the dominant part of WOM. We also found that customer satisfaction, generally considered a main determinant of WOM (when it is narrowly defined as recommendations), was a poor predictor of WOM beyond recommendations.

Keywords: Word-of-mouth, service encounters, customer satisfaction

## Introduction

Word-of-mouth activity – defined as the customer’s informal communications directed at other customers about the ownership, usage, or characteristics of particular goods and services and/or their sellers (de Matos and Rossi, 2008) – is a common variable in the service literature. Typically, it is modeled as determined by an overall evaluation variable, such as customer satisfaction, following a customer’s service encounter. In many models it is also the main dependent variable, together with loyalty, in a chain of cause and effect variables comprising several intermediate variables. Intuitively, and in general, a final dependent variable in a model would require a carefully crafted conceptualization and operationalization. This variable, after all, represents the final outcome covered by the model and is thus particularly important for answering the “so-what-questions” every researcher must keep in mind. In the case of service researchers who include word-of-mouth activity in their models, however, the dominant practice does not appear to mirror a detailed concern for word-of-mouth activity. Instead, this activity is typically defined and measured in terms of the intent to provide *recommendations* to other customers (e.g., with specific questionnaire items such as “How likely is it that you would recommend XYZ to a friend?”). Incidentally, Reichheld (2003) has even suggested that this particular aspect – recommendation intent – is the only information the marketer would need about customers.

We do not deny that recommendations are an important aspect of word-of-mouth activity. Yet we argue that recommendations, by definition (cf. the definition above), is only *one* among several aspects. In other words, when customers talk to each other about service experiences, we believe that they do not restrict their conversations to recommendations. Because of the prevalent use of recommendation items in empirical studies, however, very little is known about the extent to which such items are able to capture also other aspects of what goes on when one customer decides to communicate with another customer. The purpose of this paper, therefore, is to explore how a typical recommendation item behaves in relation to measures which do not restrict word-of-mouth activity to explicit recommendations. An exploration of this type, we believe, would serve the purpose of informing service researchers about what is actually captured and what is lost when a recommendation item becomes selected as an indicator of word-of-mouth activity. In other words, then, we explore the validity of the recommendation type item, and we do so in a context of the service encounter and with an experimental approach.

## Frame of reference

A main premise in this paper is that word-of-mouth activity following a service encounter can comprise also other aspects than explicit recommendations. Some researchers acknowledge the potential for word-of-mouth in such terms when they view word-of-mouth activity as sharing thoughts about a commercially-based experience (Söderlund and Rosengren, 2007). Obviously, sharing thoughts may not necessarily be accompanied by recommendations. What, then, is meant by sharing “thoughts”?

In this paper, we assume that a common way to share thoughts after a service encounter is to make statements about what happened in the encounter in terms of descriptions of causally and temporally related sequences of events. Such “what-happened-statements” are also likely to be presented in terms of a story about what took place. Indeed, a story format seems to be a natural way to organize one’s experiences (Adaval and Wyer, 1998; Escalas, 2004), and we believe that it is also a natural way of informing others about one’s experiences. Indeed, we

believe that recommendations, if they occur, typically are embedded in such stories and thus that a recommendation per se, devoid of a story framework, is relatively rare. A story should be understood as a narrative in which one particular person (in our case: the sender of w-o-m) is foregrounded and in which he or she interacts with other persons during an event or the succession of events (Deighton et al, 1989; Stern, 1994). It can be noted that “what-happened-statements” are not only different from a recommendation in terms of their thought content; they are also different in terms of the implications for the sender’s future relation to the receiver. More specifically, an explicit recommendation to carry out an act, or to avoid carrying out an act, may or may not be followed by the receiver. And if the receiver decides to follow the recommendation, it involves a risk that the receiver will be disappointed – which would reflect back negatively on the sender. “What-happened-statements” devoid of explicit recommendations may of course convey recommendations in an implicit way, yet they appear to involve a lower social risk of causing harm to the future relationship with the receiver. Thus, we hypothesize that “what-happened-statements” represent a greater part of word-of-mouth activity than recommendations (Hypothesis 1). In a service encounter context, this means that a research focus only on explicit recommendations is likely to underestimate the extent to which other word-of-mouth activities take place.

Furthermore, copious studies show that customer satisfaction is positively and significantly associated with the intent to provide recommendations (e.g., Brown et al, 2005; Mooradian and Olver, 1997). However, would an overall evaluation variable, such as customer satisfaction, also be a causally potent factor in the case of “what-happened-statements”? Very little is known about this. In general, however, it seems as if incongruent aspects of an event encountered by the customer (i.e., the extent to which the event deviates from the “exemplar” or prototype of the object in the customer’s mind; cf. Sujan 1985) calls for more information processing activity (Lee and Schumann, 2004) and thus incongruent aspects may elicit a particularly strong need to tell others about what happened. Indeed, the act of telling the story may serve to purpose of resolving incongruity. It can also be noted that a story worth telling, if it is to be entertaining for the receiver, needs to contain incongruity – in the sense that it needs to build tension and disequilibrium in relation to an initial state of stability (Laurence, 2007). In the light of this, it seems as if the sender’s level of customer satisfaction may be relatively less important than the level of incongruence when it comes to enhancing conversations with others about what happened. Therefore, we hypothesize that customer satisfaction is not producing “what-happened-statements” to the same extent as it is producing recommendations (Hypothesis 2). We also hypothesize that service encounter incongruity is a more causally potent factor than customer satisfaction for enhancing “what-happened-statements” (Hypothesis 3).

### **Research method**

*General design, data collection, and participants.* We used an experimental between-subjects design to generate data on customer satisfaction and word-of-mouth activity (recommendations and “what-happened-statements”) following a service encounter. The service encounter stimulus was a text-based scenario in which the participant was asked to assume the role of a passenger who interacted with a stewardess aboard an aircraft. Such scenario-based treatments appear frequently in the service literature (e.g., Bitner, 1990; Murray, 1991; Ueltschy et al, 2002), and one fundamental advantage is that they allow for a systematic manipulation of variables and contexts that cannot be easily studied in a real-life setting or are difficult to manipulate in other ways in experiments. Two factors were manipulated in the present experiment: the level of technical service quality provided by the

stewardess (poor vs. good) and the stewardess' emotional display (negative vs. positive). The manipulation of these two factors, we assumed, would create a setting with variation in both satisfaction and incongruity (the combinations good service-negative display and poor service-positive display, we assumed, would create a relatively higher level of incongruity compared to the other two combinations). The scenario text and the subsequent questions to the participants were presented on a computer screen and the responses to the questions were recorded by the computer. The participants ( $N = 400$ ), who we randomly allocated to one of the four treatment groups, belonged to a panel of respondents used by a commercial marketing research firm with which we cooperated.

*Measures.* We measured customer satisfaction with three items used in several national satisfaction barometers (Fornell, 1992; Johnson et al, 2001), namely: "How satisfied or dissatisfied are you with this airline?" (1 = very dissatisfied, 10 = very satisfied), "To what extent does it meet your expectations?" (1 = not at all, 10 = totally), and "Imagine an airline that is perfect in every respect. How near or far from this ideal do you find this airline?" (1 = very far from, 10 = can not get any closer). Cronbach's alpha for this scale was .93. We measured incongruity in terms of the stewardess' behavior in the service encounter; we used a set of adjective pairs (normal behavior-not normal behavior, easy to understand-difficult to understand, and usual-uncommon) scored on 10-point scales ( $\alpha = .89$ ). The recommendation aspect of word-of-mouth activity was measured with an intention item used in many assessments of service customers' post-consumption reactions (e.g., Bitner, 1990; Brown et al, 2005; Mooradian and Olver, 1997): "How likely is it that you would recommend this airline to a friend?". The responses were scored on a scale ranging from 1 (very unlikely) to 10 (very likely). The "tell-what-happened" aspect of word-of-mouth activity was measured by an intention item phrased as follows: "How likely is it that you would tell a friend about what happened?". The responses were scored on a scale ranging from 1 (very unlikely) to 10 (very likely). Moreover, the intention to tell-what-happened item was followed by an open-ended question designed to capture the specific content of what the participant would say; here, we thus asked the participants what they would say to friends after the service encounter. The responses were coded in terms of the following categories of statements, in such a way that each participant's statements were allocated to one and only one of these categories: (1) no statements, (2) statements about what happened without any explicit recommendations (e.g., "The service is bad and the personnel is unpleasant" and "Tell about the good service I received"), (3) statements with explicit recommendations of the positive type (e.g., "You just have to try this company" and "I would recommend them"), and (4) statements with explicit recommendations of the negative type (e.g., "Never use this airline!" and "Don't travel with this company").

### Analysis and results

We hypothesized (H1) that word-of-mouth activity after a service encounter contains more "what-happened" content without recommendations than explicit recommendations content, and this was assessed in two ways. First, we computed the level of each of these two aspects of word-of-mouth activity as they were captured by our two intention items. This revealed that the intent to tell what happened reached a higher level ( $M = 8.07$ ) than the explicit recommendation intent ( $M = 4.66$ ). This difference was significant ( $t(399) = 16.01, p < .01$ ). It can also be noted that the zero-order correlation between these two word-of-mouth activity variables was modest ( $r = .09$ ), suggesting that the two measures tap into two different constructs. Second, we used the responses to the open-ended item (i.e., the participants' statements about what to say to friends after the service encounter) to assess explicit

recommendation content versus content without explicit recommendations. In this case, the number of participants who contributed with statements of different types was used as an indicator of the extent to which different types of word-of-mouth activity would take place. The frequencies are presented in Table 1:

**Table 1**  
**Frequencies of participants who provided the different statement types**

No explicit recommendation content	212
Explicit positive recommendation content	14
Explicit negative recommendation content	63
<hr/>	
Total number of participants providing statements	289

Given the 400 participants in our study, it can be noted that 111 participants did not contribute a statement. The open-ended measure thus indicates that roughly 25 percent would not engage in any word-of-mouth activity. Of the 289 participants who did produce statements, however, Table 1 reveals that the great majority (i.e., 212 participants) provided statements *without* any recommendation content. The observed frequency distribution in Table 1 is significantly different from the expected frequencies given that each statement type has an equal share of the total number of statements ( $\chi^2(3) = 214.30, p < .01$ ). It can be noted that negative recommendation statements were relatively more frequent than positive recommendation statements, which can be seen in the light of a negativity bias observed in some word-of-mouth studies. Taken together, then, our analyses indicate that providing an explicit recommendation is only one part (and a relatively less salient part compared to no recommendation) of word-of-mouth activity, so H1 was supported.

In the light of the prevalent use of recommendation intent items in service research, one may ask to what extent such items are able to distinguish between customers who make statements in the categories presented in Table 1. To examine this issue, we computed the mean recommendation intent for the participants in each of the three statement groups (and in the group who did not make any statement). The resulting means were as follows: 4.73 for the no statement group, 5.06 for the statements without recommendation group, 2.11 for the negative recommendation group, and 9.57 for the positive recommendation group. A one-way ANOVA indicated that all means were not equal ( $F(3,396) = 26.17, p < .01$ ). Moreover, a post hoc pairwise comparison with the Sheffé test showed that the group means were significantly different between all pairs ( $p < .01$  in each case), with one exception: the difference between the group with no statements ( $M = 4.73$ ) and the group with statements without explicit recommendations ( $M = 5.06$ ) was not significant ( $p = .85$ ). This thus implies that the recommendation intent item was unable to discriminate between those who would not say anything at all after a service encounter and those that would tell what happened without any recommendations. Yet in all other cases, the recommendation intent item was indeed able to reflect the content of what would be said.

Moreover, we used a structural equations modeling approach (with AMOS 7.0) to assess if customer satisfaction after a service encounter is more strongly associated with recommendations than with “what-happened-statements” (H2), and to assess if the relative impact on word-of-mouth activity of (a) perceived incongruency in what happens to a customer vis-à-vis (b) customer satisfaction (H3). The proposed model consisted of four links;

satisfaction and incongruence were the independent variables, and each of these two variables was modeled as linked to both recommendation intent and tell-what-happen intent. The proposed model showed a good fit with the data ( $\chi^2 = 97.88$ ,  $df = 24$ ,  $p < .01$ , CFI = .98, NFI = .97, RMSEA = .09). The coefficients for the associations are presented in Table 2.

**Table 2:**  
**Standardized path coefficients in the proposed model**

Link 1: Satisfaction → Recommendation intent	.73***
Link 2: Satisfaction → Tell-what-happened intent	.15*
Link 3: Incongruency → Recommendation intent	.07
Link 4: Incongruency → Tell-what-happened intent	-.26***

\*\*\*  $p < .01$ , \*  $p < .1$

The proposed model thus indicates that the satisfaction–recommendation intent link (Link 1) was stronger than the link between satisfaction and tell-what-happened intent (Link 2). To assess this difference explicitly, we set up an alternative model in which Link 1 and Link 2 were constrained to be equal. This alternative model, however, produced a significantly lower level of fit than the proposed model ( $df = 1$ , delta  $\chi^2 = 52.31$ ,  $p < .01$ ). We therefore conclude that satisfaction had a stronger impact on recommendation intent and thus that H2 was supported. Moreover, the associations in the proposed model suggest that incongruency had a stronger impact on tell-what-happened intent (Link 4) than what satisfaction had (Link 2). An additional alternative model, in which we constrained these two links to be equal, was characterized by a significantly lower level of fit than the proposed model ( $df = 1$ , delta  $\chi^2 = 7.67$ ,  $p < .01$ ), so it can be contended that the impact on tell-what-happened intent of incongruency was stronger than the impact of satisfaction. H3 was thus supported. It should be noted that the differences in the associations between our two indicators of word-of-mouth activity thereby suggest that the two aspects should be dealt with as discrete constructs.

### Discussion

One main finding in our study was that an explicit recommendation is a relatively modest part of word-of-mouth activity; the main part of the “text” provided by the senders in our study did not contain recommendations. This means that service researchers’ prevalent use of measurement items focusing on the recommendation aspect provides a distorted picture of the content of customers’ informal communications with each other. Moreover, our finding that customer satisfaction was only weakly associated with tell-what-happened intent suggests that more work remains to be done for the marketer who wishes to encourage word-of-mouth activity: this marketer should not expect that customer satisfaction is a sufficient factor. Given that incongruency boosts tell-what-happened intent, the service marketer who wants more word-of-mouth activity may thus need to consider ways to deliberately include incongruency in service encounters. Recent research in an advertising context indeed indicates that incongruency may have several other positive effects (cf. Dahlén et al 2008). An additional question that remains to be examined, of course, is if recommendations or telling what happened has the same impact on the receiver. It should not be taken for granted that an explicit recommendation is more efficient than telling what happened, because a story-telling format has been shown to be persuasive in a communications context (Adaval and Wyer, 1998; Deighton et al, 1989).

## References

- Adaval, R. and Wyer, R.S. 1998. The role of narratives in consumer information processing. *Journal of Consumer Psychology* 7 (3), 207-245.
- Bitner, M.J., 1990. Evaluating service encounters: The effects of physical surroundings and employee responses. *Journal of Marketing* 54 (April), 69-82.
- Brown, T.J., Barry, T.E., Dacin, P.A. and Gunst, R.F., 2005. Spreading the word: Investigating antecedents of consumers' positive word-of-mouth intentions and behaviors in a retailing context. *Journal of the Academy of Marketing Science* 33 (2), 123-138.
- Dahlén, M., Rosengren, S., Törn, F. and Öhman, N., 2008, Could Placing Ads Wrong Be Right?, *Journal of Advertising* 37 (3), 57-67.
- Deighton, J., Romer, D., and McQueen, J., 1989. Using drama to persuade. *Journal of Consumer Research* 16 (December), 335-343.
- Escalas, J.E., 2004. Narrative processing: Building consumer connections to brands. *Journal of Consumer Psychology* 14 (1-2), 168-180.
- Fornell, C., 1992. A national satisfaction barometer: The Swedish experience. *Journal of Marketing*, 56 (January), 6-21.
- Johnson, M. D., Gustafsson, A., Andreassen, T.W., Lervik, L. and Cha, J., 2001. The evolution and future of national satisfaction index models. *Journal of Economic Psychology*, 22 (2), 217-245.
- Laurence, J., 2007. *Writing crime fiction*, Studymates, UK.
- Lee, E.J. and Schumann, D.W., 2004. Explaining the special case of incongruity in advertising: Combining classic theoretical approaches. *Marketing Theory* 41 (1/2), 59-90.
- de Matos, C.A. and Rossi, C.A.V., 2008. Word-of-mouth communications in marketing: A meta-analytic review of the antecedents and moderators. *Journal of the Academy of Marketing Science* 36, 578-596.
- Mooradian, T.A. and Olver, J.M., 1997. "I can't get no satisfaction: The impact of personality and emotion on postpurchase processes. *Psychology and Marketing* 14 (4), 379-393.
- Murray, K.N., 1991. A test of services marketing theory: Consumer information acquisition activities. *Journal of Marketing* 55 (January), 10-25.
- Reichheld, F.F., 2003. The one number you need. *Harvard Business Review*, December, 46-54.
- Stern, B.B., 1994. Classical and vignette television advertising dramas: Structural models, formal analysis, and consumer effects. *Journal of Consumer Research* 20 (March), 601-615.

Sujan M., 1985. Consumer knowledge: Effects on evaluation strategies mediating consumer judgments, *Journal of Consumer Research* 12 (June), 31-46.

Söderlund, M. and Rosengren, S., 2007. Receiving word-of-mouth from the service customer: An emotion-based effectiveness assessment. *Journal of Retailing and Consumer Services* 14 (2), 123-136.

Ueltschy, L.C., Laroche, M., Tamilia, R.D., and Yannopoulos. P., 2002. Cross-cultural invariance of measures of satisfaction and service quality. *Journal of Business Research* 57, 901-212.