Perceived helpfulness of online consumer reviews: The role of message content and style

ROBERT M. SCHINDLER[1]* and BARBARA BICKART[2]
[1]*School of Business, Rutgers University–Camden, 227 Penn St., Camden, NJ, 08102, USA*
[2]*School of Management, Boston University, 595 Commonwealth Ave., Boston, MA 02215, USA*

ABSTRACT

The rise of online reviews written by consumers makes possible an examination of how the content and style of these word-of-mouth messages contribute to their helpfulness. In this study, consumers are asked to judge the value of real online consumer reviews to their simulated shopping activities. The results suggest the benefits of moderate review length and of positive, but not negative, product evaluative statements. Non-evaluative product information and information about the reviewer were also found to be associated with review helpfulness. Stylistic elements that may impair clarity (such as spelling and grammatical errors) were associated with less valuable reviews, and elements that may make a review more entertaining (such as expressive slang and humor) were associated with more valuable reviews. These findings point to factors beyond product information that may affect the perceived helpfulness of an online consumer review.
Copyright © 2012 John Wiley & Sons, Ltd.

Word-of-mouth (WOM) communication—the exchange of information about goods and services among consumers—has long been recognized as a valued and influential source of consumer information (e.g., Whyte, 1954). The Internet has dramatically increased WOM communication, particularly in the form of consumer reviews on retailing websites. As of 2004, it was estimated that there were over 10 million consumer reviews on Amazon.com alone (Harmon, 2004). The growth in online consumer reviews has been motivated by consumer interest in such reviews (Kumar and Benbasat, 2006) and has benefitted online retailers with increased customer loyalty and lower costs, such as for returned products (Voight, 2007). It has been proposed that consumers who post reviews serve as "sales assistants" for online retailers (Chen and Xie, 2008), who provide other consumers with useful information, and who contribute to other consumers' satisfaction with the shopping experience.

The goal of the study reported here was to better understand the factors that make online reviews appealing to consumers. Although there is recent research on the effects of consumer reviews on product sales (e.g., Godes and Mazylin, 2004; Chevalier and Mayzlin, 2006, Liu, 2006; Forman *et al.*, 2008; Zhu and Zhang, 2010), we focus on the characteristics of online reviews that shoppers find helpful and of value. Although there is older research on "source effects" (e.g., Hovland and Weiss, 1951; McGuire, 1969; Brown and Reingen, 1987; Wilson and Sherrell, 1993), we take advantage of the fact that this modern consumer WOM is expressed in written form and look at the wording of consumer reviews rather than their source. We divide wording factors into two categories, content and style. Greater understanding of these factors can help guide managers of retail websites toward encouraging reviews that their patrons will

find more useful and may also shed some light on the nature of the long-acknowledged power of WOM communication.

CONCEPTUAL FRAMEWORK

The distinction between the content and the style of a message is one that is well established in the field of communication research (e.g., Norton, 1978). The content-versus-style distinction has been used in studying the effectiveness of personal selling communications (Williams and Spiro, 1985; Dion and Notarantonio, 1992), in communications regarding product information (Moon, 2002), and in understanding how people use electronic communications such as email (Colley and Todd, 2002). For the purposes of the present study, we define the *content* of an online review as the information it provides. Its *style*, by contrast, involves the choice of words the reviewer uses to express this information.

Content factors

To begin an examination of how the content of an online review can affect its perceived helpfulness, it is necessary to have a means of characterizing a review's content. Past efforts to characterize the content of a WOM message have tended to consider each WOM message as a whole. For example, Granitz and Ward (1996) looked at the types of comments made in online WOM messages (in a bulletin board discussion group) and found that the most common comment types were "recommendations," "how-to advice," and "explanations." Richins and Root-Shaffer (1988) classified recalled WOM comments as "positive personal experience," "advice-giving," "product news," or "negative word-of-mouth." Mangold *et al.* (1999) classified recalled WOM incidents as concerning "quality only," "price-only," or "value." Other research has examined the number of reviews for a product or the dispersion of these reviews among writers of, say, different gender or age, but have characterized the content of an individual review by

*Correspondence to: Robert M. Schindler, School of Business, Rutgers University–Camden, 227 Penn St., Camden, NJ 08102, USA.
E-mail: rschindl@camden.rutgers.edu

only general measures of valence (e.g., Godes and Mazylin, 2004; Liu, 2006; Dellarocas *et al.*, 2007) or by length (Mudambi and Schuff, 2010).

Our goals were to develop a means of characterizing WOM content that can be applicable in all types of WOM situations and that can make possible the observation of effects of content elements that are incidental to the main purpose of the WOM communication. Our approach is to consider an online review (as one could consider any type of WOM communication) to be a "verbal protocol" that can be divided into component statements (Ericsson and Simon, 1993). A *component statement* is defined as a group of words that comprise a single thought. The content of the review could then be characterized by the number and types of these component statements.

Our typology of component statements is based on whether the statement contains product evaluation words, such as "like," "dislike," "best," "worst," "high quality," and "low quality." Statements containing such words are considered to be *product evaluative* statements and can be further classified as either positive or negative in valence. Statements that do not contain product evaluative information are considered to be *descriptive* statements and can be further classified as either concerning the product or concerning the reviewer. This classification procedure results in four statement categories:

Product evaluative statements:

1. Positive evaluative statements
2. Negative evaluative statements

Descriptive statements:

3. Product-descriptive statements
4. Reviewer-descriptive statements

Statements that do not fit into any of these four categories (e.g., a digression about politics) are placed into a fifth category described simply as "other."

Number of statements in a review

Before the types of statements in a review are discussed, there could be some consideration of how the total number of statements in a review might affect the review's usefulness. On one hand, longer reviews may have more details and specifics and thus may provide more diagnostic information pertinent to the consumer's decision. Supporting this idea, Mudambi and Schuff (2010) found that longer reviews were associated with higher review helpfulness ratings. On the other hand, in the context of a WOM message, it may be possible for a review to be too long. One of the implicit rules that govern conversations is the *maxim of quantity*—the convention that speakers provide only the information that is required to make a point (Grice, 1975). Providing more information than is necessary violates this maxim and may lead to confusion (Schwarz, 1996). For example, a review that repeatedly claims that a hotel is in a very safe neighborhood might lead readers to question the safety of

the neighborhood (i.e., why do they keep mentioning safety?). Thus, we hypothesize:

> *H1:* A greater number of statements in a review will be associated with high-value reviews, but only up to a point.

Positive and negative evaluative statements

Research on the effects of the number and valence of product evaluative statements in an online review has focused on whether consumer reviews of extreme valence are found to be more helpful than equivocal reviews. This research has produced conflicting results (Forman *et al.*, 2008; Mudambi and Schuff, 2010). In the context of our study, a review with a high proportion of positive statements could be helpful to readers in that it makes the case that the product should be considered further. However, Sen and Lerman (2007) found that, at least for some products, consumers find a negative review more accurate, informative, and useful than a positive one. Further, considering the benefits of a two-sided appeal in advertising (e.g., Pechmann, 1992), it might be most useful to avoid an excessive number of either positive or negative statements. Thus, we hypothesize:

> *H2:* A greater proportion of positive evaluative statements in a review will be associated with high-value reviews, but only up to a point.

> *H3:* A greater proportion of negative evaluative statements in a review will be associated with high-value reviews, but only up to a point.

Product-descriptive statements

Our typology of review statements makes possible an examination of the helpfulness of statements that are merely descriptive of the product. Because relevant factual information presented without evaluation gives a shopper the means to make his or her own evaluation, it is hypothesized that many product-descriptive statements would also tend to enhance the value of an online review. In general, however, evaluative statements should be more useful than descriptive statements because descriptive information should be easily available from other sources (e.g., the package or marketing information), whereas evaluative statements would not be available from these sources and hence could be more informative for consumers making a decision. Thus, we hypothesize:

> *H4:* A greater proportion of product-descriptive statements in a review will be associated with high-value reviews, but only up to a point.

Reviewer-descriptive statements

The presence in a review of personal statements about the reviewer could serve several useful functions. First, personal statements could help the reader to better understand the reviewer's perspective. This possibility is consistent with the finding of Forman *et al.* (2008) that, for reviews in which product ratings were equivocal, information about the reviewer's real name and location was considered by readers

to be helpful. Other functions of personal statements are that they could help to establish the authority or knowledge base of the reviewer (Alon and Brunel, 2006) and/or could help make the writer more likeable to the reader (Walther *et al.*, 2005). On the other hand, it has also been found that as information about another person increases, the liking for that person tends to decrease, presumably because the increased amount of information causes more dissimilarities to become apparent (Norton *et al.*, 2007). Also, a large number of reviewer-descriptive statements could be interpreted as a reviewer's undesirable digression into self-expression or self-promotion. Thus, we hypothesize:

> *H5:* A greater proportion of reviewer-descriptive statements in a review will be associated with high-value reviews, but only up to a point.

Style factors

In an online WOM situation (compared with a face-to-face context), there are limited cues available for expressing communication style. The burden of style expression is carried by the communicator's specific wording choices, and thus, these choices are likely to have a strong effect on the impressions created by an online message (Hancock and Dunham, 2001). As we characterized a review's content in terms of its component statements, we characterize a review's style in terms of the specific wording choices made within each of the review's statements. We coded the statements in a review for a variety of stylistic wording choices, some of which could be expected to be associated with low strength or impact of the statement (e.g., bad grammar, inexpressive slang) and some which could be expected to be related to high statement strength or impact (e.g., first-person pronouns, emotion words).

Style factors related to weaker impact

The *maxim of manner* suggests that people will communicate clearly and recognize that their contribution to the conversation must be understood (Grice, 1975). If a communication is difficult to understand, the reader is likely to make negative inferences about the communicator. Thus, it is expected that the occurrence of style variables that reduce the readers' ability to comprehend a review will be associated with less valuable reviews. Such variables include misspellings, bad grammar, the use of inexpressive slang, the use of qualifications, and repetition. Consistent with this prediction, Jessmer and Anderson (2001) found that people writing grammatical email messages (versus ungrammatical messages) were perceived as more likeable and more competent. Thus, we hypothesize:

> *H6:* A greater use of negative style characteristics in a review will be associated with low-value reviews.

Style factors related to stronger impact

Other stylistic variables may help increase the usefulness of an online review. For example, the use of expressive slang and humor may evoke an informal tone that would help the reader feel a connection with the reviewer (Fraley and Aron,

2004). The use of emotion-laden words may convey excitement and enthusiasm (or disdain and dissatisfaction) about a product, which could simplify the decision process and thus be helpful to the reader. The use of first-person pronouns and the presence of personal information worded in a self-effacing way could give a review a sense of authenticity in the mind of the reader and lead the reader to feel a greater confidence about the information provided by the reviewer. Thus, we hypothesize:

> *H7:* A greater use of positive style characteristics in a review will be associated with high-value reviews.

Approach of the study

The goal of this study was to explore these issues by using naturally occurring online reviews, which makes appropriate the use of a quasi-experimental design. Specifically, study participants were given a decision scenario and were asked to locate and list relevant consumer reviews for making this decision. They were then asked to select and print two of these reviews: (i) the review that they felt was the most valuable for making this decision and (ii) the review that they felt was the least valuable for making this decision. This procedure of having each participant select both a valuable and not-valuable review was designed to control for differences between participants in overall preferences—for example, some people might generally prefer long reviews whereas others might prefer short reviews. By having each participant produce a most-valuable *and* a least-valuable review, we obtain an indication of the factors that, to the participant, contribute to making a review helpful to a decision process.

The reviews collected in this way were then divided up into component statements, and each statement was coded for aspects of content and style. The coded reviews were analyzed so as to test for a relationship between content and style characteristics and the perceived value of the review (high versus low). This design made possible an examination of naturally occurring reviews with some degree of research control.

METHOD

Task

Participants were asked to examine real online consumer reviews for two product categories—books and automobiles. These are two categories for which online reviews are prevalent and appear to be used by consumers.

The participants were asked to examine these reviews in the context of a decision scenario that asked them to assume that they were looking for a product that would meet a specific goal. Two decision goals were used for each product category. For example, for the book category, one scenario involved looking for a fun book to read and the other involved looking for a book about financial planning.

In each scenario, the participant was also given the names of several items that were under consideration (e.g., books by Stephen King, Michael Crichton, etc.). In addition, the scenario listed a number of websites that provided consumer

reviews on the items that they would be considering (e.g., Amazon.com, bn.com). This task was designed to simulate an actual decision-making situation in a way that would be engaging and realistic to the participants. The wording of the decision scenarios given to participants can be seen in Appendix A.

The participants were told to look for consumer comments or reviews that were relevant to the decision scenario and consisted of at least 25 words. Reviews with fewer than 25 words generally contained very little substance (e.g., "Great book, buy it, buy it!"). The instruction sheet included a place to list URL addresses of the comments examined, the name of the person who posted the comment, and a rating of the value of the comment (on a 5-point scale, where 1 is *not valuable* and 5 is *very valuable*). Participants were told to consult two or more web sites and read at least 10 comments consisting of 25 words or more. An example of a 25-word comment was provided.

After looking for and listing the reviews, participants were asked to think back over all of the comments they had read and to select the one comment that was "most valuable" and also the one comment that was "least valuable." They were told to print out those two comments. They then completed a follow-up questionnaire, which included questions about both the book and the automobile decision scenarios. The order of questions about the product categories was counterbalanced.

Participants

The participants were 42 business school students (both graduate and undergraduate) who received course credit and a chance to win $50 in a drawing for completing the study. All participants looked for reviews for both a book and an automobile, but the order in which they searched for these product categories was systematically varied (car first or book first).

The 42 participants in the study provided 78 pairs of online reviews (a pair consisting of one most valuable and one least valuable). Thirty-six participants gave two pairs of reviews, and six participants provided only one pair of reviews. The review pairs were equally distributed among the two product categories. From here on, the study's participants will be referred to as the reviews' "readers."

Coding of reviews

Each of the reviews in the 78 most-valuable/least-valuable pairs (a total of 156 reviews) was parsed by the researchers into component statements (which from here on will be referred to simply as "statements"), each expressing a single thought. Usually, the text of a review was parsed at punctuation marks, but sometimes sentences and long phrases were judged to contain separate thoughts (e.g., the parsing in the following sentence from a review is indicated by the slash: "I have read the entire series of Patricia Cornwell books where she writes about Dr. Kay Scarpetta/and all of them are excellent reading."). The 156 reviews were parsed into a total of 4901 statements, for an average of 31.4 statements per review.

The characteristics of each of these 4901 statements were determined by two raters, who were naïve to the purposes of the study. The raters first rated each statement separately, then met to work out final ratings for statements where their individual ratings did not agree. The raters' classification of each statement enabled the statements to be classified into one of the five statement types. These five types, along with example statements, can be seen in Table 1. The raters also coded each statement for the presence of 17 wording characteristics that are related, positively or negatively, to wording style. A complete list of wording style variables examined in this study is presented in Appendix B. Agreement between the two coders was 70 per cent for the statement type variables (kappa = 0.605) and averaged 98 per cent for the 17 wording style variables (kappas ranged from 0.153 to 0.927).

RESULTS

What kinds of statements comprise a review?

The occurrence percentages of each of the five statement types, shown in the third column of Table 1, suggest several

Table 1. Description and occurrence of the five statement types

Statement type	Examples	Percent occurrence in all reviews
Positive evaluative statements	"The author does a good job of developing the characters." "This is a really good car."	37
Negative evaluative statements	"Her explanation of nutritional principles was very incomplete." "Given all the problems we've had, I wouldn't recommend this car to anyone."	19
Product-descriptive statements	"Loretta had numerous close calls before finding her lost son, Horace." "This book is longer than the typical romantic novel."	15
Reviewer-descriptive statements	"I have three kids; I usually try to see the positive side of things." "I have owned this product for two months now."	16
Other statements	"It's important to do a lot of price shopping before buying a car." "Most of the other romantic novels that I've read have been overly sentimental."	13
Total		100

interesting findings. First, consistent with earlier research (e.g., Chevalier and Mayzlin, 2006), there were more positive evaluative statements than negative evaluative statements (37% versus 19%; $\chi^2(1) = 980.50$, $p < 0.001$), suggesting that people, in general, are more likely to provide WOM messages when they are generally satisfied or like a product. Second, the writers of these reviews do express opinions, in that there are more evaluative statements (positive or negative) present in the reviews than non-evaluative statements (product-descriptive, reviewer-descriptive, or other) (56% versus 44%; $\chi^2(1) = 94.07$, $p < 0.001$). Third, about one out of every six statements relates to the reviewer (versus the product being reviewed), suggesting that the writers are establishing a context for the review.

How are valuable reviews different from less valuable reviews?

To determine factors that might predict the value of consumer reviews, we compared the overall number of statements and the frequencies of each statement type for the high-value reviews with those from the low-value reviews. We ran a series of logistic regressions to examine these differences systematically. Because of the correlations between reviews due to the study's within-subjects design, the regressions were estimated using generalized estimating equations (Liang and Zeger, 1986). In these models, the dependent measure was the value of the review (coded 1 = *valuable* and 0 = *not valuable*). The predictor variables included both the linear and quadratic components of a possible relationship. The results of these analyses are shown in Table 2.

The regression examining the relationship between the number of statements in a message and its value indicated that both the linear and quadratic terms are significant (linear, $z = 5.38$, $p < 0.001$; quadratic, $z = -3.51$, $p < 0.001$). The negative coefficient of the quadratic term indicates that the relationship is described by an inverted U-shaped function. A greater number of statements are related to an increased likelihood of a review being valuable, but only up to a point, thus providing support for H1.

The regressions examining the relationships between the types of message statements and message value were complicated by this strong relation between the number of statements and message value. Thus, all of the following analyses are based on the proportion (rather than the absolute number) of each statement type among all the statements in a message.

The relation between the proportion of positive evaluative statements in a review and the likelihood that the review is perceived as valuable is also described by an inverted U-shaped function, as indicated by the significant quadratic term ($z = -2.57$, $p = 0.01$) shown in Table 2. This finding suggests that a review could consist of too many or too few positive evaluative statements and thus supports H2.

The same is not true for negative evaluative statements, however. No statistically significant relationship was found between the proportion of negative evaluative statements and message value, thus failing to support H3. (Note that the trend opposed the predicted direction, with the proportion of negative evaluative statements being inversely related to review value, $z = -1.81$, $p = 0.07$.) Given this unexpected result, it could be asked if it is important to have at least some negative information in a review. For example, the presence of even one negative evaluative statement in a review might make a review seem more balanced and thus increase its value. Of the 156 reviews in the data, 45 (29%) had no negative evaluative statements. In contrast to the expectation, these 45 reviews were equally likely to be considered most valuable and least valuable ($\chi^2(1) = 1.09$, $p = 0.30$). Thus, even the presence of any negative evaluative statements did not distinguish between more and less valuable reviews.

Increases in the proportion of product-descriptive statements are related to an increased likelihood that the message is considered valuable, as indicated by a significant linear effect ($z = 2.89$, $p = 0.004$). This finding supports part of H4, but the failure to find a statistically significant quadratic component to the relationship does not support the part of H4 that suggests this relationship would occur only to a point. Unlike evaluative statements that favor a product, there appears to be a greater tolerance for a large proportion of statements that provide product information without evaluation.

We expected that statements about the reviewer would be associated with valuable reviews, but only up to a point. Our findings are consistent with that expectation. Increases in the proportion of reviewer-descriptive statements increase the likelihood of a message being valuable, but only up to a point, as indicated by significant linear and quadratic terms (linear, $z = 2.80$, $p = 0.005$; quadratic, $z = -2.42$, $p = 0.016$). Thus, our results provide support for H5.

How are style characteristics related to review value?

Each review statement was coded for the presence of 17 wording style variables (see Appendix B). Table 3 shows

Table 2. Predicting value of review by number and types of statements

	Regression coefficients	
	Linear term	Quadratic term
Total number of statements in message	2.931[a]	−0.491[a]
Proportion of positive evaluative statements	0.074	−0.465[b]
Proportion of negative evaluative statements	−0.429	−0.034
Proportion of product-descriptive statements	0.872[b]	−0.187
Proportion of reviewer-descriptive statements	0.672[b]	−0.392[c]

[a] $p \leq 0.001$.
[b] $p \leq 0.01$.
[c] $p \leq 0.05$.

Table 3. Incidence of wording style characteristics in reviews

Style characteristic	Percent of reviews containing			
	0 occurrence	1 occurrence	2 occurrences	3+ occurrences
1st-person pronoun	3.2	12.8	8.3	75.7
Proper noun	7.7	7.7	13.5	71.1
Number or date	10.3	17.9	13.5	58.3
Misspelling	35.9	22.4	22.4	19.3
2nd-person pronoun	39.7	19.2	11.5	29.6
Exclamation	47.4	20.5	14.7	17.4
Emotion word	49.4	26.9	7.7	16.0
Qualification	57.1	18.6	14.1	10.2
Bad grammar	59.0	30.1	7.7	3.2
Superlative word	62.8	22.4	7.7	7.1
Imagery	66.7	16.7	7.1	9.5
Slang—inexpressive	71.8	19.9	4.5	3.8
Slang—expressive	72.4	9.0	7.1	11.5
Repetition	76.9	17.9	4.5	0.7
Humor	78.8	13.5	1.9	5.8
Self-effacing wording	89.1	9.6	1.3	0.0
Sequential statements	93.6	1.9	1.9	2.6

the percent of reviews that showed 0, 1, 2, and 3+ occurrences of each of the 17 style characteristics. Statements containing first-person pronouns and proper nouns occurred very often in the reviews. Statements containing self-effacing wording and sequential statements (i.e., the beginning of a narrative) occurred infrequently in reviews. It is likely that narratives occur more frequently in types of online WOM other than product reviews—for example, in postings on online forums or bulletin boards (Alon and Brunel, 2006).

We expected that negative wording style characteristics would be associated with not-valuable reviews (H6), and positive style characteristics would be associated with the reviews that were seen as valuable (H7). Specifically, we expected that five negative style characteristics—misspellings, qualifications (i.e., qualifying phrases that made the point of a statement equivocal), bad grammar, slang-inexpressive (i.e., use of slang to avoid discussion of specifics, e.g., "it sucks"), and repetition—would be more likely to occur in reviews that were perceived as not valuable. As the remaining 12 style characteristics are mostly positive, we expected that they would occur more frequently in reviews that were considered valuable.

These expectations were tested by adding together the number of times a statement in a review was coded as containing a negative or positive style characteristic and dividing the total by the number of statements in the review. For the five negative style characteristics, the values of this index for a review ranged from 0 to 0.75, with a mean of 0.15. Logistic regression indicated that negative style characteristics were indeed associated with less valuable reviews (linear effect, $B = -0.486$, $p < 0.005$), thus supporting H6. For the 12 positive style characteristics, the values of this index ranged from 0 to 1.92 (because a statement with two positive style characteristics would be counted twice), with a mean of 0.93. However, logistic regression analysis indicated that the overall index of positive style characteristics was not significantly associated with the value of a review (linear effect, $B = -0.063$, $p > 0.30$), thus failing to support H7.

To further examine possible associations of the 12 positive style characteristics with review value, we carried out a factor analysis on the 12 characteristics. The first three factors identified in this analysis had eigenvalues of 1.30 or greater; all of the other factors had eigenvalues of 1.165 or less. Considering style characteristics with factor loadings of 0.6 or greater, the first factor, which we termed "informality," consisted of the presence of expressive slang and humor. The second factor, which we termed "strong emotion," consisted of the presence of exclamations and emotion words. The third factor, which we termed "authenticity," consisted of the presence of first-person pronouns and self-effacing wording.

A logistic regression with these three factors as predictors and review value as the dependent variable indicated that "informality" showed a positive association with the value of a review (linear effect, $B = 0.893$, $p < 0.005$; quadratic effect, $B = -0.491$, $p < 0.01$). The significance of both the linear and quadratic terms suggests that some use of expressive slang and humor may contribute to the value of a review, but that a greater amount of such informality may begin to detract from review value. However, for the factors "strong emotion" and "authenticity," there appeared to be no statistically significant positive association with review value (linear effects, $B = -0.227$, $p = 0.09$; $B = -0.007$, $p > 0.4$, respectively).

DISCUSSION

The findings of this study provide a first look at the content and style characteristics that are associated with a helpful online review. We found the length of a consumer review to be positively related to its perceived value to other consumers, but only to a point. Readers may require enough information for an informed decision, but find that too many statements make the review difficult to absorb. Our results are consistent with past findings (e.g., Godes and Mazylin, 2004) that online consumer reviews are more often positive than negative, and

we found that a moderate proportion of positive evaluative statements was associated with helpful reviews. Positive evaluations could support the reader's further consideration of an alternative, but too much positive information might lead the reader to question the reviewer's motives. We also found that a higher proportion of product-descriptive statements was related to review value, suggesting that more "factual" information about products helped readers with their decision-making. Finally, we found that a moderate proportion of statements about the writer of the review was positively related to review value. Statements about the reviewer appear to provide some context within which to interpret the review.

An unexpected result concerning review content was our failure to find a relationship between review value and the proportion of, or even the existence of, negative evaluative statements. A clue to the explanation of this counterintuitive finding might be evidence from depth interviews with online shoppers that negative or mixed online WOM messages are particularly valued when consumers are engaged in ruling out decision alternatives (Schindler and Bickart, 2005). In the simulated decisions of the present study, the readers were asked to begin a shopping process. Thus, they may have used the online reviews mainly to suggest possibilities worthy of further consideration. It may be that few of the readers in the study were in the decision stage of ruling out alternatives. This finding highlights the importance of future research on how the value of review dimensions may vary over the stages of the consumer's decision process.

Our results concerning the style characteristics of reviews support the importance of wording styles in judgments of review value. As expected, negative stylistic elements were associated with less valuable reviews. These may hurt the value of a review by making it more difficult to understand and by undermining the reader's feelings about the reviewer's competence. Reviews that included more informal style characteristics (i.e., more use of expressive slang and humor) were perceived as more valuable, at least up to a point. It is possible that informality can increase the reader's felt connection with the reviewer and make the review more engaging. However, too much informality might interfere with the reader's perception of the reviewer as being intelligent and competent (Levin *et al.*, 1994). In general, there is much potential in examining the use of language in WOM messages, both in terms of when consumers tend to make specific language choices and how those choices affect the impact of the message (for example, see Schellekens *et al.*, 2010).

Implications and directions for future research

The present findings have some theoretical implications concerning WOM communication and raise some meaningful research questions. First, Grice's (1975) principles, which apply to face-to-face communication, also appear to be active in online WOM. The finding of the present study that review value increases with the number of product-descriptive statements in a review, but increases only to a point with the number of positive evaluative statements, is consistent with the maxim of quantity. At some point, too many evaluative statements could become "overkill" (perhaps making it appear that the reviewer is biased), whereas additional descriptive statements may simply provide more useful information. Likewise, the finding that stylistic elements may, in themselves, affect the usefulness of the review is consistent with the maxim of manner. Stylistic elements may affect the reader's perceptions of the communicator and can make the message easier to understand and/or remember. Online reviews differ from conversations or dialogue in that they are written as "one-to-many," rather than one-to-one, communications. Thus, it is interesting that at least some of the maxims of one-to-one conversation appear to still apply. One question for future research is to better delineate the ways in which one-to-many versus one-to-one conversations are different and the ways in which they are similar.

A question raised by this research is the degree to which the content and style factors that we found to be associated with a useful consumer review would also be factors in the usefulness of professional journalistic reviews. Consistent with our findings, journalistic reviews have been found to contain a mixture of descriptive and evaluative elements (Shrum, 1991). Further, at least one journalism textbook has noted the importance of colorful phrasing in a review, quoting a famous drama critic who stated that the work of a critic "is to present his opinions entertainingly and clearly . . ." (Allen, 2005). Although there has been research on the effects of wording-related factors in a professional review, such as whether a critic has a tough or gentle style (D'Astous and Touil, 1999), there has not been empirical confirmation of the value of being entertaining other than our finding regarding the value of expressive slang and humor. Considering both consumer and professional reviews would be a promising direction in future research on wording effects in product reviews.

The results of this study also point to possible limits of the written form of consumer-to-consumer communication. For example, our lack of results for emotion words may be due to peoples' poor ability to detect emotions in written messages such as consumer reviews. Kruger *et al.* (2005) found that readers of email messages were often inaccurate in detecting emotions such as sarcasm, humor, anger, and sadness in email messages. They suggest that written messages are impoverished relative to verbal messages and lack the nonverbal cues necessary for readers to effectively decode emotions in these messages. Future research could examine cues that might help people understand and decode emotions in WOM messages.

Limitations

Although these results provide some initial insight into what makes an online WOM message valuable, there are some weaknesses associated with the procedures used in this study. First, the research focused on the perceived value of the review in a simulated decision context, rather than the actual value of the review in a real decision-making context. The results would be more compelling in a context with actual decision outcomes.

Second, the sample size was relatively small and homogeneous, and there was consideration of only two product categories. Student participants may be more interested in online WOM than older consumers (Yin, 2003), but we

saw no reason why they would value different aspects of the review. We also allowed participants to select high and low value reviews. Although this procedure increased the realism of the study, it obviously reduced control. Our understanding of what comprises review value could be enhanced by a more comprehensive study in which a larger probability sample of online reviews is content coded. The link between coded content and the rated helpfulness of the reviews could be examined. A larger amount of data could lead to more definitive results, particularly for results concerning rarer review characteristics.

Third, it is likely that the relative importance of various content and style factors in predicting review helpfulness may depend on the type of product being reviewed (Mudambi and Schuff, 2010) or the consumers' decision goal (Smith *et al.*, 2005). Understanding the role of these potential moderators is an important direction for future research.

Finally, the method used in this study does not allow us to examine the direction of causality—it is possible to say only that there is an association between certain review characteristics and review value. Future research could confirm the present results by experimentally varying the factors found here to be associated with the perceived value of an online review.

CONCLUSION

The results of this study indicate several aspects of an online review's content and style that distinguish more valuable reviews from those that consumers feel are less valuable. These findings increase our understanding of online reviews and also provide suggestions concerning the broader question of how consumers respond to WOM communication. Although only an early step, this work illustrates that study of online consumer-to-consumer communication can be a potent method of gaining a deeper understanding of the factors that make WOM communication such a powerful influence on the consumer.

ACKNOWLEDGEMENTS

BIOGRAPHICAL NOTES

Robert M. Schindler is Professor of Marketing, Rutgers University–Camden. He is interested in consumer motivation and the psychological processes underlying the effects of marketing techniques. He has published in the areas of price promotions, 9-ending prices, word-of-mouth communication, and the effects of early experience on consumer tastes.

Barbara Bickart is Associate Professor of Marketing, Boston University. She is interested in context effects on judgment and decision-making processes, how people retrieve and use information about others in answering survey questions, and developing methods for reducing measurement error in surveys via questionnaire design.

REFERENCES

Allen R. 2005. The art of reviewing. In Keeble R (ed). *Print Journalism: A Critical Introduction.* London: Routlege; 180–187.

Alon AT, Brunel FF. 2006. Opening the word-of-mouth black box: an analysis of word-of-mouth rhetorical methods in group conversations. Working paper #2006-7. Boston University School of Management, Boston, MA.

Brown JJ, Reingen PH. 1987. Social ties and word-of-mouth referral behavior. *Journal of Consumer Research 14*: 350–362.

Chen Y, Xie J. 2008. Online consumer review: word-of-mouth as a new element of marketing communication mix. *Management Science 54*: 477–491.

Chevalier JA, Mayzlin D. 2006. The effect of word of mouth on sales: online book reviews. *Journal of Marketing Research 63*: 345–354.

Colley A, Todd Z. 2002. Gender-linked differences in the style and content of e-mails to friends. *Journal of Language and Social Psychology 21*: 380–392.

D'Astous A, Touil N. 1999. Consumer evaluations of movies on the basis of critics' judgments. *Psychology and Marketing 16*: 677–694.

Dellarocas C, Zhang M, Awad NF. 2007. Exploring the value of online product reviews in forecasting sales: the case of motion pictures. *Journal of Interactive Marketing 21*: 23–45.

Dion PA, Notarantonio EM. 1992. Salesperson communication style: the neglected dimension in sales performance. *Journal of Business Communication 29*: 63–77.

Ericsson KA, Simon H. 1993. *Protocol Analysis: Verbal Reports as Data*, revised edition. Cambridge, MA: MIT Press.

Forman C, Ghose A, Wiesenfeld B. 2008. Examining the relationship between reviews and sales: the role of reviewer identify disclosure in electronic markets. *Information Systems Research 19*: 291–313.

Fraley B, Aron A. 2004. The effect of a shared humorous experience on closeness in initial encounters. *Personal Relationships 11*: 61–78.

Godes D, Mazylin D. 2004. Using online conversations to study word-of-mouth communication. *Marketing Science 23*: 545–560.

Granitz NA, Ward JC. 1996. Virtual community: a sociocognitive analysis. *Advances in Consumer Research 23*: 163–166.

Grice PH. 1975. Logic and conversation. In Cole P, Morgan JL (eds). *Syntax and Semantics 3: Speech Acts.* New York: Academic Press; 41–58.

Hancock JT, Dunham PJ. 2001. Impression formation in computer-mediated communication revisited: an analysis of the breadth and intensity of impressions. *Communication Research 28*: 325–347.

Harmon A. 2004. Amazon glitch unmasks war of reviewers. *New York Times*.

Hovland CI, Weiss W. 1951. The influence of source credibility on communication effectiveness. *Public Opinion Quarterly 15*: 635–650.

Jessmer SL, Anderson D. 2001. The effect of politeness and grammar on user perceptions of electronic mail. *North American Journal of Psychology 3*: 331–346.

Kruger J, Epley N, Parker J, Ng Z. 2005. Egocentrism over e-mail: can we communicate as well as we think? *Journal of Personality and Social Psychology 89*: 925–936.

Kumar N, Benbasat II. 2006. The influence of recommendations and consumer reviews on evaluations of websites. *Information Systems Research 17*: 425–439.

Levin H, Giles H, Garrett P. 1994. The effects of lexical formality and accent on trait attributions. *Language and Communication 14*: 265–274.

Liang K, Zeger SL. 1986. Longitudinal data analysis using generalized linear models. *Biometrika 73*: 13–22.

Liu Y. 2006. Word of mouth for movies: its dynamics and impact on box office revenue. *Journal of Marketing 70*: 74–89.

Mangold GW, Miller F, Brockway GR. 1999. Word-of-mouth communication in the service marketplace. *Journal of Services Marketing 13*: 73–89.

McGuire WJ. 1969. The nature of attitudes and attitude change. In Lindzey G, Aronson E (eds). *The Handbook of Social Psychology*, 2nd edition. Reading, MA: Addison-Wesley; 136–314.

Moon Y. 2002. Personalization and personality: some effects of customizing message style based on consumer personality. *Journal of Consumer Psychology 12*: 313–326.

Mudambi SM, Schuff D. 2010. What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly 34*(1): 185–200.

Norton RW. 1978. Foundation of a communicator style construct. *Human Communication Research 4*: 99–112.

Norton MI, Frost JH, Ariely D. 2007. Less is more: the lure of ambiguity, or why familiarity breeds contempt. *Journal of Personality and Social Psychology 92*: 97–105.

Pechmann C. 1992. Predicting when two-sided ads will be more effective than one-sided ads: the role of correlational and correspondent inferences. *Journal of Marketing Research 24*: 441–453.

Richins ML, Root-Shaffer T. 1988. The role of involvement and opinion leadership in consumer word-of-mouth: an implicit model made explicit. *Advances in Consumer Research 15*: 32–36.

Schellekens GAC, Verlegh PWJ, Smidts, A. 2010. Language abstraction in word of mouth. *Journal of Consumer Research 37*(August): 207–223.

Schindler RM, Bickart B. 2005. Published word of mouth: referable, consumer-generated information on the Internet. In Haugtvedt CP, Machleit KA, Yalch RF (eds). *Online Consumer Psychology: Understanding and Influencing Behavior in the Virtual World*. Hillsdale, NJ: Lawrence Erlbaum Associates; 35–61.

Schwarz N. 1996. *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Sen S, Lerman D. 2007. Why are you telling me this? An examination into negative consumer reviews on the Web. *Journal of Interactive Marketing 21*: 76–94.

Shrum W. 1991. Critics and publics: cultural mediation in highbrow and popular performing arts. *The American Journal of Sociology 97*: 347–375.

Smith D, Menon S, Sivakumar K. 2005. Online peer and editorial recommendations, trust, and choice in virtual markets. *Journal of Interactive Marketing 19*(3): 15–37.

Voight J. 2007. Getting a handle on customer reviews. *Adweek 48*(26): 16–17.

Walther JB, Loh T, Granka L. 2005. Let me count the ways: the interchange of verbal and nonverbal cues in computer-mediated and face-to-face affinity. *Journal of Language and Social Psychology 24*: 36–65.

Whyte WH, Jr. 1954. The web of word of mouth. *Fortune 50*: 140–143.

Williams KC, Spiro RL. 1985. Communication style in the salesperson–customer dyad. *Journal of Marketing Research 22*: 434–442.

Wilson EJ, Sherrell DL. 1993. Source effects in communication and persuasion research: a meta-analysis of effect size. *Journal of the Academy of Marketing Science 21*: 101–112.

Yin, S. 2003. Degree of challenge. *American Demographics 25*(4): 20–23.

Zhu F, Zhang X. 2010. Impact of online consumer reviews on sales: the moderating role of product and consumer characteristics. *Journal of Marketing 74*: 133–148.

APPENDIX A
WORDING OF DECISION SCENARIOS

Book decision scenario (second decision-goal wording in parentheses)

> Assume you are shopping for a book to read for pleasure during an upcoming vacation (to help you better manage your stock market investments). You may or may not buy the book online, but you want to shop for it online. One of the things you do when you shop online is check out what other consumers say about the various options that are available.

Automobile decision scenario (second decision-goal wording in parentheses)

> Assume you are shopping for a new car. Assume that, for you, a car is not just a means to get from here to there. You want to get there in style, and have a good time on the ride. (Assume that, for you, a car is just a means to get from here to there. Thus, you are looking for a car that is reliable, practical, and fuel-efficient.) Before you start going to dealerships, you want to do some new-car shopping online. One of the things you do when you shop online is check out what other consumers say about the various options that are available.

APPENDIX B
THE 17 WORDING STYLE CHARACTERISTICS CODED

IMPACT-STRENGTHENING CHARACTERISTICS

1st-person pronoun
E.g., I, me, my, we, our

Proper noun
Proper nouns or direct quotes; e.g., brand name, name of character in a book

Number or date
E.g., 400 horsepower, written in 1997

2nd-person pronoun
E.g., you, your

Exclamation
Exclamation point or use of capital letters for emphasis

Emotion word
E.g., paranoid, hated, annoyed, thrilled, disappointed

Superlative word
E.g., most, least, the best, the worst

Imagery
E.g., left me hanging; characters leap off the page; the car never ceases to put a smile on my face; it is a grown-up go cart

Slang-expressive
Informal usage where it is hard to think of a more descriptive non-slang word; e.g., makes an annoying swooshing sound; the reverse gear is notchy

Humor
In the rater's opinion

Self-effacing wording
Personal information expressed in a way that could be considered humble or unflattering; e.g., I'm not very good with numbers; I got into a serious accident

Sequential statements
A statement describing an event sequential to one described in the previous statement; e.g., I was pretty satisfied with the test drive/so I returned to that dealership to negotiate a price

IMPACT-WEAKENING CHARACTERISTICS

Misspelling
E.g., confusion of "their" and "there"

Qualification
E.g., I'd recommend this car for people who do a lot of commuting

Bad grammar
Ungrammatical word sequence

Slang-inexpressive
Informal usage that could easily have been more descriptive; e.g., it sucks; it's a piece of crap; it's awesome

Repetition
Obvious repetition or contradiction with an earlier statement in the review