Computers in Human Behavior 48 (2015) 17–27

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

A study of factors that contribute to online review helpfulness

Albert H. Huang [a,1], Kuanchin Chen [b,2], David C. Yen [c,*], Trang P. Tran [d,3]

[a] Eberhardt School of Business, University of the Pacific, 3611 Pacific Avenue, Stockton, CA, USA
[b] Department of Business Information Systems, Western Michigan University, Kalamazoo, MI 49008, USA
[c] School of Economics and Business, 226 Netzer Administration Bldg., SUNY College at Oneonta, Oneonta, NY 13820, USA
[d] School of Economics and Business, 326 Netzer Administration Bldg., SUNY College at Oneonta, Oneonta, NY 13820, USA

ARTICLE INFO

Article history:
Available online 9 February 2015

Keywords:
Online reviews
Product reviews
Review helpfulness
Review sidedness

ABSTRACT

Helpfulness of online reviews is a multi-faceted concept that can be driven by several types of factors. This study was designed to extend existing research on online review helpfulness by looking at not just the quantitative factors (such as word count), but also qualitative aspects of reviewers (including reviewer experience, reviewer impact, reviewer cumulative helpfulness). This integrated view uncovers some insights that were not available before. Our findings suggest that word count has a threshold in its effects on review helpfulness. Beyond this threshold, its effect diminishes significantly or becomes near non-existent. Reviewer experience and their impact were not statistically significant predictors of helpfulness, but past helpfulness records tended to predict future helpfulness ratings. Review framing was also a strong predictor of helpfulness. As a result, characteristics of reviewers and review messages have a varying degree of impact on review helpfulness. Theoretical and practical implications are discussed.

1. Introduction

Product reviews have become an essential part of both electronic and traditional commerce. Google Scholar shows that there was much interest in product reviews in the literature (with 15,600 hits for "product reviews" and 13,200 hits for "online reviews"). Similarly ABI/INFORM shows 1015 articles for "product reviews" and 1442 articles for "online reviews" published in scholarly journals. Reading product reviews has been a popular step for consumers in pre-purchase information gathering. As a result, both retailers and manufacturers are paying an increased level of attention to online reviews as they can either be a threat or an opportunity for businesses (Anderson & Magruder, 2012; Chen & Xie, 2008; Chevalier & Mayzlin, 2006; Hu, Zhang, & Pavlou, 2009; Li & Hitt, 2010).

Many studies show that online product reviews and related factors significantly affect sales under certain conditions, and for some product categories (Chen, Dhanasobhon, & Smith, 2007; Duan, Gu, & Whinston, 2008; Forman, Ghose, & Wiesenfeld, 2008). Recent studies suggest that characteristics of reviews and reviewers collected through variables such as reviewer identity, reviewer location, information quantity, and semantic factors (Cao, Duan, & Gan, 2011) may add more insights to the line of research. Since reviewer identity is difficult to come by when collecting reviews directly from publicly available sources (such as Amazon.com), Mudambi and Schuff (2010) suggest that future extensions of their work may focus on other forms of disclosed reviewer's status, such as Amazon's "top reviewer" designation.

Among the many variables associated with online product reviews, "review helpfulness" is particularly important, as it represents the subjective valuation of the review judged by others, and is also the aggregate perceived utility of the information contained in the review (Cao et al., 2011; Baek, Ahn, & Choi, 2012; Li, Huang, Tan, & Wei, 2013). A favorable helpful review adds perceived value to the product, but a critical review is also an opportunity to perform customer service. However, "helpfulness" is quite a complex concept as one would easily equate length (number of words) of review message to helpfulness, while others might consider message length as "effort" instead. Added to the complexity is that helpful messages are often lengthy fill with details, but the other way around may not be true (i.e., a lengthy message is not necessarily a helpful one). Even within helpful reviews, the effect of message length on helpfulness could diminish when it reaches a certain threshold. Beyond the threshold, it may not get read in detail. A straight relationship between helpfulness and message length is explored in a recent study (Mudambi & Schuff, 2010), where the authors reported a linear relationship between

* Corresponding author. Tel.: +1 607 436 3458 (O); fax: +1 607 436 2543.
E-mail addresses: ahuang@pacific.edu (A.H. Huang), kc.chen@wmich.edu (K. Chen), David.Yen@oneonta.edu (D.C. Yen), trang.tran@oneonta.edu (T.P. Tran).
[1] Tel.: +1 (209) 946 7494.
[2] Tel.: +1 269 387 5795.
[3] Tel.: +1 607 436 3535.

http://dx.doi.org/10.1016/j.chb.2015.01.010
0747-5632/

helpfulness and aspects of message and product (rating, product type, work count and total votes). Despite that past behavioral patterns might have an effect on review helpfulness, existing studies have not paid enough attention to them. An example of such pattern includes the effect of past review performance on future review quality. Past helpfulness ratings could become a form of incentive for those who have done good reviews before. As a result, we will be able to shed some light by including past review patterns to study review helpfulness.

In short, it is better for an online reviews study to take a more holistic view by incorporating aspects of messages, reviewers and other characteristics. Therefore, the purpose of this paper is to examine message length together with aspects of review patterns and reviewers for their joint effects on review helpfulness. As message length should not just be taken at the face value (too long or too short is not useful), part of our goal also extends to exploring a possible threshold for message length. Built on the relevant online review literature, six hypotheses were proposed (H1a, H1b, H2 through H5) to study factors relating to review helpfulness. Depending on the nature of the variables involved, two data sets were used to test the hypotheses. The first data set including 1375 reviewers (a.k.a., all reviewers) was employed to study how word count as hypothesized in H1a relates to the review helpfulness. The remaining hypotheses concern about what makes a top reviewer. Therefore, we collected the second data set consisting of 60 top ranking reviewers (hereafter called top reviewers) to test the rest of hypotheses. Reviews in this research were obtained for six focal products (cell phone, printer, camera, music player, music CD, and video game) that are popular in the consumer's market. Along with two data sets, our use of additional reviewer-related variables has developed new findings and conclusions, which offer a theoretical extension to existing literature as well as useful practical implications.

Theoretically, the results of the current research have contributed to relevant literature by providing further understanding of quantitative and qualitative predictors of online review helpfulness. More specifically, the paper takes a step further to uncover the threshold of word count that sheds light on the empirical relationship between this variable and review helpfulness. Additionally, the findings of the paper have extended the results found in existing research (i.e., Mudambi & Schuff, 2010) by looking at also the reviewer aspects of online reviews (i.e., reviewer experience, impact, cumulative helpfulness, and review sidedness) to see whether each of those aspects influences online review helpfulness. Practical implications are also discussed for marketers to incorporate the findings into their marketing strategies in attempt to make reviews more meaningful to their customers.

2. Literature review

Online product reviews attract interest from several academic disciplines. From the marketing perspective, online product reviews can be a valuable tool for promoting products, collecting consumer feedback and boosting sales (Chu & Roh, 2014; Forman et al., 2008; Hu, Liu, & Zhang, 2008). As these studies show, there is a direct relationship between product ratings and sales. For example, online movie reviews and ratings significantly correlate with box office revenues (Liu, 2006), and online book reviews positively affect book sales (Chevalier & Mayzlin, 2006). However, the effects of online reviews vary significantly by category, location, and other factors (Mudambi & Schuff, 2010; Zhu & Zhang, 2010).

A growing body of research has paid attention to review helpfulness. Using Amazon.com data, Baek et al.'s (2012) study finds that both peripheral cues, including review rating and reviewer's credibility, and central cues, such as the content of reviews, influence the helpfulness of reviews. Cao et al. (2011), employing data from CNET Download, state that the semantic characteristics are more influential than other characteristics in affecting how many helpfulness votes reviews receive. Reviews with extreme opinions receive more helpfulness votes than those with mixed or neutral opinions. Based on Amazon.com data, Mudambi and Schuff (2010) conclude that review extremity, review depth, and product type affect the perceived helpfulness of the review. Review depth has a positive effect on the helpfulness of the review, but the product type (search or experience) moderates the effect of review depth on the helpfulness of the review. More recently, Li et al. (2013) conducted a study using Bulletin Board System and found that the source- and content-based review features have a direct impact on product review helpfulness. Consumers perceive customer-written product reviews as more helpful than those written by experts. A customer-written product review with a low level of content abstractness yields the highest perceived review helpfulness. Current research of review helpfulness and possible gaps in the literature are summarized in Table 1.

Although review helpfulness has become an important topic in marketing and information technology literature, little research has explored the effects of both quantitative and qualitative factors on review helpfulness. The current research is developed to bridge the gap in literature by shedding more light on this connection. In terms of quantitative factors, the current research uses word count as a predictor of review helpfulness. Additionally, the quality of information is extremely crucial in online reviews, since high quality information provides reliable, current and concise information (Arazy & Kopak, 2011; Yaari, Baruchson-Arbib, & Bar-Ilan, 2011). In the online review context, quality of information relates to the qualifications and credibility of reviewers (Li & Zhan, 2011; Sotiriadis & van Zyl, 2013). Qualifications and credibility usually take time to establish. This is the reason Amazon uses the total review helpfulness votes to determine the quality of reviewers. Top reviewers receive badges at one of the six levels to encourage them to continue contributing quality reviews. Therefore, the qualifications, credibility and other aspects of reviewer quality are embedded in the review helpfulness ranking for these top reviewers. Since there is currently no objective metrics to quantify qualifications, credibility and other similar factors, top ranking reviewers may be considered as a surrogate representation of a group of reviewers who possess the desired quality and credibility stated in the studies mentioned above. Other qualitative aspects of reviewers including reviewer experience, reviewer impact, and reviewer cumulative helpfulness are also used as qualitative factors in this present study.

2.1. Review helpfulness and measurement

To maintain the value of online reviews and to address concerns about their credibility and quality (Cheung, Sia, & Kuan, 2012), some online review sites allow readers to "review the reviews." The most common approach is to rate a review as "Helpful" or "Not Helpful" (Baek et al., 2012; Li et al., 2013). A helpfulness score is then calculated as the percentage of "Helpful" votes among all votes. Such a check-and-balance measure provides a certain degree of quality assurance, and allows readers to more quickly find helpful reviews among the thousands that may exist. Reviews with a higher number of helpfulness votes were found to have a higher correlation with sales (Chen, 2013; Chen et al., 2007).

In addition to being a quality assurance tool, helpfulness can also be regarded as a subjective measurement of the potential value of the information contained in a review. A review that influences potential customers could logically lead to a purchase. Theoretically, one could calculate the net economic value of a review by summing the net financial outcomes for all consumers who

Table 1
Summary of key findings of recent literature.

Authors (year)	Contribution	Literature gap	Data
Baek et al. (2012)	The results show that both peripheral cues, including review rating and reviewer's credibility, and central cues, such as the content of reviews, influence the helpfulness of reviews	The study does not focus on top reviewers as well as reviewer cumulative helpfulness	Amazon
Cao et al. (2011)	The semantic characteristics are more influential than other characteristics in affecting how many helpfulness votes reviews receive. Reviews with extreme opinions receive more helpfulness votes than those with mixed or neutral opinions	The paper focuses more on how to received more helpfulness votes than on how to crease review helpfulness	CNET Download
Koh et al. (2010)	Significant differences exist across reviewers of online movies in two countries: China and USA. Online reviews are a better movie perceived quality proxy in China and Singapore than in USA	Although online reviews are supposed to be associated with movie perceived quality, review helpfulness is not taken into account	IMDB & Douban
Li and Zhan (2011)	Language style, organizational structure, and other content features of online product reviews affect perceived helpfulness of reviews	The study investigated different factors affecting online review helpfulness	Amazon
Li et al. (2013)	The source- and content-based review features have direct impact on product review helpfulness. Consumers perceive customer-written product reviews as more helpful than those written by experts. A customer-written product review with a low level of content abstractness yields the highest perceived review helpfulness	Although the study focuses on three factors: source-, content-based, and vicarious expression review, other critical factors of review like word count, experience are not taken into account	Bulletin Board System
Ludwig et al. (2013)	The influence of positive affective content on conversion rates is asymmetrical (i.e. different effects between positive affective versus negative content in customer reviews)	The study does not focus on online review helpfulness	Amazon
Mudambi and Schuff (2010)	Review extremity, review depth, and product type affect the perceived helpfulness of the review. Review depth has a positive effect on the helpfulness of the review, but the product type (search or experience) moderates the effect of review depth on the helpfulness of the review	Although word count is found to have a positive effect on review helpfulness, the study does not specify the threshold of word count on helpfulness	Amazon
Sotiriadis and van Zyl (2013)	Three factors that most affect the use of tourism services information retrieved from Twitter include (1) Reliability of Twitter followers; (2) Degree of involvement—Posting; and (3) Expertise of Twitter followers	Review helpfulness is not the main target of the study	Twitter

acted on it. Unfortunately, such information would be nearly impossible to obtain. However, what *may* be obtainable would be the aggregate helpfulness score of the review, submitted by consumers who read and voted on it. This could then be seen as a subjective assessment of the review. While the result would not have a monetary value, it would represent the perceived utility of the information contained in the review.

2.2. Factors contributing to review helpfulness

Understanding the factors that affect helpfulness of online reviews is important to both vendors and reviewers. Vendors continue to invest substantial resources in content engineering and are always looking for better ways to manage and improve the digital content of their portion of cyberspace. Therefore, knowing what makes a review helpful is critical to good content engineering and the overall success of their Internet marketing efforts. A recent study (Mudambi & Schuff, 2010) found a high correlation between the number of words in a review and review helpfulness. It seems that the lengthier of a review, the more likely readers perceive it to be helpful. This finding can be explained intuitively, because more words are needed to convey multiple aspects of a detail, which increases the review quality. In most cases, a *short* review simply does not have the necessary capacity to include all the required elements of a *good* review (Keller & Staelin, 1987).

Yet, quantified information (such as word count) may be at best a necessary factor that is not, of itself, sufficient to gauge the utility value (such as helpfulness) of a review. Other factors such as the content of a review is also important for its effects on helpfulness (Cao et al., 2011), because a large number of words does not always equate to useful or helpful information. An unsatisfied consumer could rant by writing a lengthy review that is neither objective nor relevant to the readers. Additionally, reviewers may differ widely in their expertise and writing skills (Mackiewicz, 2010), with some experienced reviewers providing concise, helpful information, and inexperienced ones providing long but less useful content. Based on this argument, this study aims to study review helpfulness and factors that contribute to it.

3. Theoretical background and research hypotheses

3.1. Information quality

Research (Keller & Staelin, 1987) has suggested that both quantitative and qualitative factors are relevant to study the quality of information. Quantitative factors of product review refer to the amount, length, volume, and other quantity-related aspects of information. Examples of quantitative factors include word count, page count, length of a video or audio recording, size of a digital storage medium, bandwidth of a streaming digital medium, dimensions of a picture, number of cues in the environment, frequency of exposure, and so on (Adler, de Alfaro, Pye, & Raman, 2008; Aizawa, 2000; Van Rooy, 2003). Qualitative factors are more subjective and thus are harder to define. These could refer to the content, writing style, meaning, quality, source, and any other non-quantity aspects of information. Examples of qualitative factors include relevance, accuracy, reliability, timeliness, source credibility, readability, conciseness, sidedness, and others (Alkhattabi, Neagu, & Cullen, 2011; Arazy & Kopak, 2011; Leung, 2001; Wang & Strong, 1996; Yaari et al., 2011).

3.2. Quantitative factors of information and review helpfulness

Not all quantitative factors are relevant to all contexts of study. For video and audio recordings, the length or duration would seem to be an appropriate measure. Alphabetic languages, such as English and Spanish, may also be measured using the number of letters. Logographic languages, such as Chinese or Korean, may be measured using the number of characters (Workman, 2007). Still or motion pictures may be measured by the number of pixels or resolutions. For all digital formats, the number of bits or bytes of data could be used as a common measurement (Weaver & Shannon, 1963). Text-based information can be measured using the number of words or pages. Therefore, word count becomes a natural way to start looking at insights of online reviews because these reviews are typically composed and delivered in text forms.

Furthermore, it has been well established that decision makers' ability to process information is restricted when the amount of information is either extremely abundant or extremely scarce (Martin & Sell, 1980; Schroder, Driver, & Streufert, 1967). The negative effect of insufficient information on decision making is easily understood. The effect of too much information, or information overload, has also been documented in empirical studies. "Information overload" refers to the variety and quantity of stimuli that exceeds the receiver's ability to integrate and process them (Jackson & Farzaneh, 2012; Jacoby, 1977; McCormick, 1970). With the increasing use of information technology, the overabundance of information has become a bigger issue than lack of information (Simperl et al., 2010). Information overload impairs comprehension (Hildon, Allwood, & Black, 2012; Lipowski, 1975) and hampers performance (Driver & Mock, 1975; O'Reilly, 1980; Jakoby, Speller, & Kohn, 1974; Schultz, Schreyoegg, & von Reitzenstein, 2013).

Therefore, the quantified metric of online reviews can be considered a predictor for helpfulness, but the linear relationship as explored in the literature (e.g., Word count in Mudambi & Schuff, 2010) may only hold to a certain extent until it becomes non-linear or no relationship. For example, review word count may be a positive factor of helpfulness as suggested in previous studies. When the word count reaches a certain level, its influence on review helpfulness may start to diminish or become insignificant. The literature has very limited guidance on this threshold. The following hypothesis is then postulated:

H1a. For reviews written by all reviewers, word count is a significant predictor of review helpfulness when the review is shorter than average.

As suggested in previous research, information quantity and quality are interdependent factors. The increase in word count could be an increase in qualitative factors such as relevance and completeness, rather than simply as an increased level of quantity. Consequently, it is questionable whether word count would remain a significant predictor of helpfulness if we hold the qualitative aspects at a more constant level. To verify this hypothesis, we focus specifically on top reviewers and their reviews. This is because top reviewers already enjoy a track record of good review quality. By looking at the word count of their reviews, we are able to more appropriately reduce the confounding effect from other variables (such as ranting, intentional sabotage of manufacturer's reputation, and irrelevant complaints) that may be present in the data set collected from all reviewers. A good review may be filled with details that could make it lengthy, but a lengthy review is not necessarily a good review. Additionally, reviewer qualifications and credibility that plague review quality (see Li & Zhan, 2011; Sotiriadis & van Zyl, 2013) are less of a concern for this group of users. Otherwise, this would be reflected in votes they received, which ultimately determines their being ranked as top reviewers. By focusing on this group, we were able to distill a better relationship among the variables. Therefore, the following hypothesis was developed:

H1b. For the reviews written by top reviewers, word count is a significant predictor of review helpfulness.

3.3. Qualitative factors of information and review helpfulness

The foremost qualitative factor used in the present study is information quality. "Quality" is a construct that is subjective and hard to define. High-quality information might be characterized as accurate, reliable, current, concise, fair, easy to understand, organized, and many other things (Alkhattabi et al., 2011; Arazy & Kopak, 2011; Yaari et al., 2011). Since the purpose of information in general is to inform, the maxim of quality is defined as "appropriately informative" (Grice, 1967).

Similar to other types of communication, one motive of an online review is to influence another person's behavior in accordance with one's own preferences, meaning that reviewers may already have their own goals, preferences, and strategic considerations before reviewing a product (Van Rooy, 2003). As a result, product reviewers may not always be maximally rational in their reviews causing variations in their review quality, quantity and relevance. Reviewers may not provide everything they know, as assumed by the cooperative principle (Ganu, Kakodkar, & Marian, 2013; Grice, 1967; Van Rooy, 2003). Some reviewers may even have ulterior motives (Dellarocas, 2003, 2006; Li & Zhan, 2011; Sotiriadis & van Zyl, 2013). Because of this reality, the qualification and credibility of reviewers is an important qualitative factor of online reviews.

A "qualified source" is someone possessing the expertise relevant to the subject of communication, such as the product being reviewed. A "credible source" is someone who can be trusted to provide a reliable and objective opinion on the subject (Belch & Belch, 1994; Goldsmith, Lafferty, & Newell, 2000; Ohanian, 1990). Credibility of information source has proved to be an important concept in related fields of research. Particularly in marketing, endorser or spokesperson credibility has received considerable attention in academic literature (Bochner & Insko, 1966; Goldberg & Hartwick, 1990; Sternthal, Phillips, & Dholakia, 1978).

As expounded in the previous section, reviewer qualification and credibility may be reflected indirectly in the votes a reviewer has reviewed. A review composed by someone lacks of knowledge or out of an ill intent (termed "trolling") can be spotted sooner or later. Such person is unlikely to receive votes into the top reviewers. There are also other more direct measures for credibility than review helpfulness votes. In the case of online reviews, reviewer qualification and credibility may also be inferred based on the reviewer's track record. Thus, we selected reviewer experience, reviewer impact, and reviewer cumulative helpfulness as indicators of reviewer qualification and credibility. *Reviewer experience* refers to the total number of reviews contributed by a reviewer up to the date of data collection. The more experienced a reviewer with writing reviews, the more likely he or she will be familiar with the aspects of a good review. *Reviewer impact* is defined as the number of votes that a reviewer receives on a certain review. The more impactful the reviewer, the higher number of votes pertaining to a particular review the reviewer receives. *Reviewer cumulative helpfulness* represents a ratio of the total number of YES ("Helpful") votes to the total number of votes. If a reviewer has a higher percentage of Yes votes out of the total number of votes, he has a higher level of cumulative helpfulness. Top reviewers are those who have high scores in those variables. With respect to reviewer experience, reviewer impact, and reviewer cumulative helpfulness, three hypotheses are proposed below:

H2. For top reviewers, reviewer experience is a significant predictor of review helpfulness.

H3. For top reviewers, reviewer impact is a significant predictor of review helpfulness.

H4. For top reviewers, reviewer cumulative helpfulness is a significant predictor of review helpfulness.

The framing of a message may be neutral, positive, or negative toward a product, and may affect the perceived value of the message (Grewal, Gotlieb, & Marmorstein, 1994); but findings on the relationship between framing and consumer attitude have been conflicting: Some studies suggest that negatively framed information is more credible because it is unlikely to be contributed by the product's sellers or manufacturers (Kanouse, 1984). However, some people may prefer messages that present both the negative and positive aspects of a product (Hastak & Park, 1990), and so may view online reviews that contain both the pros and cons of a product to be more objective and thus more believable. Two-sided arguments were also more persuasive than one-sided positive arguments when the initial attitude of the consumer was neutral or negative (Crowley & Hoyer, 1994). Some studies have also shown that reviews that carry both positive and negative opinions receive more helpfulness votes than those with neutral opinions (Cao et al., 2011).

Similar to review framing, the star rating of the product being reviewed has been shown to correlate with review helpfulness. A review that rates a product five stars would logically contain more positive information than the review that rates only one-star (Mudambi & Schuff, 2010; Poston & Speier, 2005). Some readers depend more on low-rating reviews because they feel that they are more diagnostic and thus more useful (Ahluwalia, Burnkrant, & Unnava, 2000), and positive reviews as less helpful because of their weaker perceived depth (Hao, Li, & Zou, 2009). There is also evidence that product ratings are positively associated with review helpfulness (Mudambi & Schuff, 2010).

Preference for extremity might be due to the fact that readers may prefer the reviewer to have a definite opinion, whether positive or negative, rather than to be wishy-washy. The second possible explanation is that many readers who read a review may already have a positive opinion of the product. Studies have shown that when the opinions of reviewers are more aligned with the opinions of readers, more communication is possible (Crawford & Sobel, 1982). That is, extreme reviews (particularly positive reviews) may be more aligned with the existing opinions of some readers, and so are perceived to be more helpful.

To examine this relationship, we set up the fifth hypothesis that illustrates the relationship between a reviewer's product rating and review helpfulness. Within the scope of the study, product rating indicates whether reviews of a product are positive or negative. When a product is evaluated positively, the product has high product rating. To control the variation of review quality, we focused only on reviews written by top reviewers. The hypothesis is stated as:

H5. For the top reviewers, product rating is a significant predictor of review helpfulness.

To summarize, this study addresses two general constructs: quantitative and qualitative factors of online reviews. For quantitative factor, we examined the effect of word count at several levels to uncover the threshold of word count as it relates to review helpfulness. For qualitative factors, we examined the review framing (sidedness), reviewer experience, reviewer impact, and past average helpfulness. As a result, these variables together capture three essential aspects of online reviews: review message, product, and reviewer, thus providing a more comprehensive coverage of online reviews. All hypotheses developed on the basis of the group of top reviewers (H1b through H5) are presented in Fig. 1.

4. Method

To address the five hypotheses, we collected two data sets. The first data set has a total of 2209 reviews for six products (a cell phone, printer, camera, music player, music CD, and video game). The selection of these products is based on Mudambi and Schuff (2010) that uncovers initial insights regarding the relationship between word count and helpfulness. By following the categories studied in existing literature, we can then offer contributions to the literature in the forms of theory expansion and generalization. After removing reviews with no votes on helpfulness, 1375 reviews remained in the sample. This data set was used to address H1a, since this hypothesis concerns about the general relationship between review helpfulness and word count.

The remaining hypotheses were formulated to study behaviors of top reviewers. Therefore, the second data set was collected from top-ranking reviewers who had an established track record in review writing. Inexperienced reviewers were excluded. To locate these experienced reviewers, we used one of the reviewer ranking systems on Amazon.com. (Amazon.com has two reviewer ranking systems: a classical ranking system, and a new ranking system that ranks reviewers using different criteria.) For this study, we used the new reviewer ranking system, because this system determines a reviewer's rank by a weighted score of cumulative helpfulness and the total number of reviews the reviewer has written.

We drew a sample of 60 reviewers using the "Top 10,000" reviewer list of the new ranking system as the sampling frame. Top reviewers were chosen using 60 random numbers between 1 and 10,000. The variables that we collected for these top reviewers include the total number of reviews written, the total number of helpful votes received, and the overall percentage of helpful votes on their reviews. A random sample of 30 reviews was obtained from each top reviewer. For reviewers who had written less than 30 reviews, all reviews were collected. As a result, the total number of reviews obtained was less than 1800. The variables collected for the reviews by these top reviewers were product ratings and word count.

5. Results

5.1. Descriptive statistics

Table 2 shows the descriptive statistics of reviews collected for this study. The data show significant differences between the top reviewer group and the all reviewers in terms of word count and helpfulness. On average, top reviewers wrote almost 106 words more than other reviewers in each review. Their average helpfulness was about 18% higher. Table 3 contains descriptive statistics about the top reviewers selected.

5.2. Effect of word count

Previously, word count was found to be a significant predictor of review helpfulness (Mudambi & Schuff, 2010). However, this correlation may have been different if:

1. The number of words had already reached a sufficient amount to adequately deliver the necessary information. Anything more added to the message could likely create a marginal, insignificant or even negative effect on the helpfulness of the review.
2. The reviews had been limited to those written by the top reviewers. Generally, these top reviewers are experienced in writing reviews more effectively and efficiently. In other words, messages with the same word count may contribute to a varying degree of review helpfulness depending on who wrote the messages. As a result, studies focusing on the general population for their review helpfulness may be a starting point to help understand the general relationships among variables. This approach, however, may fall short when one wishes to study the underlying reasons for helpful reviews. As we explained, one large confounding effect is the variation of review helpfulness (even under the same message length) due to the effectiveness between top and other reviewers in writing

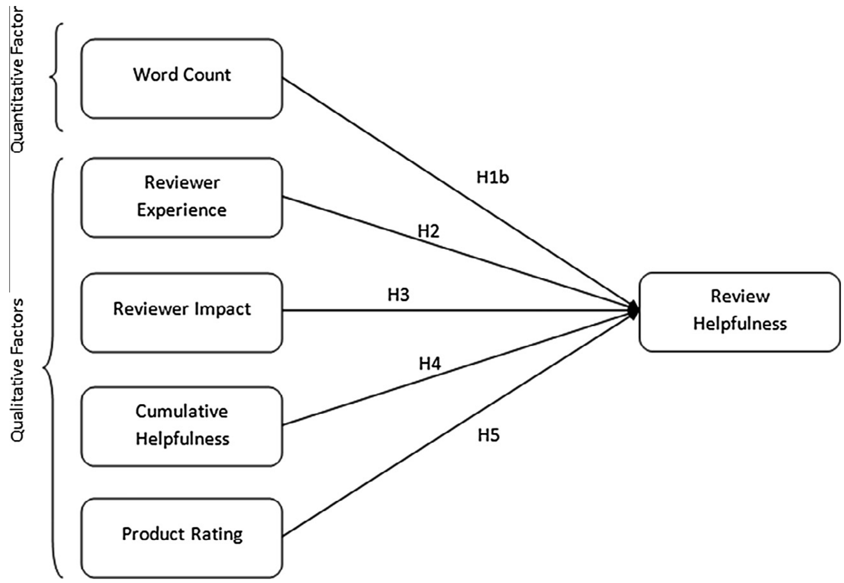

Fig. 1. Conceptual model of online review helpfulness (applied to top reviewers).

Table 2
Descriptive statistics for Data Sets 1 and 2.

	Data Set 1: Collected from all reviewers of 6 categories of products ($N = 1375$)		Data Set 2: Collected from top reviewers ($N = 1327$)	
	Mean	SD	Mean	SD
Rating	4.07	1.33	4.12	1.16
Word count	144.00	152.35	250.55	209.79
Helpfulness%	64.07	36.32	82.27	27.82

Table 3
Descriptive statistics of the selected reviewers sample ($N = 60$).

	Mean	SD
Numbers of reviews written	219.95	326.76
Total votes received	1496.42	2446.53
Average helpfulness of reviewers%	85.58	8.55

reviews. Focusing on top reviewers allows us to avoid a large part of this confounding effect by better modeling review helpfulness and its relationship with other variables.

Data Set 1 was analyzed using Tobit Regression (Croissant & Millo, 2008; Tobin, 1958) because of the dependent variable, review helpfulness, never becomes negative. The general construction of Tobit Regression assumes the following form (see Amemiya, 1973; Tobin, 1958):

$$y_i = \begin{cases} \beta x_i + u_i, & \text{if RHS} > 0 \\ 0, & \text{if RHS} \leqslant 0 \end{cases}$$

$$u_i \sim \text{IIDN}(0, \sigma^2)$$

When all reviews were included in the analysis, the relationship between word count and helpfulness was significant (see Table 4, Regression I). This result confirmed the findings of the previous study (Mudambi & Schuff, 2010).

To test hypothesis H1a, we split the data into two subsets using the average word count of 144 as the cut-off point after experimenting with different cut-offs. This is a point beyond which the relationship among variables begins to change as reported below. Using the mean value as the cut-off is similar to the common practice of mean-split used to compare group differences.

The relationship between word count and helpfulness remained significant ($\beta = 0.321$, $p < 0.001$, Table 4, Regression II) for reviews with 144 or less words. When the word count exceeds 144, the relationship became statistically insignificant ($\beta = 0.012$, $p = 0.076$, Table 4, Regression III). As a result of this difference between the two sub-samples, H1a was supported. This confirms our expectation in that information quantity as measured with word count is important, but only to a certain extent. Once the word count exceeds a certain level, its effect is no longer significant in predicting review helpfulness.

Tobit regression was also performed to test H1b, using data collected from the top reviewers (Data Set 2). The result showed that word count had a very small coefficient ($\beta = 0.0066$, $p = 0.070$, Table 5) and it was not statistically significant. Moreover, the R^2 was less than 1%, indicating that very little variance was explained by word count. Therefore, H1b was not supported. In other words, word count was not relevant to review helpfulness among top reviewers.

Table 4
Regression output of word count as the only predictor (using reviews for 6 products).

	Estimate	Standard deviation	t-value	p-value
Regression I: All reviews				
(Intercept)	54.964	1.277	43.030	0.000
Word count	0.064	0.006	10.565	0.000***
Regression II: Reviews with 144 words or less				
(Intercept)	36.490	2.466	14.797	0.000
Word count	0.321	0.034	9.476	0.000***
Regression III: Reviews with 145 words or more				
(Intercept)	74.809	2.468	30.306	0.000
Word count	0.012	0.007	1.778	0.076

5.3. Reviewer characteristics

As the above analysis shows, the effect of word count was negligible for top reviewers, which suggests that other factors may have played a more important role than word count in predicting review helpfulness. We further report the result of including three additional predictors (reviewer experience, reviewer impact, and reviewer cumulative helpfulness) in this section.

Reviewer experience was measured using the total number of reviews contributed by a reviewer up to the date of data collection. A reviewer who wrote more reviews has accumulated more experience writing reviews. Reviewer impact was measured using the total votes of all reviews written by a reviewer. However, this was only a surrogate measure of popularity, because the number of people who read the reviews but did not vote remains unknown. Reviewer cumulative helpfulness was measured by the ratio of total "Helpful" votes to total votes received based on all reviews written by the same reviewer. We hypothesized that more-experienced reviewers write more-helpful reviews, that popular reviewers are the ones who wrote helpful reviews; and that reviewer cumulative helpfulness has a positive correlation with the helpfulness of individual reviews.

The results of Tobit regression showed no significant relationship between reviewer experience and review helpfulness ($\beta = 6.032\text{e}{-04}$, $p = 0.800$, Table 6). That is, reviewers who had written *more* reviews did not necessarily write *more-helpful* reviews. Therefore, productivity (or quantity) did not correlate with either quality or effectiveness. Under Amazon.com's new ranking system, reviewers are ranked according to two criteria: reviewer experience (the number of reviews written) and helpfulness. Our finding suggests that the two criteria used by Amazon.com do indeed measure different dimensions of a reviewer.

We were also unable to discern any significant relationship between reviewer impact and review helpfulness ($\beta = 3.393\text{e}{-04}$, $p = 0.250$). Reviewers who received more votes on their reviews did not necessarily write more-helpful reviews. In other words, reviews that were read by more people were not necessarily the more-helpful ones.

Reviewer cumulative helpfulness was the only reviewer characteristic that was a significant predictor of review helpfulness ($\beta = 70.88$, $p < 0.001$). The cumulative helpfulness was measured as the total number of YES ("Helpful") votes divided by the total number of votes out of all reviews by the same reviewer. The result is similar to a weighted average of individual review helpfulness. Reviews that received more votes were weighted more. The significant positive correlation indicates that reviewers with a high cumulative helpfulness score were more likely to receive a higher helpfulness scores for their reviews. Although this finding cannot assure that reviewers who were helpful in the past would be helpful in the future, it is an indication that reviewers were generally consistent in the helpfulness of their reviews.

5.4. Review framing

Product rating was used as a surrogate of the review framing. The results (Tables 7 and 8) show that the product rating was still a significant predictor of review helpfulness when the data were obtained from the top reviewers. The slope of the regression line was positive; that is, positive reviews were more likely to be helpful reviews no matter who the reviewers were.

These findings also revealed a built-in bias by some review sites against negative reviews. For example, Amazon.com allows users to organize the reviews using two different criteria: the most-helpful ones first or the most-recent ones first. By default, the most-helpful reviews are listed first, but can be manually changed. Consequently, consumers may be exposed to more positive reviews than to negative ones so long as they simply read the reviews in the default order shown on the site.

5.5. A combined regression model

As shown in the analysis above, product rating and reviewer cumulative helpfulness were independently found to be significant predictors of helpfulness for reviews written by top reviewers. To determine whether the two predictors remained significant when combined, a Tobit regression model with these two predictors was constructed and tested. Both predictors remained statistically significant in the same model at the 1% significance level. We also tested the model with interaction terms between the two variables, but did not find a significant interaction effect.

$$\begin{aligned}&\text{Helpfulness \% of Reviews Written By Top Reviewers}\\&= 9.90 + 52.10 * \text{Reviewer Cumulative Helpfulness} + 6.87\\&\quad * \text{Rating} + \text{Error}\end{aligned}$$

6. Conclusions

The purpose of this paper is to examine both message length together with aspects of review patterns and reviewer characteristics for their joint effects on review helpfulness. As message length should not just be taken at the face value (too long or tool short is not useful), part of our goal also extends to exploring a possible threshold for message length. Built on relevant online review literature, six hypotheses were proposed (H1a, H1b, H2 through H5). Two data sets were used to test hypotheses. The first data set including 1375 reviews was employed to test H1a, while the second consisting of 60 top ranking reviewers (hereafter called top reviewers) was used to test the remaining hypotheses. Amazon's ranking system was utilized to identify the top reviewers by using a weighted score of cumulative helpfulness and the total number of reviews the reviewer has written. Reviews were obtained from customers' evaluation of six focal products (cell phone, printer, camera, music player, music CD, and video game). The results show that H1a proposing that word count is a significant predictor of review helpfulness when word count reaches a certain level is supported. Although previous literature has suggested that word count positively affects review helpfulness (Mudambi & Schuff, 2010), the current research points out that this relationship is valid until the number of word count reaches a certain threshold of 144 words. Specifically, our empirical results show that the relationship between word count and helpfulness remained significant for reviews with 144 or less words. When the word count exceeds 144, the relationship became statistically insignificant.

H1b stating that for the reviews written by top reviewers, word count is a significant predictor of review helpfulness is not supported. The empirical findings show something interesting that shed light on the difference between general reviewers and top reviewers. The positive relationship between word count and review helpfulness that exists among all reviewers (Mudambi & Schuff, 2010) does not exist among top reviewers. In other words, among top reviewers, quantitative factors (i.e., word count) are not key drivers leading to review helpfulness.

Table 5
Regression output of word count as the only predictor (using "Top 10,000" reviewer data).

	Estimate	Standard deviation	*t*-value	*p*-value
(Intercept)	80.60	1.19	67.85	<0.001***
Word count	0.0066	0.0036	1.83	0.07

*** Significant at 1% level; R-Square = 0.003.

Table 6
Regression output of reviewer characteristics as predictors of review helpfulness (with each variable tested independently).

Variable	Estimate of coefficient	Standard error	*t*-value	*p*-value	R^2
Reviewer experience	6.032e−04	2.397e−03	0.25	0.80	0.000
Reviewer impact	3.393e−04	2.923e−04	1.16	0.25	0.001
Reviewer cumulative helpfulness	70.88	8.80	8.06	<0.001***	0.047

*** Significant at 1% level.

Table 7
Regression output using rating as the only predictor (using "Top 10,000" reviewer data).

	Estimate	Standard deviation	*t*-value	*p*-value
(Intercept)	50.90	2.66	19.12	<0.001***
Rating	7.62	0.62	12.24	<0.001***

*** Significant at 1% level; *R*-Square = 10.2%.

Table 8
Regression model of helpfulness of reviews written by top reviewers.

	Estimate	Standard deviation	*t*-value	*p*-value
(Intercept)	9.90	7.26	1.36	0.173
Cumulative helpfulness	52.10	8.60	6.06	<0.001***
Product rating	6.87	0.63	10.96	<0.001***

*** Significant at 1% level; *R*-Square = 12.8%.

Four qualitative factors are selected as surrogate measures for reviewer quality and credibility, including reviewer experience, impact, cumulative helpfulness, and product rating. The relationship between each factor and review helpfulness is captured in four hypotheses (H2 through H5). However, the empirical results show various effects of each factor on review helpfulness. Particularly, H2 postulating that for top reviewers, reviewer experience is a significant predictor of review helpfulness is not supported, H3 suggesting that for top reviewers, reviewer impact is a significant predictor of review helpfulness is not supported either, H4 proposing that for top reviewers, reviewer cumulative helpfulness is a significant predictor of review helpfulness is supported, and H5 suggesting that for the top reviewers, product rating is a significant predictor of review helpfulness is supported. Stated differently, among top reviewers, only do reviewer cumulative helpfulness and product rating affect review helpfulness while reviewer experience and product rating do not. This means that an online review becomes more helpful either when the ratio of the total number of YES ("Helpful") votes to the total number of votes for a particular review is higher, or when a product is evaluated positively. The findings are crucial because they provide a unique perspective regarding how qualitative factors among top reviewers impact review helpfulness. Related literature discusses how credibility leads to trustworthy advice that is perceived by the audience (Belch & Belch, 1994; Bochner & Insko, 1966; Goldsmith et al., 2000; Ohanian, 1990; Goldberg & Hartwick, 1990). Reviewer cumulative helpfulness and product rating are important antecedents of online review helpfulness.

All results of hypotheses testing are summarized in Table 9. Further theoretical and managerial implications are discussed in the accompanying paragraphs.

7. Theoretical implications

The results of the current research have contributed to relevant literature by providing a deeper understanding of quantitative and qualitative predictors of online review helpfulness. More specifically, the paper takes a step further to uncover the threshold of word count that sheds light on the empirical relationship between this variable and review helpfulness. Additionally, the findings of the paper have extended the results found in existing research (i.e., Mudambi & Schuff, 2010) by looking at also the reviewer aspects of online reviews (i.e., reviewer experience, impact, cumulative helpfulness, and review sidedness) to see whether each of those aspects influences online review helpfulness.

Previous studies support word count to be a significant predictor of review helpfulness, but little is known whether this will hold true for different word counts. Based on the online product reviews in the six product categories selected to be consistent with existing literature, findings in the present study suggest that word count was indeed a significant predictor of review helpfulness, but only to a certain extent. Word count was a significant predictor when all reviews short in length (less than 144 words) were considered. Its effect becomes nearly non-existent for longer reviews (with more than 144 words). Such a threshold of 144 words was empirically located after testing different variations of the word count. The implication of our finding is that the relationship between information quantity (i.e., work count) and information utility (i.e., review helpfulness) may be linear with a positive slope until a threshold is reached. After this threshold, the relationship turns flat. As a result, the benefit of word count on review helpfulness diminishes as reviews became longer. This phenomenon may be explained by the theory of information overload. Initially, words count was important in order to make the review message informative. As the length of a review grows longer, the increase in the utility presented in the review may be offset by an increased demand of readers' mental capacity (Jackson & Farzaneh, 2012; Kanouse, 1984). When this happens, readers could choose to skim through the message or abandon it prematurely, giving the message less chance to be helpful. One possible result is a lowered level of review helpfulness.

Another possible cause of this phenomenon is that some Internet users may not be able to finish reading long reviews before their attention wanders away. Past studies have noted that the Internet has affected our ability to stay focused (Sparrow, Liu, & Wegner, 2011). People who are used to reading books and lengthy articles find themselves quickly losing interest or losing focus when reading the same type of information online likely because the endless stream of seemingly unlimited information and the constant distractions from various intruding applications all make it difficult to stay on one task for a long time. In the case of online reviews, it is common to find hundreds of reviews for the same product; and review sites such as Amazon.com are jam-packed with advertisements and links designed to grab the attention of the users. So, it is probably not realistic to expect consumers to read long reviews in their entirety; it is more likely that they will jump rapidly from one thing to another. Some users will not even take time to vote on the helpfulness of the review before their attention switches to other matters. Consequently, the marginal effect of information quantity disappears when the review length exceeds a certain word count (144 words). This translates to about half of a double-spaced page as the ideal length of reviews.

Our data also showed that top reviewers wrote longer reviews than average reviewers, meaning that longer reviews were more

Table 9
Results of hypotheses testing.

H1a: For reviews written by all reviewers, word count is a significant predictor of review helpfulness when the review is shorter than average	Supported
H1b: For the reviews written by top reviewers, word count is a significant predictor of review helpfulness	Not supported
H2: For top reviewers, reviewer experience is a significant predictor of review helpfulness	Not supported
H3: For top reviewers, reviewer impact is a significant predictor of review helpfulness	Not supported
H4: For top reviewers, reviewer cumulative helpfulness is a significant predictor of review helpfulness	Supported
H5: For the top reviewers, product rating is a significant predictor of review helpfulness	Supported

likely to have come from top reviewers. Thus, the increase in the helpfulness of longer reviews, as shown in previous studies (such as Hu et al., 2008) could also be the result of other factors such as reviewer characteristics and other qualitative variables. In other words, it could be quality, not quantity (such as word count) that really affects helpfulness as the length of reviews grows longer.

Of the three qualitative factors that we analyzed (reviewer experience, reviewer impact and reviewer cumulative helpfulness), reviewer cumulative helpfulness turned out to be statistically significant at 0.001. Although past performance may not always be indicative of future results, our findings support that there were some traces of consistency in review quality by top reviewers. Reviewer experience and reviewer impact, however, were not significantly correlated with review helpfulness. Judging by the very small beta coefficients, their practical effect on review helpfulness is quite negligible. This lack of statistical relationship between review experience and impact on review helpfulness has some interesting implications. Since reviewer experience was measured through the proxy variable of total review count, it seems that the volume of reviews does not translate into review helpfulness. The same goes to reviewer impact that was measured as the total votes on all reviews by a reviewer. As the findings suggest, strategies to increase review volume and/or attract votes will unlikely be successful in affecting the review helpfulness rating.

Review framing, measured by product rating, was also a significant predictor of helpfulness for reviews written by top reviewers. Combining our findings and the findings of prior research (Mudambi & Schuff, 2010), product rating is a consistent predictor for review helpfulness, no matter whether the reviews were written by top or all reviewers. This has practical implications for e-businesses. Previous studies have shown that reviews with extreme ratings received more YES votes on helpfulness (Cao et al., 2011), and most ratings of online reviews tend to be positive (Hu et al., 2009). In other words, most reviews are positive, and the positive ones are more likely to be seen as helpful by consumers. Our work also confirms existing findings. This provides a solid foundation for viral marketing. One way to take advantage of positive reviews is to develop a strong initial response from a small group of influential reviewers when the product is first introduced. Their early positive reviews may be shared in many modern forms of distribution channels (e.g., social networking), which eventually become early influences of opinions (Koh, Hu, & Clemons, 2010) that could become viral in the online community.

8. Managerial implications

Several practical implications may be derived from our findings. First, the results reported may be used as a guideline to create more meaningful online reviews, especially for those who wish to compose helpful reviews. For example, the myth of "the longer review the better" has an empirical support in our study, but we went a step further to identify the upper limit of its effect on helpfulness. Together with the result of no statistical support for the relationship between review volume (termed reviewer experience in our study) and helpfulness, one could conclude that "content is king" for online reviews – a similar conclusion that has long been adopted in online usability studies (Nielsen, 1999). The resemblance is rooted in the fact that both helpfulness of online reviews and perceived online usability are derived from and built for online viewers, who have been characterized to have little time to handle a large amount of online/offline information using a communication medium (i.e., Internet) that even slows down the speed of online reading (Al-Othman, 2003). Although the above recommendation is primarily constructed based on review contributor's point of view, the same findings can also be used to the advantage of review readers. The upper bound of the word count uncovered in the present study is even more relevant when a helpful review is viewed through electronic forms. When a reader come across a lengthy review (of more than 144 words or roughly half a double-spaced page), the reader might not need to read the message in detail due to the likelihood of it being less informative.

Second, review rating was found to be a predictor of review helpfulness. Although this applies to both positive and negative reviews, merchants should not fear the negative reviews as it may present an opportunity for a possible online customer service. Responsiveness, reliability and ease of use have been identified as the top three reasons why customers are satisfied with an online merchant (Yang & Fang, 2004). Similarly, the service quality literature (e.g., Calabrese & Scoglio, 2012) also suggests that reliability, responsiveness, service assurance, empathy and tangibles are key dimensions of service quality. All point to the fact that responsiveness could be a key to remedy the damage done by negative reviews. Key personnel may be assigned to respond to negative reviews in order to foster a sense of care and quality customer service. The result may not turn a negative rating to positive, but the purpose is to reduce the impact of the top three concerns identified above.

Third, review helpfulness is a continuing effort as shown in the effect of past review helpfulness on the overall helpfulness rating. For the top reviewers, word count, reviewer experience, and reviewer impact all had a very limited effect on overall helpfulness. Reviewers wishing to be helpful are reminded that consistency in review quality is the key to review helpfulness. Another "me too" review message does not necessarily contribute to review helpfulness much, but it may still have some effect on the perception of "vote" to the readers.

Overall, this study shows that review helpfulness is a complex construct, and that it is correlated with review length to a certain extent. Qualitative factors, including reviewer characteristics (e.g., cumulative helpfulness) and individual review characteristics (e.g., sidedness or product ratings) play a role in estimating the helpfulness of reviews. Furthermore, the overall R^2 indicates that other factors may have been responsible for a significant portion of the variation in review helpfulness. For example, studies have shown that messages that are engaging, imaginative, fun, and intriguing are important factors in attracting readers (Dobele, Toleman, & Beverland, 2005), but these perceptions of fun, engaging, etc. are rather personal. A fun review message for one reader may be considered offensive by another. More studies of these qualitative factors are needed to further understand this complex and important construct in e-business.

9. Limitations

As with all empirical studies, this study is not without limitations. First, data were collected from the Amazon.com review site. Although Amazon.com is a famous online retailer that provides customers with opportunities to leave their comments on anything they purchase, several retailers including other online retailers, like Newegg.com, or Rakuten.com, follow the trend by creating similar functionalities on the websites for customers to review. Additionally, customer reviews are not only limited to online retailers but also to physical stores, such as such as Wal-Mart and Best buy, who now take advantage of Internet to add an online channel to the traditional marketing channel. Therefore Amazon.com is one among several retailers welcoming customer reviews on the website. So using the data collected from Amazon.com to generalize to the general market would be biased. Therefore, readers are cautioned not to generalize the results beyond the intended context. However, the size of this organization and the volumes of trades that it processes have made it one of the largest online retailers. The sample drawn from this organization nonetheless still represents a large proportion of the true online retailing population. Second, the data for testing the effect of word count were collected by comparing six products (cell phone, printer, camera, music player, music CD, and video game). Hence the results of this study are more appropriately applicable to these types of products. Although our goal was to be consistent with existing studies in product categories to maintain consistency and the correct base for benchmarking, future studies can include more products or different brands in order to confirm or contradict our findings.

Additionally, there were two types of products used in this study: experience type (music player, music CD and video game) and search type (cell phone, camera, and printer). According to Mudambi and Schuff (2010), product types might affect the perceived helpfulness of review. However, the moderation effect of product types on customer reviews is not a major consideration of the current research. Hence this would open doors for future research to investigate and see whether the results still exist.

Last, the data used in the current study were collected from Amazon.com where demographic features of reviewers were unknown. Amazon is among the many retail web sites (such as BestBuy.com and Walmart.com) that does not strictly enforce real names for product reviews. Although this promotes an environment conducive to a more comfortably review atmosphere, it makes it difficult to tap into the demographics of the reviewers reliably without hampering the quality of the reviews. Additionally, some people behave differently when they know they are being watched – a situation called the Hawthorne effect. When this happens during the more direct data collection methods (such as experiments), researchers gain the benefit of being able to collect reviewer demographics at the expense of the quality of product reviews. The gain does not outweigh the risk for the context of our present study.

However, the way that customers rate or review a product or a service is affected by cultural factors (Koh et al., 2010). According to this study, customers living in a collectivist culture tend to rate movies differently from their counterparts in individualist culture. Therefore, a potential avenue for future research is to explore the moderating effect of cultural factors on online review helpfulness.

References

Adler, B., de Alfaro, L., Pye, I., & Raman, V. (2008). Measuring author contributions to the Wikipedia. In *Proceedings of the 4th international symposium on Wikis*.

Ahluwalia, R., Burnkrant, R., & Unnava, R. (2000). Consumer response to negative publicity: The moderating role of commitment. *Journal of Marketing Research, 37*, 203–221.

Aizawa, A. (2000). The feature quantity: an information theoretic perspective of TFIDF-like measures. In *Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 104–111).

Alkhattabi, M., Neagu, D., & Cullen, A. (2011). Assessing information quality of e-learning systems: A web mining approach. *Computers in Human Behavior, 27*, 862–873.

Al-Othman, N. M. A. (2003). The relationship between online reading rates and performance on proficiency tests. *The Reading matrix, 3*(3), 120–136.

Amemiya, T. (1973). Regression analysis when the dependent variable is truncated normal. *Econometrica, 41*(6), 997–1016.

Anderson, M., & Magruder, J. (2012). Learning from the crowd: Regression discontinuity estimates of the effects of an online review database∗. *The Economic Journal, 122*(563), 957–989.

Arazy, O., & Kopak, R. (2011). On the measurability of information quality. *Journal of American Society of Information Sciences and Technology, 62*(1), 89–99.

Baek, H., Ahn, J. H., & Choi, Y.-S. (2012). Helpfulness of online consumer reviews: Readers' objectives and review cues. *International Journal of Electronic Commerce, 17*(2), 99–126.

Belch, G., & Belch, M. (1994). *Introduction to advertising and promotion: An integrated marketing communications perspective* (3d ed.). Homewood, IL: Irwin.

Bochner, S., & Insko, C. (1966). Communicator discrepancy, source credibility, and opinion change. *Journal of Personality and Social Psychology, 4*, 614–621.

Calabrese, A., & Scoglio, F. (2012). Reframing the past: A new approach in service quality assessment. *Total Quality Management and Business Excellence, 23*(11–12), 1329–1343.

Cao, Q., Duan, W., & Gan, Q. (2011). Exploring determinants of voting for the "helpfulness" of online user reviews: a text mining approach. *Decision Support Systems, 50*, 511–521.

Chen, H.-N. (2013). An investigation into online reviewers' behavior. *European Journal of Marketing, 47*(10), 1758–1773.

Chen, P., Dhanasobhon, S., & Smith, M. D. (2007). All reviews are not created equal: The disaggregate impact of reviews and reviewers at amazon.com. In *Proceedings of the 28th international conference on information systems, Montreal, Quebec, Canada* (pp. 9–12).

Chen, Y., & Xie, J. (2008). Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management Science, 54*(3), 477–491.

Cheung, C. M.-Y., Sia, C.-L., & Kuan, K. K. Y. (2012). Is this review believable? A study of factors affecting the credibility of online consumer reviews from an ELM perspective. *Journal of the Association of Information Systems, 13*(8), 618–635.

Chevalier, J., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research, 43*(3), 345–354.

Chu, W., & Roh, M. (2014). Exploring the role of preference heterogeneity and causal attribution in online ratings dynamics. *Asia Marketing Journal, 15*(4), 61–101.

Crawford, V., & Sobel, J. (1982). Strategic information transmission. *Econometrica, 50*(6), 1431–1451.

Croissant, Y., & Millo, G. (2008). Panel data econometrics in R: The PLM package. *Journal of Statistical Software, 27*(2), 1–43.

Crowley, A., & Hoyer, W. (1994). An integrative framework for understanding two-sided persuasion. *Journal of Consumer Research, 20*(4), 561–574.

Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science: Special Issue on E-Business and Management, 49*(10), 1407–1424.

Dellarocas, C. (2006). Strategic manipulation of internet opinion forums: Implications for consumers and firms. *Management Science, 52*(10), 1577–1593.

Dobele, D., Toleman & Beverland, M. (2005). Controlled infection! Spreading the brand message through viral marketing. *Business Horizons, 48*(2), 143–149.

Driver, M., & Mock, T. (1975). Human information processing decision style theory, and accounting information systems. *Accounting Review, 50*(3), 490–508.

Duan, W., Gu, B., & Whinston, A. (2008). The dynamics of online WOM and product sales – An empirical investigation of the movie industry. *Journal of Retailing, 84*(2), 233–242.

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research, 19*(3), 291–313.

Ganu, G., Kakodkar, Y., & Marian, A. (2013). Improving the quality of predictions using textual information in online user reviews. *Information Systems, 38*(1), 1–15.

Goldberg, M., & Hartwick, J. (1990). The effects of advertiser reputation and extremity of advertising claim on advertising effectiveness. *Journal of Consumer Research, 17*, 172–179.

Goldsmith, R., Lafferty, B., & Newell, S. (2000). The impact of corporate credibility and celebrity credibility on consumer reaction to advertisements and brands. *Journal of Advertising, 29*(3), 43–54.

Grewal, D., Gotlieb, J., & Marmorstein, H. (1994). The moderating effects of message framing and source credibility on the price-perceived risk relationship. *Journal of Consumer Research, 21*(1), 145–153.

Grice, H. (1967). Logic and conversation. In P. Grice (Ed.), *Studies in the way of Worlds*. Cambridge, MA: Harvard University Press.

Hao, Y., Li, Y., & Zou, P. (2009). Why some online product reviews have no usefulness rating? In *Proceedings of the 2009 Pacific Asia conference on information systems*.

Hastak, M., & Park, J. (1990). Mediators of message sidedness effects on cognitive structure for involved and uninvolved audiences. *Advances in Consumer Research, 17*(1), 329–336.

Hildon, Z., Allwood, D., & Black, N. (2012). Impact of format and content of visual display of data on comprehension, choice and preference: A systematic review. *International Journal for Quality in Health Care, 24*(1), 55–64.

Hu, N., Liu, L., & Zhang, J. (2008). Do online reviews affect product sales? The role of reviewer characteristics and temporal effects. *Information Technology Management, 9*(3), 201–214.

Hu, N., Zhang, J., & Pavlou, P. (2009). Overcoming the J-shaped distribution of product reviews. *Communications of the ACM, 52*(10), 144–147.

Jackson, T. W., & Farzaneh, P. (2012). Theory-based model of factors affecting information overload. *International Journal of Information Management, 32*(6), 523–532.

Jacoby, J. (1977). Information load and decision quality: Some contested issues. *Journal of Marketing Research, 14*, 569–573.

Jakoby, J., Speller, D., & Kohn, C. (1974). Brand choice as a function of information load. *Journal of Marketing Research, 11*(1), 63–69.

Kanouse, D. (1984). Explaining negativity biases in evaluation and choice behavior: Theory and research. *Advances in Consumer Research, 11*(1), 703–708.

Keller, K., & Staelin, R. (1987). Effects of quality and quantity of information on decision effectiveness. *Journal of Consumer Research, 14*(2), 200–213.

Koh, N., Hu, N., & Clemons, E. (2010). Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electronic Commerce Research & Application, 9*(5), 374–385.

Leung, H. (2001). Quality metrics for intranet application. *Information & Management, 38*(3), 37–152.

Li, X., & Hitt, L. (2010). Price effects in online product reviews: An analytical model and empirical analysis. *MIS Quarterly, 34*(4), 809–831.

Li, M., Huang, L., Tan, C.-H., & Wei, K.-K. (2013). Helpfulness of online product reviews as seen by consumers: Source and content features. *International Journal of Electronic Commerce, 17*(4), 101–136.

Li, J., & Zhan, L. (2011). Online persuasion: How the written word drives WOM-evidence from consumer-generated product reviews. *Journal of Advertising Research – New York, 51*(1), 239–257.

Lipowski, Z. (1975). Sensory and information inputs overload. *Comprehensive Psychiatry, 16*(3), 199–221.

Liu, Y. (2006). Word-of-mouth for movies: Its dynamics and impact on box office revenue. *Journal of Marketing, 70*(3), 74–89.

Ludwig, S., de Ruyter, K., Friedman, M., Brüggen, E. C., Wetzels, M., & Pfann, G. (2013). More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates. *Journal of Marketing, 77*(1), 87–103.

Mackiewicz, J. (2010). Assertions of expertise in online product reviews. *Journal of Business & Technical Communication, 24*(1), 3–28.

Martin, M., & Sell, J. (1980). The marginal utility of information: Its effects upon decision-making. *The Sociological Quarterly, 21*(2), 233–242.

McCormick, E. (1970). *Human Factors Engineering*. New York: McGraw-Hill Book Company.

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

Nielsen, J. (1999). User interface directions for the web. *Communications of the ACM, 42*(1), 65–72.

Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising, 19*(3), 39–52.

O'Reilly (1980). Individuals and information overload in organisations: Is more necessarily better? *Academy of Management Journal, 23*(4), 684–696.

Poston, R., & Speier, C. (2005). Effective use of knowledge management systems: A process model of content ratings and credibility indicators. *MIS Quarterly, 29*(2), 221–244.

Schroder, H., Driver, M., & Streufert, S. (1967). *Human information processing: Individual's and group's functioning in complex social situations*. New York: Holt, Rinehart and Winston.

Schultz, C., Schreyoegg, J., & von Reitzenstein, C. (2013). The moderating role of internal and external resources on the performance effect of multitasking: Evidence from the R&D performance of surgeons. *Research Policy, 42*(8), 1356–1365.

Simperl, E., Thurlow, I., Warren, P., Dengler, F., Davies, J., Grobelnik, M., et al. (2010). Overcoming information overload in the enterprise: The active approach. *IEEE Internet Computing, 14*.

Sotiriadis, M. D., & van Zyl, C. (2013). Electronic word-of-mouth and online reviews in tourism services: The use of twitter by tourists. *Electronic Commerce Research, 13*(1), 103–124.

Sparrow, B., Liu, J., & Wegner, D. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science, 333*(6043), 776–778.

Sternthal, B., Phillips, L., & Dholakia, R. (1978). The persuasive effect of source credibility: A situational analysis. *Public Opinion Quarterly, 42*, 285–314.

Tobin, J. (1958). Estimation of relationship for limited dependent variables. *Econometrica, 26*, 24–36.

Van Rooy, R. (2003). Quality and quantity of information exchange. *Journal of Logic, Language and Information, 12*(4), 423–451.

Wang, R., & Strong, D. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems, 12*(4), 5–33.

Weaver, W., & Shannon, C. (1963). *The mathematical theory of communication*. Univ. of Illinois Press.

Workman, M. (2007). Cognitive load research and semantic apprehension of graphical linguistics. In *Proceedings of the 3rd Human–computer interaction and usability engineering of the Austrian computer society conference on HCI and usability for medicine and health* (pp. 375–388).

Yaari, E., Baruchson-Arbib, S., & Bar-Ilan, J. (2011). Information quality assessment of community generated content: A user study of Wikipedia. *Journal of Information Science, 37*(5), 487–498.

Yang, Z., & Fang, X. (2004). Online service quality dimensions and their relationships with satisfaction. *International Journal of Service Industry Management, 15*(3), 302–326.

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.