The Research on Effects of Information Content Quality and Consumer Knowledge on Online Review Helpfulness from Readers' Perspectives

Xuan Zhang [1,*], Lingzhi Wang [2]

[1] School of Business, Jiaxing University, Jiaxing 314001, China
[2] School of Economics and Management, Hubei Polytechnic University, Huangshi 435000, China
*Correspondence: zhangxuan623@sina.com

Abstract: Online consumer reviews play an increasingly important role in e-business and the function of filtrating high-quality reviews is indispensable for online commenting system. In order to optimize review recommending and filtrating functions, many scholars conducted studies on the impact factors of perceived review helpfulness from different aspects such as data mining, information systems and consumer behavior areas. This paper explored the impact of review information content quality and consumer knowledge on review helpfulness across laboratory experiment and empirical testing based on real review data from JD.com. Consumer knowledge is an important characteristic of online review readers as the information receivers. The results illustrate that the higher quality the review text is of, the more helpful the review is to readers with high consumer knowledge level, while the effect of which is not obvious for readers with low consumer knowledge level. Both existing studies on review helpfulness and designs for review recommending system mostly focus on the features of reviews and reviewers instead of readers, our findings will contribute to deepening the understanding of readers' roles among scholars and appealing to sellers and review system designers to attach importance to the role of readers.

Keywords: review helpfulness; information content quality; consumer knowledge; product attributes; text analysis

1. Introduction

Accompanied by rapid development of e-business, functions within posting reviews are provided by various e-business platforms, crowd-sourced online rating and reviewing communities and other virtual communities for buyers, sellers as well as potential buyers to communicate with each other. Compared with product information published by sellers, online reviews of purchasers are more credible and persuasive and gradually become the main reference source of consumers' online consumption. The first online review system was established by Amazon in 1995 to encourage purchasers to post comments, after which other e-business websites and reviewing communities build online reviewing systems by degrees and constantly optimize the function. However, cost of consumers' viewing about getting helpful information is heavily increased as massive reviews accumulating online [1], in case of which, recommending and filtrating functions over reviews have been developed in reviewing system. It is achievable to evaluate helpful reviews and selectively present them to consumers quickly through text mining and big data analysis. In order to optimize review recommending and filtrating functions, plenty of researchers studied the characteristics of online reviews which are helpful to consumers from different aspects in data mining, information systems and consumer behavior areas [2-8]. Our research also focused on the characteristics of review texts and readers, excavating the impact of review information content quality and consumer knowledge on review helpfulness.

According to the persuasion researches proposed by Hovland et al. [9], communication effects are influenced by characteristics of content, sources (information producers), and information receivers. We found that most studies focused on the factors of review content characteristics affecting review helpfulness according to previous literatures. Some studies have noticed the influence of reviewers' identity, reputation and status, but few studies deal with influencing factors of helpfulness from the readers' perspectives. Our research found differences in perceived review helpfulness caused by different levels of information content quality among readers with different consumer knowledge levels. As existing review recommendation system designs stick to review features and characteristics of reviewers, we hope to help review system designers and sellers further understand and value the influence of review readers.

2. Literature Review and Hypotheses

2.1. Impact Factors on Online Review Helpfulness

Review helpfulness is the important variable to research review information quality. Hu et al. [10] pointed out that identifying helpful product reviews is helpful for customers to get more valuable information and reduce the time spent on searching for needful information. Mudambi and Schuff [11] were the first to define the helpfulness of online reviews from the perspective of perceived value, which is widely recognized by scholars. They defined that online review helpfulness is the perceived value of consumers over online reviews during making a shopping decision. By reason of the review helpfulness being equivalent to perceived value, it can be either measured by the number of helpful votes from review readers on websites or subjectively measured by perceived potential value of information contained in reviews [3]. In research fields such as information systems and data mining, review helpfulness was usually measured by the number or proportion of helpful votes of a review [4, 12]. In the fields of e-business and consumer behavior researches, online reviews helpfulness was measured by scale of perceived helpfulness of online reviews [6,11,13-14]. Perceived helpfulness is a subjective measurement to the degree of perceived product reviews helpfulness in making a shopping decision. Mudambi and Schuff [11] thought that the helpfulness of such subjective perception and the perceived value of online reviews both were a kind of subjective perception of consumers over online reviews thus there was no substantial difference between them. That is to say, subjective measurement of scale of perceived helpfulness and helpful votes on reviews both are valid measurements.

Researches about review helpfulness were aimed at exploring the impact factors and its influences on reviews helpfulness. Factors appearing in current studies were mainly about review content, reviewers' characteristics, reviewing platform features (such as platform type and recommendation system), and product features (such as product type and brand reputation). Yoon-Joo Park [8] summarized and divided the review features of previous studies into three categories: linguistic characteristics (the number of words, word per sentences, etc.), the content of reviews (positivity/negativity, subjectivity/objectivity, etc.) and other peripheral factors (product rating score, review time, etc.).Product types are mainly parted into two types: experience goods and search goods. For experience goods, it is relatively difficult and costly to obtain information on product quality prior to interaction with it; for search goods, it is relatively easy to obtain information on product quality prior to interaction [11].The most researches about review helpfulness in recent years pay more attention to characteristics of review content, among which characteristics of review text mining is the maximum. Product types mainly appear to be experience goods and search goods, among which mobile phone products are mostly studied. Review helpfulness is generally measured by helpful votes of reviews.

2.2. Review Information Content Quality and Consumer Knowledge

Mudambi and Schuff [11] described review length by the number of words in review text. Many studies also demonstrated that the number of words in a review had an impact on review helpful votes [3-6,11,15]. Wang et al. [16] explained the role of length from the perspective of information content quality as the number of words in a review was associated with the recognizability of the information. The more words in a review, the more information readers are able to obtain, and the more helpful is the review to consumers. However, some studies indicated that effects of review length were conditional and not the more, the better as it is [7]. Ghose and Ipeirotis [4] and Wu and Liu [17] both defined that review information content quality was able to eliminate users' uncertainty about goods and services to some degree and measured it by the number of attributes words of the goods and services in a review. From the perspective of readers' dealing with information, we think although long reviews are usually more informative, it should also be considered that not every word can play a role in eliminating uncertainty and providing more effective information. By contrast, the number of product attribute words included in review text is more likely to reflect the diagnostic value of review information and review information content quality. Therefore, review information content quality in this study is calculated by the number of words on attribute descriptions of products and services contained in review text.

Park and Lee [18] divided EWOM into two forms which were attribute-based and experience-based by their nature. An attribute-based product review centers on examining product attribute (e.g., memory capacity of a laptop), which is more specific, rational and objective. An experience-based product review contains the descriptions of consumers' firsthand experiences after purchasing and using products, which is more abstract, subjective and affectively evaluated. Park and Lee [18] discovered that the attribute-based reviews were more helpful and convincing for search products. Huang et al. [14] found that experience-based reviews helped consumers perceive higher helpfulness and paid less cognitive efforts for experience products. To avoid interference of product type, only search products were adopted in our studies. Mobile phone products are the most popular search products, thus both Study 1 and Study 2 regard reviews of mobile phone products as research materials.

The influence of readers' characteristics have been neglected all long in the studies of factors influencing review helpfulness. Park and Lee [18] found that readers' online shopping experience and involvement had a positive impact on online purchasing intentions. Gupta and Harris [19] studied and analyzed the impact of EWOM receivers' information processing motivations on their information process behavior. From the perspective of cognitive science, the personal characteristics of readers as information receivers are the important internal factors [20]. Consumer Knowledge acts as an important characteristic of information receivers [21-22], which is presented to affect decision-making framework [23] and information process (including product information collection and processing) in existing findings [21,24-26]. Consumer knowledge includes consumer understanding of product structure, materials used in the product, and existing product technologies in the market. The more knowledgeable consumers become, the better they understand the products structure and technical relevance [27]. More commonly consumer knowledge is distinguished into objective knowledge and subjective knowledge [21-22]. Objective knowledge refers to product information stored in a consumer's long-term memory, while subjective knowledge refers to the degree of consumer perceived understanding of the product and they are both measured by self-reporting [22]. Self-reported subjective knowledge was used to measure consumer knowledge in the study of Beatty and Smith [28]. Alba and Hutchinson [24] measured consumer knowledge by familiarity, experience, and expertise to the product. Cowley and Mitchell [29] combined previous definition and developed a consumer knowledge scale consisting of four items of familiarity, experience, skillful and subjective knowledge. Study 2 of this paper applied this scale to measure consumer knowledge.

2.3. Hypotheses

As mentioned above, the number of product attribute words included in review text is more representative for review information content quality than the number of review words. The diagnostic value of information largely weights the value of information, namely the helpfulness. Feldman and Lynch [30] insisted that diagnostic value of information was the helpful degree in which relevant information emerging in consumers' minds helped them make judgments or decisions. If the information assists consumers better in understanding the product and making a right decision, it has a relatively good diagnostic value. On the contrary, the diagnostic value is relatively poor. Generally the information that contains more detailed product attribute words or more specific attitude tendency was in a higher diagnostic value and relatively affected other people's shopping decisions more easily. While processing information readers required more powerful argument information which was provided directly by the more product attribute words involved in the review text [16]. As products were made up of a series of attributes and online products reviews mostly presented users' evaluation of product attributes [31], readers were inclined to focus on those attributes and evaluation words, and then evaluated product attributes by referring to related information when purchasing [32]. As a result, readers would estimate that reviews containing more product attribute words were worth of higher review information content quality and greater helpfulness. As the review information content quality was measured by the number of product attribute words in this study, Hypothesis 1 is stated as followed:

H1: High information content quality reviews are perceived to be more helpful to readers than low information content quality reviews

In the field of consumer behavior consumers with low consumer knowledge are called novices, and consumers with high consumer knowledge are called experts. Park and Lessig [33] found that experts were more likely to grasp product attributes. Park et al. [22] studied more deeply and explained product knowledge structure helped experts to get necessary information relatively easily; while novices could not grasp information related to product attributes for lack of basic product cognitive structure knowledge. More researches revealed that experts preferred to process information through Schema-based system while collected information and evaluated product basing on attitude. By contrast, novices tended to collect and process external information as well as were more susceptible to contextual information and persuasion models [21,25-26].

Combining all the previous findings, we infer that the more product attribute words in a review, the more likely experts are able to grasp and judge the product attribute information because of better product-related knowledge structure and consider the review to be more helpful. On the contrary, reviews with low information content quality are not very helpful to experts. Due to the lack of basic knowledge of product cognition structure, novices don't attach much importance to product attribute information in a review, instead of which, they may pay more attention to external and contextual information of irrelevant product attributes. Therefore, effects of product attribute words in a review text on review helpfulness are not obvious to novices. Based on which, Hypothesis 2 is proposed:

H2: Consumer knowledge moderates the impact of review information content quality on perceived helpfulness

H2a: High information content quality reviews is significantly more helpful than low information content quality reviews to experts.

H2b: Difference on perceived helpfulness between high information content quality reviews and low information content quality reviews is not significant for novices.

In addition, some researches implied the impact of posting time of reviews and responses of consumers to reviews [3, 34]. And there were other influence factors of peripheral characteristics of reviews on helpfulness, such as pictures reviews, videos reviews, additional reviews, etc. [17]. As a consequence, these impact factors were under control in Study 1 and Study 2.

3. Study 1: Impact of Review Information Content Quality on Review Helpfulness

In Study 1 the reviews were crawled from websites and a dataset was set up by extracting data from the crawled reviews for empirical analysis. To verify H1 empirically, the text analysis software was used to extract and count the product attribute words in each review, then the effect on helpful votes of the review was validated. In view of consumer knowledge, the moderator variable in H2 which defined the degree of consumer's self-perception about their products familiarity [28], it was hard to acquire self-reported data of readers' consumer knowledge from website reviews. In consequence, Study 1 only verified H1, and Study 2 validated both H1 and H2 through self-reported data investigated in laboratory experiment.

3.1. Data Collection

The dataset in Study 1 was originally collected from JD.com, the largest B2C platform in China. In order to exclude the influence of product type on review information content quality and review helpfulness, only reviews of search products were collected in our research. We used Octopus, the data crawling software to capture the top ten mobile phones sorted by the number of reviews in Mobile Phones on JD.com, gathering 300 reviews for each with 3000 reviews in total. The prices of these ten mobile phones range from 600RMB to 6,000RMB, covering mainstream models on sale in different positioning markets. All these mobile phones are self-managed products on JD.com with the same after-sales and logistics services. Within 0.98-1.45 million reviews in total for each of the ten mobile phones, these millions of reviews are able to avoid differences on amount of reviews of different products affecting research results effectively. Lukyanenko et al. [35] advocated that it was necessary to exclude interference of invalid and false reviews when studying the information quality of User-Generated Content (UGC).There is a review recommendation system on JD.com by which readers can browse reviews in a recommended order on the review page. The recommendation system would screen out and block invalid and false ones automatically among valuable reviews and sort the remaining reviews through series data mining and analysis. Thus we choose to capture reviews presented in recommended order as original dataset in Study 1.

Min et al. [36] mentioned that the manipulation of website's reviewing system affected the probability of each review being viewed and its credibility. The intervention of this system manipulation included the main factors such as recommending, refining, and authenticating of the website on reviews, as well as rearranging on sorted reviews which do not sort reviews chronologically. In Study 1 we effectively circumvented the inconstant impact of recommending system due to taking reviews captured in recommended order mode through Octopus. And only the top 300 reviews were adopted to eliminate impact from review sorting by recommendation system. On the one hand, it's nearly impossible for readers to view reviews after 300; on the other hand, it is found that sorted reviews recommended by website are dynamically updated every day after tracking review pages from JD.com. It guarantees all these top 300 reviews of each product we captured to be read in similar probability.

Typically, a review in captured review data contains the followings information: (1) reviewer account name, (2) reviewer status, (3) star rating, a five-point scale(with 1-5 corresponding to poor, average, good, very good, and excellent), (4) the time when the review was posted, (5) text review (mostly in Chinese), (6) pictures review number; (7) videos review (only one video permitted on JD.com),(8) additional review content and time; (9) "thumb up" votes (likes), (10) the number of responses to review. Filtering out 27 irrelevant and similar reviews, 2973 valid reviews were obtained in the end as the original dataset of Study1.

3.2. Variable Design and Analysis

3.2.1. Variable design

To validate influence of review information content quality over review helpfulness in Study 1, review helpfulness was taken as the explained variable which was measured by the number of helpful votes in a review ("thumb up" votes on JD.com). There are 8 explanatory variables in total, among which the review information content quality is the core explanatory variable, measured by the number of attribute words related to product attributes in review text. The variable explanations of all variables are shown in Table 1.

Table 1. Explanation of the research variables

Variable nature	Variable	Variable explanation	Abbr.
Explained variable	Review helpfulness	Number of helpful votes in review (votes in JD.com is "thumb up" meaning likes)	RH
Explanatory variables	Review information content quality	The number of attribute words related to product attributes in review text	RICQ
	Review length	Word count of review text	RL
	Response to review	The number of responses to the review	RTR
	Review time	The number of days between review posted time and review crawled and extracted time	RT
	Picture review	Number of images in comments	PR
	Video review	Video comment is recorded as 1, otherwise as 0	VR
	Additional review	Additional comments is recorded as 1, otherwise as 0	AR
	Reviewer status	The member of JD.com is recorded as 1, otherwise as 0	RS

In Table1 only the review information content quality is measured by text analysis. The process is as the following steps. First, we constructed the corpus of attribute words related to mobile phone attributes. Second, to finish the work of word segmenting and attribute words corpus tagging, ICTCLAS was used. ICTCLAS is a Chinese text analysis software program developed by Doctor Zhang Huaping. The corpus we constructed was imported in ICTCLAS and then the review text from dataset was segmented and tagged. Finally, the tagged attribute words related to mobile phone attributes in each review text sample was counted and its number was as the measurement of the review information content quality. Wu and Liu [17] ever constructed the corpus of attribute words related to mobile phone attributes in their research. Their corpus gathered the extracted attribute words related to mobile phone attributes, sales and logistics services from the reviews of ZOL.com and Amazon.com (China). The ZOL.com is the largest crowd-sourced online rating and review community of IT products in China. Their corpus can be regarded as a more perfect and detailed Chinese corpus about the mobile phone products attributes. When we constructed corpus in this study, we first extracted keywords from all review text in our research dataset by ICTCLAS. Then we manually compared the extracted keywords with the Wu and Liu [17]'s corpus and supplemented the new words in the latter. Finally we obtained the extended corpus used in our study.

3.2.2. Analysis method

Table2 showed the description statistics results of sample data. Among them, review helpfulness, explained variable, was counted data, the min value of which was zero and the max was 916. There were 901 reviews that got 0 “thumb up” vote, taking 30.32% of all; 495 reviews got 1 vote, which took 16.65%. Reviews with “thumb up” votes among 2-20 reached 42.61% and only 10.42% of the reviews' votes were over 20. It was more prominent for the votes of the number of responses to reviews, a explanatory variable. Reviews with 0 response were occupied 47.4% and 44.2% of the reviews covered responses within 2-10. Only 8.4% of the reviews' responses exceeded 10. It showed an obvious discrete feature. The max value of RICQ was 112 and the min was 0. Reviews containing within 20 product attribute words reached 97.7%, and most reviews carried 4 or 5 attribute words.

Table 2. Descriptive statistics of variables

Variable	Abbr.	Mean	Std.Dev	Min	Max
Review helpfulness	RH	9.9617	35.5128	0	916
Review information content quality	RICQ	7.9983	4.8512	0	112
Review length	RL	101.6044	65.8204	28	500
Response to review	RTR	4.6885	21.1709	0	604
Review time	RT	199.8881	200.5775	0.4306	971.2667
Picture review	PR	2.4369	2.0735	0	10
Video review	VR	0.3031	0.4597	0	1
Additional review	AR	0.0350	0.1838	0	1
Reviewer status	RS	0.4575	0.4983	0	1

As it is not able to meet the demand of OLS regression model, three kinds of regression model are widely used in metrological analysis for the counting characteristics of interpreted variables, including Poisson regression, negative binomial regression, Zero-expansion Poisson regression or Zero-expansion negative binomial regression. Poisson regression requires the mean and variance of the sample data to be equal. Zero-expansion Poisson regression or zero-expansion negative binomial regression is mainly applicable to the case where most zero values exist in the sample data. By further examining the data distribution of the explained variable variable, it is found that the sample data variance (1261.16) is much larger than the mean (9.96) in the case of only 30.3% of the zero value existing. Therefore, the negative binomial regression model is more suitable.

Model setting is shown in Formula (1).

$$RH = \beta_0+\beta_1 RICQ+\beta_2 RL+\beta_3 RTR+\beta_4 RT+\beta_5 PR+\beta_6 VR+\beta_7 AR+\beta_8 RS+\varepsilon, \qquad (1)$$

Incomplete multicollinearity between variables would reduce the accuracy of parameter estimation and increase the variance of the estimator, resulting in the failure of the test conclusion. We used a correlation coefficient test to perform a multicollinearity test on explanatory variables. Table 3 shows the matrix of the perforation correlation coefficient of the explanatory variables. The results show that the correlation coefficient between the review information content quality and the review length reaches 0.7861. There is a collinearity problem between the two variables, and the correlation coefficients of other variables are less than 0.5.

Mudambi and Schuff [11] proposed the text length was able to represent quality of review arguments. Many researches demonstrated that the number of words in the review text had an impact on the review helpfulness [3-6,11,15]. Ghose and Ipeirotis [4] measured the quality of reviews information content by the number of product attribute words included in the review text, and considered it should be more helpful to measure the quality of the effective review information than the absolute number of words of the review. Study 1 took the same measurement as Ghose and Ipeirotis [4] while set up Formula 2 (Model 1 that retains the RICQ and rejects the RL) and Formula 3 (Model 2 that replaces the RICQ by the RL) for regression analysis separately to compare the effects.

Table 3. Pearson correlation coefficient matrix

	RICQ	**RL**	**RTR**	**RT**	**PR**	**VR**	**AR**	**RS**
RICQ	1.0000							
RL	0.7861	1.0000						
RTR	0.0847	0.1302	1.0000					
RT	0.1859	0.2225	0.0167	1.0000				
PR	0.2115	0.2033	0.0359	0.1886	1.0000			
VR	-0.1273	-0.0917	0.1519	-0.0588	-0.4874	1.0000		
AR	0.0423	0.0617	-0.0717	-0.0181	0.0535	-0.0060	1.0000	
RS	0.0635	0.0149	-0.0127	-0.0205	0.0892	-0.0987	0.0016	1.0000

Model1 setting is shown in Formula (2).

$$RH = \beta_0+\beta_1 RICQ+\beta_2 RTR+\beta_3 RT+\beta_4 PR+\beta_5 VR+\beta_6 AR+\beta_7 RS+\varepsilon, \quad (2)$$

Model2 setting is shown in Formula (3).

$$RH = \beta_0+\beta_1 RL+\beta_2 RTR+\beta_3 RT+\beta_4 PR+\beta_5 VR+\beta_6 AR+\beta_7 RS+\varepsilon, \quad (3)$$

3.3. Results and Discussion

Table 4 depicts the results of regression analysis based on Model 1 and Model 2. From the results of regression analysis in Model 1, we can see that the estimated coefficient of variables are all positive except the additional review which is negative (-0.1399) and unconspicuous. The estimated coefficient of the PR and RS is not significant while that of the RICQ, RTR, RT as well as VR on remains significant at the level of 1%. The RICQ has a significant positive effect on the RH which means that reviews including more attribute words get more helpful votes, thus H1 makes sense. From Model 2 which specified the RL instead of the RICQ, we can see that the positive or negative sign of estimated coefficient of the variable as well as its significance stay aligned to Model 1, and the gap between Model 1 and 2 of the Loglikelihood remains small, which means that the two models' overall effects of fitting are acceptable. However, in regard to the degree of effects of core explanatory variables on review helpfulness, the estimated coefficient of RICQ (0.049) in Model 1 remained much higher than that of RL (0.0041) in Model 2. It illustrates that the RICQ is better predicted for the RH than RL and further confirms H1. Study 1 illustrated that the review information content quality had positive effect on review helpfulness votes. As it hasn't shown the influence of readers such as consumer knowledge yet, we further carried out Study 2 which was a lab experiment.

Table 4. Regression results

	Model 1			**Model 2**		
	Coef.	**Std.Err.**	**Z**	**Coef.**	**Std.Err.**	**Z**
RICQ	0.0490^{***}	0.0070	6.97			
RL				0.0041^{***}	0.0005	8.73
RTR	0.0896^{***}	0.0106	8.47	0.0879^{***}	0.0105	8.4
RT	0.0005^{***}	0.0001	3.83	0.0005^{***}	0.0001	3.35
PR	0.0004	0.0153	0.03	0.0004	0.0154	0.23
VR	0.4019^{***}	0.0811	4.96	0.4027^{***}	0.0804	5.01
AR	-0.1399	0.1440	-0.97	-0.1201	0.1524	-0.79
RR	0.0331	0.0582	0.57	0.0428	0.0583	0.73
_CONS	0.5686^{***}	0.0809	7.03	0.5500^{***}	0.0760	7.23
Loglikelihood	-7904			-7897		

$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

4. Study 2: The Moderator Effect of Consumer Knowledge

The simulated laboratory experiment was performed in Study 2 to further validate the impact of review information content quality on the perceived helpfulness, as well to verify the moderator effect of consumer knowledge from readers' characteristic perspective. In the experiment, the participants were asked to image they were browsing online shopping websites to purchase a mobile phone. Then they were informed of product descriptions and the treatment review which was a well-designed product review. Finally participants reported the related variables and completed basic demographic information to end experiment procedure.

4.1. Stimulus Materials and Pretest

Preparation of stimuli for Study 2 involved the product descriptions and a review with manipulation over review information content quality. As reasons mentioned in Study 1, this study continued to select one typical search product, mobile phone as the formal experimental product. The online shopping website for studying still was JD.com. In consideration of participants being all undergraduate students, researchers chose a mobile phone named " Huawei Honor 9 Youth " targeting on young people at about 1,000 RMB from well-sold list of *Mobile Phones* on JD.com as the experiment product. We captured its product descriptions from its product page as experiment About the design of product reviews, after referring to a large quantity of real reviews of *Huawei Honor 9 Youth* mobile phone on JD.com, researchers adapted a real review of this mobile phone by adding some product attributes words such as appearance, tactility, screen, photographing, system, booting speed, headset, charger cable, cost performance and price to it. These product attributes words were picked up among common product attributes words in our corpus in study1 meanwhile in comparison with those in real reviews of this product. The manipulation of review information content quality was to control the number of attributes words above. The high information content quality reviews involved all the above product attributes information, while the low information content quality reviews only carried features about screen, cost performance, appearance, price, quality and charger cable. Therefore, there were 10 attribute words in high information content quality reviews and just 6 attribute words in low information content quality reviews in Study 2. We controlled the number of words of the review in the two groups to be equal in order to avoid interference of review length on results. To supplement the number of words of the low information content quality review, it was along with some feeling descriptions about the product when users were chatting with family members which do not belong to product attributes words. Other aspects of the review, such as star rating, review time, number of images and videos, votes of “thumb up” and responses, etc., were designed to be the same for both treatment reviews, controlling effects caused by the differences of these factors.

Researchers recruited 40 undergraduate students (the average age is 19.9; half male and half female) to conduct a pretest to ensure the validity of the stimulus materials. Those participants were randomly assigned into two groups, high-quality information content group and low-quality information content group, with 20 people for each. As the cover story in the experiment procedure, descriptions of a *Huawei Honor 9 Youth* mobile phone were introduced to participants and the product review manipulated was in turns. The cover story told them to image they chose a mobile phone from JD.com, and then descriptions and a review of this product would be shown to them to help make a purchasing decision. After reading the descriptions and treatment review, participants reported their perceptions of review valence, credibility of review, and manipulation check of information content quality. At the end of procedure they completed the basic demographic information lists. Perceived review valence was measured by a five-point scale ranging from 1 to

5 on behalf of complete negative, half negative, neutral, half positive and total positive respectively as using one item as Jin [37]. And four items were used for perceived credibility as Cheung et al. [38]. The measurement of perceived information content quality used one item adapted from the definition of Wu and Liu [17] and the measurement of product attribute words was to count its number as “The review provides lots of product attribute information” according to Ghose and Ipeirotis [4]. The latter two variables were measured by a seven-point scale ranging from” complete disagreement” to “total agreement”.

Through ANONA analysis, it was found that there wasn’t significant difference in review valence between high-quality information content group and low-quality information content group (M_{High} =3.950, M_{Low} =3.700, F(1,38)=1.568, p=0.218) . Namely the participants of the two groups considered the reviews to be positive with no much difference on degree, and it was successful to control the effect of valence. In the meantime the same to the credibility of review for both groups (M_{High} =5.163, M_{Low} =5.263, F(1,38)=0.091 p=0.765) and reliability testing presents Cronbach’s α=0.767, indicating adequate internal consistency reliability for this variable measurement. It meant that participants of two groups didn’t think the reviews were fake and there was not significant difference in the degree of credibility. As Lukyanenko et.al [35] advocated that it was necessary to exclude interference of invalid and false comments when studying the information quality of User-Generated Content (UGC). The pretesting results displayed that rewritten reviews content in the two groups successfully eliminated interference caused by fake reviews. For perception of information content quality, there was a significant difference rising within the two groups (M_{High} =5.900, M_{Low} =4.55, F(1,38)=21.147, p=0.000), while perception over high informative content quality review was significantly better than that over low information content quality review. From all above, the manipulation and design of stimulus materials of reviews were successful and met the requirements of the formal experiment.

4.2. Experiment Design and Variable Measurement

The formal experiment was a review information content quality level (high vs. low)×consumer knowledge (continuous variables) between-subjects design. 120 undergraduates from a Jiaxing university in China participated in this experiment at one laboratory of the college. These participants all carried online shopping experience and were randomly divided into two groups, high-quality review group and low-quality review group, with 60 people for each. There were 105 valid samples finally (the average age is 20.3; 59.5% female) after excluding unfinished and invalid questionnaires, including 53 in high-quality group and 52 in low-quality group. The procedure of formal experiment was the same with the pretest except that the variables were reported by participants. In formal experiment the perception of review information content quality, helpfulness, consumer knowledge and credibility of review were measured. All items of variables measurement of pretest and formal experiment and their literature sources are shown in Table5. Except the review valence, all variables were measured by a seven-point scale ranging from “complete disagreement” to “total agreement”.

Table 5. Variables measurement and literature sources

Variables	Measurement items	Sources
Review valence	Regarding to this review, reviewers' attitudes towards the phone	Jin [37]
Perception over information content quality	This review offers much product attribute information	Summarized by Ghose and Ipeirotis [4]and Wu and Liu [17]
Credibility of review	This review is reliable	Cheung et al. [38]
	This review is impartial	
	This review is faithful	
	This review is credible	
Perceived helpfulness	This review is helpful to me	Davis [13]
	This review helps me to know the phone more effectively	
	This review helps me a lot to make a purchasing decision	
Consumer knowledge	I am much familiar with the phone	Cowley and Mitchell [29]
	I am more knowledgeable to the phone	
	I am more skillful to use the phone	
	I have rich experience to judge the reviews of the phone	

4.3. Analysis and Results

4.3.1. Manipulation Check and Testing of Validity and Reliability

Manipulation check results revealed significant difference in perception information content quality of these two groups according to 105 valid samples obtained from the formal experiment. The high-quality review was obviously perceived to be more informative than low-quality review (M_{High} =6.226, M_{Low} =4.481， $F(1,103)$= 146.423, p=0.000). The results about credibility of review of two groups were approaching (M_{High} =5.7311, M_{Low} =5.5625, $F(1,103)$=0.900, p=0.345), which signed that participants considered both high-quality and low-quality reviews to be true and reliable with no significant difference on degree of reality. Thus the manipulating and controlling on the information content quality in the formal experiment were successful.

Table 6 comes out with the results of reliability and validity testing of perceived helpfulness, consumer knowledge and credibility of review in formal experiment. Cronbach's α for these variables were greater than 0.7, demonstrating adequate internal consistency reliability for these constructs. KMO were almost all greater than 0.7 and Bartlett's test were significant, stating that each variable was suitable for factor analysis. Within the component matrix (just extracting one principal component), loadings of items on the principal component were higher than 0.7 in credibility of review and perceived helpfulness while loadings of items were higher than 0.6 in consumer knowledge. It demonstrated that these scales contained good convergent validity.

Table 6. Results of reliability and validity testing

Scales	Item No.	Cronbach's α	KMO	Sig. of Bartlett's test (Approx-Square)	Loadings of items
credibility of review	4	0.804	0.761	0.000 (131.035)	0.799; 0.762; 0.810; 0.809
Perceived Helpfulness	3	0.792	0.687	0.000 (95.421)	0.824; 0.819; 0.880;
Consumer Knowledge	4	0.790	0.720	0.000 (128.566)	0.847; 0.757; 0.830; 0.694

4.3.2. Results Analysis

Firstly, ANONA analysis was performed to examine the difference in perceived helpfulness across treatment reviews to verify H1. It was found that high-quality review was considered to be more helpful to participants (M_{High} =4.359, M_{Low} =3.641, $F(1,103)$=10.323, p=0.002) and H1 was verified.

Secondly, moderator effect of consumer knowledge was validated through regression analysis on perceived helpfulness through Bootstrap method and PROCESS program proposed by Hayes [39]. In the PROCESS, we chose Model 1 with review information content quality as independent variable, consumer knowledge as moderator variable and perceived helpfulness as dependent variable. Bootstrap samples was set as 5,000 and confidence level for confidence intervals was 95%, meanwhile choosing Bias Corrected within Bootstrap CI method. Consumer knowledge, the continuous variable, was split into high-level and low-level samples representing experts and novices respectively. After the conditioning Pick-a-Point was set as Mean and +/- SD from Mean, the program of PROCESS automatically split the participants by mean and +/- standard deviation according to Spotlight method [40]. In addition we marked 0 as high-quality review sample level, 1 as low-quality review sample level in the regression analysis.

Results showed there was significant negative moderator effect of consumer knowledge by predicted interaction (β= -0.3918, t=-4.1305, p=0.0001). Examining the results on one standard deviation above and below the mean provides more insight into the pattern of results (shown in Fig. 1). Among participants with high consumer knowledge level (experts), negative moderator effect was remarkable as β= -0.6296 (t=-5.2654, p<0.0000, 95%CI= [-0.8668, -0.3924]), which implied experts agreed that the more product attribute words in review text, the more helpful this review was. Among participants with low consumer knowledge level (novices), there was insignificant effect of review information content quality as β=0.0914 (t=0.7398, p=0.4612, 95%CI= [-0.1537, 0.3366]), which declared attribute words in review text exerted little influence on perceived review helpfulness for novices. Therefore, H2, H2a and H2b were all verified.

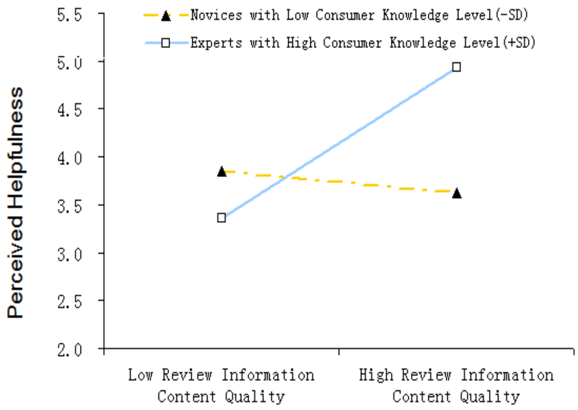

Figure 1. The moderator effects of consumer knowledge

5. Conclusions and Future Work

In this paper, two conclusions were conducted. First, we examined the influence of review information content quality on review helpfulness across laboratory experiment and empirical testing on real review data, and results of these two measurements on perceived review helpfulness and helpful votes were consistent. Our research weighted the quality of review information content through the number of product attribute words contained in a review text. In Study 1 we respectively applied the number of product attribute words and length of reviews as explanatory variables to establish model for helpful votes of reviews. Comparing two models it was discovered that the number of product attribute words could better predict helpful votes than the length of review. In Study 2, we still obtained a conclusion about the number of product attribute words in review text positively affecting perceived review helpfulness on the condition of the length and other review features being controlled. It suggests that the effect of review information content quality on review helpfulness existed separately from length and other impact factors. It meant great importance was supposed to be attached to the impact of review information content quality on review helpfulness, rather than only recognizing the effect of review length and ignoring the role of effective information quality involved in reviews as existing researches. And the results of this study also supported the opinion of Ghose and Ipeirotis [4], which was, the number of product attribute words of review text was more appropriate to measure review information content quality than word count of a review. Second, it validated the moderation role of readers' consumer knowledge in the laboratory experiment of Study 2, which demonstrated that readers with high consumer knowledge level were more sensitive to differences in high/low information content quality of reviews, perceiving greater helpfulness within high information content quality review and less helpfulness within low information content quality review. While product attribute words of review text insignificantly affected perceiving review helpfulness for readers with low consumer knowledge level. The result explained the difference of impacts on review helpfulness existed due to different characteristics of readers. Thus the role of readers' personal characteristics should be paid more attention to in related researches on influence factors of review helpfulness. After all, both existing studies and review recommending system designs mostly focused on review features and reviewers' characteristics while ignored readers' features. Through the conclusion of our study it was expected to deepen the understanding of readers' roles among scholars and appeal to sellers and review system designers to pay more attention to the role of readers.

There are several limitations of this study. First, this paper only studied the reviews of search goods such as mobile phones in order to focus on research topic. Whether the research conclusions are applicable to other products and product categories, it requires further verification. Second, in this study, only the number of product attribute words of review text was selected to measure review information content quality. Although being found more suitable for measuring review information content quality than review length by comparison, it was not considered in combination with other review content characteristics for attribute words. It makes measurement dimension of review information content quality relatively simplex. With the maturity of review text mining technology, we are able to take multiple dimensional composite construction of measurement of review information content quality into account in future, which greatly elevates its prediction effects on review helpfulness. Finally, we only implemented the study by using datasets from JD.com. In order to get more general results, we would like to expand the study over various datasets.

There are several possible future works related to this study. First, delving into the constituent elements of review information content quality variable through collecting more text content features by text mining technology, and constructing a more effective measurement of review information content quality to enhance prediction effects on review helpfulness, which would be meaningful. Second, verifying research conclusions among more product categories, or seeking out more suitable text content features to measure review information quality over different product categories, which would contribute to optimizing existing review recommendation system. Lastly, in terms of different characteristics of readers, we would like to consider designing a customized review recommendation system based on readers' characteristics, which shall enormously strengthen utilization efficiency of review information and achieves a win-win development between counter parties and network platforms.

Acknowledgment

This research was supported by the research projects founded by the Humanity and Social Science on Young Fund of the Ministry of Education of China, grant No. 14YJC630204.

References

[1] Jones, I.A.; Kim, K.Y. Systematic service product requirement analysis with online customer review data. Journal of Integrated Design and Process Science. 2015, 19, 25-48.
[2] Cao, Q.; Duan, W.; Gan, Q. Exploring determinants of voting for the 'helpfulness' of online user reviews: A text mining approach. Decis. Support Syst. 2011, 50, 511–521.
[3] Pan, Y.; Zhang, J.Q. Born Unequal: A study of the helpfulness of user-generated product reviews. J. Retail. 2011, 87,598-612.
[4] Ghose, A.; Ipeirotis, P. G. Estimating the helpfulness and economic impact of product reviews: mining text and reviewer characteristics. IEEE Trans. Knowl. Data Eng. 2011, 23, 1498-1512.
[5] Korfiatis, N.; Garcia-Bariocanal, E.; Sanchez-Alonso, S. Evaluating Content Quality and Helpfulness of Online Product Reviews: The Interplay of Review Helpfulness vs. Review Content. Electron. Commer. Res. Appl. 2012, 11, 205–217.
[6] Yin, D.Z.; Bond, S. D.; Zhang, H. Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. MIS Q. 2014, 38, 539-560.
[7] Kuan, K.K.; Hui, K.L.; Prasarnphanich, P.; Lai, H.Y. What makes a review voted? An empirical investigation of review voting in online review systems. J. Assoc. Inf. Syst. 2015, 16, 48-71.
[8] Yoon-Joo, P. Predicting the Helpfulness of Online Customer Reviews across Different Product Types. Sustainability. 2018,10,1735,1-20.
[9] Hovland, C.I.; Janis, I.L.; Kelley, H.H. Communication and persuasion; Yale University Press: Newhaven, USA, 1953.
[10] Hu, Y.H.; Chen, K.; Lee, P.J. The effect of user-controllable filters on the prediction of online hotel reviews. Information & Management. 2016, 54, 728-744.
[11] Mudambi, S. M.; Schuff, D. What makes a helpful online review? A study of customer reviews on amazon.com. MIS Q. 2010, 34, 185-200.
[12] Singh, J.P.; Irani, S.; Rana, N. P.; Dwivedi, Y.K.; Roy, P.K. Predicting the "helpfulness" of online consumer reviews. J. Bus. Res. 2016, 70, 346-355.
[13] Davis, F. D. Perceived Usefulness, Perceived Ease of Use and User Acceptance of Information Technology. MIS Q. 1989, 13, 319-340.
[14] Huang, L.Q.; Tan, C.H.; Ke, W.L.; Wei, K.K. Comprehension and assessment of product reviews: A review product congruity proposition. J. Manage. Inform. Syst. 2013, 30, 311-343.
[15] Schindler, R. M.; Bickart, B. Perceived helpfulness of online consumer reviews: the role of message content and style. J. Consum. Behav. 2012, 11, 234-342.
[16] Wang, Z.S.; Li, H.Y.; Sun, R. Determinants of Votes of Helpfulness for Chinese Online Customer Reviews: A Moderating Effect of Product Type. Management Review. 2016, 28, 143-153.
[17] Wu, J.; Liu, W. A research of factors affecting the perceived helpfulness of online product based on the information adoption theory. Journal of Information Resources Management. 2017, 1, 47-55.
[18] Park, D.H.; Lee, J. EWOM overload and its effect on consumer behavioral intention depending on consumer involvement. Electron. Commer. Res. Appl. 2008, 7, 386-398.
[19] Gupta, p.; Harris, J. How e-WOM recommendations influence product consideration and quality of choice: a motivation to process information perspective. J. Bus. Res. 2010, 63, 1041-1049.
[20] Foard, C. F.; Kemler, D .G. Holistic and analytic modes of processing: the multiple determinants of perceptual analysis. J. Exp. Psychol. 1984, 113, 94-111.
[21] Brucks, M. The Effects of Product Class Knowledge on Information Search Behaviour. J. Consum. Res. 1985, 12, 1-16.
[22] Park, C.W.; Mothersbaugh, D.L.; Feick, L. Consumer Knowledge Assessment. J. Consum. Res.1994, 21, 71-82.
[23] Moorman, C.; Diehl, K.; Brinberg, D.; Kidwell, B. Subjective Knowledge, Search Location, and Consumer Choice. J. Consum. Res. 2004, 31, 673-680.
[24] Alba, J.W.; Hutchinson, J.W. Dimensions of Consumer Expertise. J. Consum. Res.1987, 13, 411-454.
[25] Alba, J.W.; Hutchinson, J.W. Knowledge Calibration: What Consumers Know and What They Think They Know. J. Consum. Res. 2000, 27, 123-156.
[26] Mandel, N.; Johnson, E.J. When Web Pages Influence Choice: Effects of Visual Primes on Experts and Novices. J. Consum. Res. 2002, 29, 235-245.
[27] Schreier, M.; Prugl, P. Extending Lead-User Theory: Antecedents and Consequences of Consumers' Lead Userness. J. Prod. Innovat. Manag. 2008, 25, 331-46.
[28] Beatty, S.E.; Smith, S.M. External search effort: An investigation across several product categories. J. Consum. Res. 1987, 14, 83-95.
[29] Cowley, E.; Mitchell, A. "The Moderating Effect of Product Knowledge on the Learning and Organization of Product Information". J. Consum. Res. 2003, 30, 443-454.
[30] Feldman, J.M.; Lynch, J.G. Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. J. Appl. Psychol. 1988, 73, 421-435.
[31] Bertini, M.; Ofek, E.; Ariely, D. The impact of add-on features on consumer product evaluations. J. Consum. Res. 2009, 36, 17-28.
[32] Lee, T.Y.; Bradlow, E.T. Automated marketing research using online customer reviews. J. Mark. Res. 2011, 48, 881-894.
[33] Park, C.W.; Lessig, V.P. Familiarity and Its Impact on Consumer Decision Biases and Heuristics. J. Consum. Res.1981, 8, 223-230.
[34] Berger, J.; Sorensen, A.T.; Rasmussen, S.J. Positive effects of negative publicity: When negative reviews increase sales. Market. Sci. 2010, 29, 815-827.
[35] Lukyanenko, R.; Parsons, J.; Wiersma, Y, F. The IQ of the Crowd: Understanding and Improving Information Quality in Structured User-Generated Content. Inf. Syst. Res. 2014, 25, 669-689.

[36] Min, Q.F.; Qin, L.; Zhang, K.l. Factors Affecting the Perceived Usefulness of Online Reviews. Management Review. 2017, 29, 95-107.
[37] Jin, L.Y.; The effects of online WOM information on consumer purchase decision: An experimental study. Economic Management. 2007, 29, 36-42.
[38] Cheung, M.Y.; Luo, C.; Chen, S.H. Credibility of electronic word-of-mouth: informational and normative determinants of on-line consumer recommendations. Int. J. Electron. Commer. 2009, 13, 9-38.
[39] Hayes, A.F. An Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-based Approach; Guilford Press: New York, USA, 2013.
[40] Fitzsimons, G. J. Death to dichotomizing. J. Consum. Res. 2008, 35, 5-8.