Study on Information Quality of Online Shopping Reviews

Zheng Dequan and Wang Enyan

Abstract—Online shopping reviews allow the buyers to exchange reviews and experiences for all to see. It has important value in rising sales and retaining customers for e-commerce platforms and the venders to estimate reviews’ information quality. Focus on the incidence relation between online reviews and sales volume, this study collects shopping reviews from eight stores on Taobao and the sales volume of each store, so as to develop an indicator configuration of influencing factors for online reviews. The study takes regression analysis to identify the causation between indicators and sales volume by R language, then determines the indicator weight by AHP method. The maturity evaluation value is determined by fuzzy comprehensive evaluation method (FCE). By comparing two venders’ maturity value, the results show that there is a positive relationship between online reviews quality and sales volume. That is, the higher the maturity score of online reviews quality, the higher the sales volume.

Index Terms—AHP, fuzzy comprehensive evaluation method, information quality, online shopping reviews.

I. INTRODUCTION

According to China Internet development statistics report published by CNNIC (China Internet Network Information Center), China has 989 million Internet user, and the Internet penetration rate is 70.4% at the end of 2020. Currently, the number of instant messaging user reached 980 million and the number of online shopping users reached 782million in China. With the expanding of the scale of online shopping market, the number of online reviews is increasing rapidly in the same time. Estimating reviews’ information quality is provided with important practical value for rising sales and retaining customers for e-commerce platforms and the venders.

Online shopping has replaced physical store shopping to a certain extent for its advantages in convenient shopping, low price and full range of products. It has become one of the irreplaceable shopping modes in today's information society. Compared with traditional shopping method, online shopping possesses some irreplaceable advantages. In the meantime, there are still some uncertainties to be reckoned with in the online shopping. Because of the using convenience of internet technology, consumers often refer to third-party platforms on buyers’ reviews and commodity quality while they are online shopping. On the other hand, the buyer actually may not get the live experience on internet, and that becomes one of the biggest obstacles for shopping online. Then, the shopping reviews will be the effective method for buyers to overcome the information asymmetry. It has become an essential step for online shoppers to reading commodity comments during online shopping. The consumer might decide whether to buy or not based on the reading online reviews. For the venders, those reviews are very valuable reference resources for increasing the commodity quality and improving marketing strategy.

On reading reviews, the most meaningful step is to make a value judgment on the information quality; that is, the consumers will assess whether there is the information needed. Consumers will consciously or unconsciously search for the information they want when reading the comments, such as chromatic aberration, delivery speed. We call this is the information adoption process. Information adoption will strengthen the initiative of the information users, and that is the key reason for the difference between information adoption and information seeking or information retrieval. When the consumers face excessive commodity reviews, the difficulty of information adoption will increase significantly. There are three reasons for this problem: first, the difficulty of receiving information will increase when facing excessive comments; second, unwarranted comment information will cause redundancy of information; third, the false information and erroneous information increases the difficulty and imposes an extra burden for consumers to receive information. Reviews’ information quality has important value for e-commerce platforms. While the information quality is a multi-dimensional complex variable, and it cannot be quantified directly and presented numerically. By configuration model, this study focuses on the research of the dimensions and characteristics of information quality of online reviews.

II. LITERATURE REVIEWS

According to the present research, most studies have focused on the online reviews’ effect on the purchase preference [1]-[4]. Prasad and Anja researched the influence of online appraisals on the buyers’ purchase intention, and their study shows that buyers’ online comments will have a very strong impact on consumers' purchase decisions [5]. Some scholars are interested in which comment form may have the greatest impact on consumer behavior. Emoji is the fastest growing form of comment language for most young consumers. Emmanuela et al. have investigated the effect of emoji in the reviews and the effect of consumer personal characteristics on the reservation intention. The result shows...
that the interactions effect between price comment and emotional emoji will influence the consumer reservation intention [6]. The presence of pictures is also an important reason to impact the buyer’s decision. Woo and Ian examined the consequence of picture reviews on the consumer evaluation. The consumer trust is influenced by pictures, and comments containing pictures are considered more effective decision reference. In particular, when the image quality is higher, the trust level of buyers will be higher. This study confirmed the influence of reviews trust on purchase intention [7].

In recent years, the information quality of online reviews has become the forefront of issues in internet marketing. Clemes researched the quality of online reviews, and the results show that the length of online comments and the information quality of comments have a significant impact on purchase behavior [8]. Consumers generally believe that accessibility and usefulness of online reviews will also have an impact on their consuming behavior [9]. Based on the 192-sample data from a questionnaire, Hu found that the comments related with economy and services have a significant impact on online shopping behavior [10]. Chen and Li investigated the correlation between online comments and online shopping intention using 782-sample data. The results show that after controlling gender, occupation and online shopping age, the higher the quality of online reviews, the stronger the purchase intention [11].

On the topic of information quality, the researchers choose three dimensions to make empirical research: quality of information sources, quality of information content and quality of information expression [12]-[15]. The study of Lin indicted that high quality online comments have a more significant impact on consumers purchasing decision [16]. Based on the perception usefulness, Guo and Qian found that the higher the quality of comments seen by consumers, the more significant the impact on their purchase intention [17].

Do negative reviews have a negative impact on consumers purchasing decision? The research of Park and Kim answered this question. Their results show that not all negative comments will dampen consumers' purchase enthusiasm [18].

The quality of online reviews should be measured according to the impact of online comments on consumers or potential consumers’ decisions. The greater the impact of comments on consumers’ behaviors, the higher the quality of comments.

This paper takes the content and form of online review information as the benchmark dimension of research indicators. The objective feeling dimension of timeliness and the subjective feeling dimension of readability are added into the benchmark model. Then we will get four dimensions of online comment quality as the secondary indicators:

1) Objective authenticity: Objective authenticity mainly refers to the length of comments. Because the length of comments can reflect the quality of consumers' purchase experience to a certain extent. On the other hand, it is whether the online comment information is the real information commented by the consumer under the condition of respecting the objective facts and the inner real feelings, not the false information deviating from facts.

2) Overall specificity: On the one hand, whether the evaluation information completely and specifically describe the attributes of goods and experience information. On the other hand, whether contains specific service or logistics service information other than goods and businesses information. The indicator of overall specificity includes the attributes of the product characteristics, the real experience of use, venders’ services and the logistics service.

3) Diversity of forms: This indicator means that the description of online comment information should not be limited to one type, and various description methods should be considered in the decision process. Taobao e-commerce platform provides many types of comments when designing the comment methods, such as text, pictures, videos, etc. It is very helpful for the potential consumers, if buyers use multiple expressions when writing comments. Also, this conduct will decrease the information asymmetry and help other buyers to fully understand the properties of goods and refrain from impulsive consumption.

4) Timely understandability: This indicator refers that the comment time is limited to a certain range. On the other hand, it means that the comment information should be recognized by potential consumers. The reviews with timely understandability are the information with correct sentence structure and easy to understand for users. The comment information is expressed logically and accurately without ambiguity, which is published within a certain time range.

The indicator of overall specificity is expressed by tow dimensions: use experience and service evaluation. The tow dimensions are expressed by following evaluation factors: product features, use experience, vendors service evaluation, platform logistics service. Those indicators are comprehensive evaluation indicators.

The indicator of diversity of forms includes tow dimensions: various types of comment information and interpersonal interaction. The various types of comment include text, pictures, videos, etc. The interpersonal interaction includes the consumer-consumer interaction and
The indicator of timely understandability is expressed by tow dimensions: clear and understandable expression and timeliness. The indicator means that the content of comments is constructed logically and no contradictory expressions. The timeliness here refers to the different comments within a certain range.

Based on the research on the dimensions and characteristics of reviews information quality and the four principles, this study put forward the first-rank evaluation indicators. By After subdividing the first-rank indicators, we will get the second-rank evaluation indicators. In the same way, we will get the third-rank evaluation indicators: length of comment, truthful expression, using experience, service evaluating, form diversity, interaction function, understandability and timeliness. The evaluation indicators are shown in Table I.

### Table I: Indicators Hypothesis for Information Quality Factors

| First-rank indicator | Second-rank indicators | Third-rank indicators and description |
|----------------------|------------------------|---------------------------------------|
| Objective authenticity | B1 | Length of comment |
| Overall specificity | B2 | Quality of online reviews |
| Diversity of forms | B3 | Interaction function, understandability |
| Timely understandability | B4 | |

B. Hypothesis Test for Influencing Factors of Information Quality

By collecting the vendors data on Taobao platform, this study investigates the impact of information quality of online comment on sales volume. This research focuses on the reviews number when using the web crawler software. The products with insufficient comments are excluded from the data, and the false data and invalid data of each commodity are filtered out. The data are scored by the eight third-rank indicators, and the grade 1-5 scoring system is used. According to the average score for each indicator, we get the monthly sales by crawler software. Then regression analysis was carried out between the score of each indicator and the sales volume. The results are shown in Table II.

As shown in the Table II, there was a positive correlation between indicator score and sales volume. That means the higher score may increase the sales when the other conditions were fixed. With the exception of timeliness indicator (C42), all other indicators passed the test. After eliminating timeliness indicator (C42), the final indicators of influencing factors are as showed in Table III.

### Table II: Regression Analysis of Indicator Score and Sales Volume

| Estimate Std. | Error | t value | Pr(>|t|) |
|---------------|-------|---------|---------|
| C11 Intercept | -28727 | 10581 | -2.715 | 0.042 |
| C12 Intercept | -38344 | 5870 | -6.532 | 0.001 |
| C21 Intercept | -7263 | 10754 | -0.704 | 0.051 |
| C C | -15647 | 10975 | -1.426 | 0.123 |
| C22 Intercept | 15024 | 7474 | 2.010 | 0.011 |
| C31 Intercept | -13493 | 8031 | -1.680 | 0.154 |
| C32 Intercept | -12855 | 10268 | -1.252 | 0.266 |
| C41 Intercept | -44040 | 11523 | -3.822 | 0.012 |
| C42 Intercept | 27462 | 59177 | 0.464 | 0.659 |
| x | 7844 | 3140 | 2.599 | 0.044 |
| x | 12693 | 38344 | 2.715 | 0.042 |

### Table III: Evaluation Indicators of Information Quality for Online Shopping Reviews

| First-rank indicator | Second-rank indicators | Third-rank indicators |
|----------------------|------------------------|-----------------------|
| B1 | Objective authenticity | C11:Length of comment |
| B2 | Overall specificity | C12:Truthful expression |
| B3 | Diversity of forms | C21:Using experience |
| B4 | Timely understandability | C22:Service evaluating |
| | | C31:Form diversity |
| | | C32:Interaction function |
| | | C41:understandability |

C. Calculating Weight Values of Indicators

1) Configuration of fuzzy evaluation matrix

According to the above discussion, the evaluation index set includes A, B and C three-level indicator. A refers to the quality of online reviews; B1, B2, B3 and B4 respectively correspond to objective authenticity, overall specificity, diversity of forms and understandability indicator. C11, C12, C21, C22, C31, C32 and C41 B4 respectively correspond to length of the reviewer’s comment, whether describe the true experience, describing the quality of the products used, experience of logistics service, including pictures, videos comment, follow-up evaluation or vendor responding, clear and understandable expression and length of time the buyer commented.

The relationship between level A indicators and level B indicators is expressed by the following equation:

\[ A = \{B_1, B_2, B_3, B_4\} \]

The relationship between level B indicators and level C
indicators is expressed by:

\[ BI = \{ C_{i1}, C_{i2}, \ldots, C_{ij} \} \quad (i=1, \ldots, 4; j=1, \ldots, 4) \]

The judgment matrix can be got by comparing the different level indicators separately, then the consistency of the judgment matrix is tested.

### TABLE IV: COMPARISON OF QUANTITATIVE VALUES BETWEEN INDICATORS

| Indicator comparing to indicator | Quantitative value |
|---------------------------------|-------------------|
| Equal Importance                | 1.00              |
| Slight importance               | 3.00              |
| Quite importance                | 5.00              |
| Strong importance               | 7.00              |
| Extreme importance              | 9.00              |
| Slight unimportance             | 0.33              |
| Quite unimportance              | 0.20              |
| Strong unimportance             | 0.14              |
| Extreme unimportance            | 0.11              |
| Intermediate value of two adjacent judgments | 2, 4, 6, 8 |

Table IV indicates the comparison of quantitative values between indicators. The judgment matrix of category indicators is expressed by: \( M = (a_{ij})_{n \times n} \). It satisfies the condition of positive reciprocal matrix, and that means: \( a_{ij} > 0 \), \( a_{ij} = 1/a_{ji} \) (\( i, j = 1, 2, \ldots, n \)).

#### 2) Calculating weight values of all level indicators

The AHP method is used to calculate the values of indicators. The weight values \( w_1, w_2, w_3, w_4, w_{11}, w_{12}, \ldots, w_{21}, w_{22}, w_{31}, w_{32} \) and \( w_{41} \) are used to measure the significance of indicators B-level and C-level. This method includes the following steps:

- Calculating matrix values
- Index vector \( W \) for B-level with \( n \) matrix order is indicated by:

\[
W = \sqrt[n]{\prod_{i=1}^{n} r_{ij}} \quad (i, j = 1, 2, 3, \ldots, n) \quad (1)
\]

The consistency test

In order to ensure the accuracy and reliability of the data, the consistency of the judgment matrix is tested after the calculating matrix values. If the test result is \( CR \leq 0.1 \), that means the consistency ratio less than 0.1. And the weight of the judgment matrix is the effective weight. Firstly, the maximum eigenvalue of the matrix is indicated by \( \lambda_{max} \):

\[
\lambda_{max} = \frac{1}{n} \sum_{i=1}^{n} \frac{a_{ij} \times W_i}{W_j} \quad (2)
\]

Secondly, the average random consistency index \( R_i \) can be get by checking look-up table. Among them, \( C_i \) is the consistency index. \( R_i \) is the average random consistency index as shown in Table V. Finally, consistency ratio is calculated by following equation:

\[
C_i = \frac{\lambda_{max} - n}{n-1} \times \frac{1}{R_i} \quad (3)
\]

### TABLE V: AVERAGE RANDOM CONSISTENCY INDEX

| Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-------|---|---|---|---|---|---|---|---|
| Ri    | 0 | 0 | 0.58 | 0.89 | 1.12 | 1.24 | 1.32 | 1.41 |

#### D. Configuration of Index Fuzzy Comprehensive Evaluation Model

1) Configuration of membership matrix for category index

Firstly, it is necessary to determine the evaluation matrix and evaluation set by:

\[
v = (v_1, v_2, v_3, v_4)
\]

These refers to the indicators average, medium, good, excellent. Secondly, the membership matrix for certain category index is determined. The evaluation index are judged by the membership matrix according to the judgement standard, namely the relationship from B to C. The membership matrix is regulated by \( R_i \):

\[
R_i = \begin{bmatrix}
 r_{11} & r_{12} & r_{13} & r_{14} \\
 r_{12} & r_{12} & r_{13} & r_{14} \\
 r_{13} & r_{13} & r_{13} & r_{14} \\
 r_{14} & r_{14} & r_{14} & r_{14}
\end{bmatrix} \quad (i = 1, 2, 3, 4, 5, 6, 7)
\]

In this equation, \( R_i \) \( (i = 1, 2, 3, 4) \) respectively refers to the corresponding membership matrix of objective authenticity, overall specificity, diversity of forms, and timely understandability of online reviews.

Here, \( r_{11}, r_{12}, r_{13}, r_{14} \) respectively correspond to the membership degrees of comment length for the first level (level average), second level (level medium), third level (level good), and fourth level (level excellent). The same is true for other \( r \) values.

2) Configuration of fuzzy comprehensive evaluation model

The membership matrix \( R_i \) and index weight are grounded on the category index scoring. The single factor fuzzy evaluation vector \( B_i \) is calculated based on the following equation:

\[
0 \quad (5)
\]

According to the single factor fuzzy evaluation vector \( B_i \), we can compute the corresponding multi factor fuzzy evaluation vector \( B \):
Finally, the management maturity value of information fuzzy comprehensive evaluation value is determined according to the AHP method. Secondly, the calculation process.

Firstly, the weight values of individual indicators are determined according to the AHP method. Secondly, the fuzzy comprehensive evaluation model is constructed. Finally, the management maturity value of information quality of online reviews based on the maturity formula and the calculation process.

Firstly, the weight values of individual indicators are determined according to the AHP method. Secondly, the fuzzy comprehensive evaluation model is constructed. Finally, the management maturity value of information quality of online reviews based on the maturity formula and the calculation process.

The consistency test for the indicators at all levels is carried out by the following equation:

$$C_i = C_i = \frac{\lambda_{\text{max}} - n \cdot \frac{1}{R_i}}{n - 1} = 0.03526406 \ 3 < 0.1$$

### TABLE VI: CALCULATION WEIGHT

| A  | B1 | B2 | B3 | B4 | Row product | Extraction root n-times | Weight |
|----|----|----|----|----|-------------|-------------------------|--------|
| B1 | 1.00 | 5.00 | 9.00 | 0.50 | 22.5 | 2.177938587 | 0.35704768 |
| B2 | 0.20 | 1.00 | 3.00 | 0.20 | 0.12 | 0.588566191 | 0.096488576 |
| B3 | 0.11 | 0.33 | 1.00 | 0.11 | 0.0041 | 0.253278562 | 0.04152072 |
| B4 | 2.00 | 5.00 | 9.00 | 1.00 | 90.0 | 3.080072028 | 0.504941672 |

The fuzzy evaluation vector of single factor is shown in Table IX.

$$R_i = W_i * R_i = (0.049 0.062 0.281 0.098)$$

$$R_i = W_i * R_i = (0.012 0.054 0.069 0.045)$$

$$R_i = W_i * R_i = (0.024 0.026 0.028 0.002)$$

$$R_i = W_i * R_i = (0.0676 0.026 0.026 0.182)$$

The fuzzy evaluation vector of multiple factors is:

$$B = W * R = (0.033791 0.4154 0.12324 0.1313)$$

Then we will get the product result of fuzzy evaluation vector B and L (maturity level matrix), and can calculate the comprehensive evaluation value for maturity A:

$$A = B * L = 1.025812$$

### B. Policy Suggestions

This study is based on online shopping reviews from Taobao e-commerce platform, and researches the influencing effect of the length of comment, truthful expression, using experience, service evaluating, interaction function, understandability on sales volume. The research result shows that there is a positive relationship between online reviews quality and sales volume. Meanwhile, the information quality of most consumer’s reviews is in a low state (lower than level medium). There is a quite gap between the current quality level and what we expected. In order to improve marketing strategy of the venders, this study proposes the following recommendations:
1) **Encouraging consumers to make actual comments about goods**

The e-commerce platforms should encourage consumers to write down the real experience of shopping rather than fake reviews or order-faking scams. The traditional way for vendors to encourage consumers to comment is adding WeChat and sending red packet money. However, this method is too rampant and is lacking of incentive effect. There should be innovative in incentives mechanism for consumers reviews.

2) **Formulating information schemes based on needs of specific customers**

With the aid of information analysis tools, it is necessary to investigate the multi-dimensional customer information. After clarifying customer’s needs, the scientific marketing strategies should be formulated. Using the network, the vendors may update product plans for customers and improve the pertinence of customers relationship maintenance.

3) **Keeping diversity of communication way**

Customers management is an omnidirectional and tridimensional managing style. The vendors should consider the comprehensive service experience and using convenience for customers. It has significant value to communicate with customers by e-mail, telephone, WeChat. By providing customers with more convenient services, the vendors can develop a sense of customer loyalty.

4) **Prompt responding to reviews**

If the vendor wants more detailed comments, it is a good method to increase the responding times then get the length of customer’s comment. Most studies have proved that picture reviews have a positive impact on sales volume. Thus, the rising of valuable picture reviews is conducive to maintain a good reputation in e-commerce platform.

V. CONCLUSION

Focus on the online shopping reviews from Taobao e-commerce platform, this study investigated the influence in sales volume from length of comment, truthful expression, using experience, service evaluating, form diversity, interaction function, understandability and timeliness. The research result shows that there is a positive relationship between online comment quality and sales volume. The higher the quality maturity score of online reviews, the higher the sales volume. And then, the higher the information quality of online comments, the higher the sales volume. Notwithstanding, the information quality for most online comments is still in a low level. As a significant value of this study.

**CONFLICT OF INTEREST**
The authors declare no conflict of interest

**AUTHOR CONTRIBUTIONS**
Zheng Dequan conducted the research; Wang Enyan analyzed the data; Zheng Dequan wrote the paper; all authors had approved the final version.

**REFERENCES**

[1] A. Rolinn and G. R. Milne, “Cumemeurs privacy concerns about direct marketers use of personal medical information,” in Proc. Association for Health Care Research Conference, vol. 2, pp. 7–9, 2017.

[2] G. Ellen and S. Michal. “Grant differences in the perceived risk of buying online and the effects of receiving a site recommendation,” Journal of Business Research, vol. 57, pp. 768–775, 2014.

[3] Z. Qiang and L. Naihe, “Study on the current situation and future development of internet word of mouth,” Journal of Jiangxi Agriculture, vol. 20, pp. 22–30, 2016.

[4] Y. B. Chert and J. H. Xie, “Online consumer review: word of-mouth as a new element of marketing communication mix,” Management Science, vol. 54, pp. 477–491, 2016.

[5] P. Vana and A. Lambrecht, “The effect of individual online reviews on purchase likelihood,” Marketing Science, vol. 40, no. 4, pp. 593-812, May,2021.

[6] E. M. Emmanouela, E. Mourelatos, and E. Dimara, “Beyond the lexical sense of online reviews: The role of emoticons and consumer experience,” Interacting with Computers, vol. 32, pp. 5-6, 2021.

[7] P. C. Woo, S. Ian, and L. Ki, “Effects of online reviews, trust, and picture-superiority on intention to purchase restaurant services,” Journal of Hospitality and Tourism Management, vol. 47, pp. 228-236, June 2021.

[8] M. Clenes, C. Gun, and J. Zhang, “An empirical analysis of online shopping adoption in Beijing, China,” Journal of Retailing and Consumer Services, vol. 21, pp. 364-375, 2014.

[9] A. R. Kanitra and A. Kusumawati, “Pengaruh country of origin dan online consumer review terhadap trust dan keputusan pembelian,” Admistration Bisnis, vol. 61, pp. 281-291, July 2018.

[10] F. Hu, “The relationship analysis between online reviews and online shopping based on B2C platform technology,” Cluster Computing, vol. 22, pp. 3365–3373, March 2018.

[11] Y. Cheng, X. Li, “The relationship between online comments and online shopping intention: an analysis of moderating mediating effect,” Hubei Agricultural Science, vol.60, pp. 187–191, March 2020.

[12] H. Beales, M. B. Mazis, and S. C. Salap, “Consumer search and public policy,” Journal of Consumer Research, vol. 8, pp.11–22, 2017.

[13] E. Kim and B. Lee, “E-CRM and digitization of word of mouth,” International Journal of Management Science, vol. 11, pp. 47–60,2015.

[14] W. L. Stephen, E. G. Ronald, and P. Bing, “Electronic word-of-mouth in hospitality and tourism management,” Tourism Management, vol. 11, pp. 458–468, May 2017.

[15] H. Xiaojun and X. Weijing, “Basic research orientation of word of mouth communication,” Marketing, vol. 24, pp. 51–56, June 2015.

[16] X. Lin, “An empirical study on influential factors of credibility of internet word-of-mouth,” Financial Research, vol. 5, pp. 113–117 2017.

[17] G. Guo and M. Qian, “Influence of internet word of mouth on consumers,” Economic Management, vol. 5, pp. 56–59, 2017.

[18] D. Park and S. Kim, “The effects of consumer knowledge on message processing of electronic word of mouth via online consumer reviews,” Electron. Commerce, vol. 6, pp. 8–10, 2017.

Copyright © 2022 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).

Zheng Dequan is an assistant professor / lecturer at Harbin Institute of Technology where he is affiliated to the Department of Public Administration. He is involved in public policy and industry economics. Regional development and game theory have his theoretical interest.