Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-026-06884-y

OPEN

The factors that influence online review quality: the example of new energy vehicle reviews

Xiaoguang Wang[1✉], Yijun Gao[2], Lei Fan[1] & Nan Li[1]

China's new energy vehicle industry has gradually shifted from relying on policy support in the past to a state where enterprise survival is determined by the market, which has driven the industry to accelerate its move toward market orientation. This study aims to help potential consumers quickly locate and obtain valuable, high-quality reviews, while also assisting new energy vehicle companies in improving product quality and service levels to promote the healthy development of the industry. To achieve these goals, review data were first collected from online forums, and a corresponding corpus was subsequently constructed. Next, an evaluation system for review quality was established, with values assigned to review quality accordingly. Finally, a model of the influencing factors of review quality was developed; through multiple regression analysis, specific influencing factors were identified, facilitating an exploration of the impact exerted by relevant factors on review quality. The results indicate that the statistical features and semantic features of online reviews can serve as evaluation indicators for constructing a quality evaluation system. In the influencing factor model, the time interval, number of purposes, and reviewers' expertise exert a positive impact on review quality; driving mileage has a negative impact on review quality; while the positive impact of certified users on review quality is not significant. The study results provide a simple criterion for consumers' perception of review values and verify the influence of review quality on consumers' purchase intention. Additionally, they can help automotive companies identify factors affecting the quality of online reviews and enhance the effectiveness of new energy vehicle companies in utilizing online reviews for marketing purposes.

[1] Lianyungang Normal University, Lianyungang, China. [2] Dominican University, River Forest, USA. ✉email: xwang@cumt.edu.cn

Introduction

According to the 54th Statistical Report on China's Internet Development, compiled by the China Internet Network Information Centre (CNNIC), 92% of potential consumers often search for online reviews of relevant products before making a purchase decision. And 77.5% consider online reviews the most critical factor influencing their purchase decisions. To reduce the perceived risk of products, potential consumers often seek relevant signals or clues from online reviews to infer the potential quality of products (Tsai et al. 2020). The review quality primarily refers to how online reviews influence potential consumers' purchasing decisions and shopping preferences. High-quality online reviews often contain objective and rich product information or service experiences, which are highly persuasive (Li et al. 2020a). When high-quality online reviews are constructive and positive, consumers are more likely to increase their willingness to purchase, driven by increased trust, which ultimately translates into actual sales of the product. When high-quality online reviews are passive or negative, consumers can also be persuaded to abandon their procurement, impacting product sales in the opposite direction.

Experts and scholars in various fields have different understandings of the concepts and functions of online review quality (Nelson et al. 2005; Barbara et al. 1997; Wang and Strong 1996). Zhao et al. (2017) elaborated that online review quality refers to a user's evaluation of the system's performance in providing user reviews, based on their user experience. Hu and Krishen (2019) investigated how online reviews can lead to information overload and difficulty, ultimately affecting decision satisfaction. Lu et al. (2023) developed a model to examine the effects of website-generated quality and fit-related review keywords on consumers' clothing purchase decisions. This paper defines online review quality as follows: Online review quality is a combination of the objective presentation of the review content itself and the user's subjective expression of the review content. That is, the higher the objective presentation of the review content and the user's subjective expression, the higher the review quality; the lower the objective presentation of the review content and the user's subjective expression, the lower the online review quality. Therefore, the measurement of online review quality should consider both objective factors, such as the media presentation form of review content and product performance descriptions, as well as subjective factors, including the emotions and semantics expressed by consumers through their reviews.

In the vehicle industry, the replacement speed of vehicle products has gradually accelerated, and product selection has become a decision-making problem that most consumers will face (Liu and Teng 2021). For big-ticket consumer goods, such as vehicles with high prices and strong durability, the high mismatch cost makes potential consumers more inclined to obtain valuable word-of-mouth information through online reviews to aid their purchasing decisions. Efficient mining of valuable information in complex online reviews is essential for potential consumers to understand and evaluate the performance of vehicle products, thereby forming a comprehensive understanding of these products. The Chinese government initiated the New Energy Vehicles (NEV) subsidy policy in 2009 to promote the adoption of clean energy vehicles. However, to encourage market-driven growth and reduce fiscal burdens, the government introduced the subsidy retreat policy in 2020. This policy gradually reduces financial subsidies for NEV manufacturers and consumers, aiming to phase out subsidies entirely by the end of 2022 (Zhang et al. 2021). The subsidy retreat policy marks a significant shift in the NEV industry, as it forces manufacturers to compete based on product quality, technological innovation, and cost efficiency rather than relying on government support. This transition has intensified market competition and accelerated the industry's evolution toward a more sustainable and consumer-driven model (Chen et al. 2023). Based on the above background, this study aims to utilize online reviews of NEVs as the research object and explore the factors influencing online review quality.

Existing studies on the influencing factors of online review quality have two main aspects that need further improvement: 1) Most literatures directly use the number of votes (Sun et al. 2019, Jiang and Zou 2023) or the proportion of votes to the total votes as indicators of review quality (Mudambi and Schuff 2010), and a few literatures use the proportion of useful votes to the number of days since the review was posted as an indicator (Tsai et al. 2020). However, different measurement methods often lead to contradictory results when testing the same influencing factors of usefulness (Hong et al. 2017), which is mainly caused by statistical differences brought about by vote sparsity; 2) The evaluation of review quality mainly focuses on specific dimensions such as relevance and usefulness; the semantic analysis method for online reviews is relatively simple, with problems such as inaccurate quantitative basis and the need for timely updates of the lexicon.

NEVs are high-involvement, hybrid (search/experience) products. To address the two aforementioned shortcomings in the research of review quality influencing factors and demonstrate how electronic word-of-mouth (eWOM) theories perform differently in the unique context of NEV reviews, this study first constructs review quality evaluation indicators from the media dimension and content dimension of online reviews, then uses the entropy weight-TOPSIS method to determine the weights of each indicator and assign values to review quality. Finally, it defines and quantifies more comprehensive influencing factors of review quality, determines the coefficients of these factors through sentiment analysis and machine learning models, and explores the impact of multi-dimensional factors on review quality.

Concepts, models, and related research

Vehicle reviews. Vehicle reviews have become crucial decision-making tools in the consumer journey. Recent studies have demonstrated the significant impact of automotive review characteristics on consumer perception and purchasing behavior. Park et al. (2021) revealed through text mining analysis that while technical specifications in professional reviews enhance credibility, excessive jargon creates comprehension barriers for non-expert readers. Chen and Wang (2022) experimentally established that user-generated reviews outperform professional evaluations in predicting long-term reliability, particularly for post-ownership vehicle conditions. The multimedia dimension has been examined by Lee et al. (2020), whose analysis of YouTube automotive content showed that dynamic driving footage generates 37% higher emotional engagement than static imagery, especially for experience-focused vehicle segments. Cross-cultural differences emerge prominently, as Brand et al. (2022) found that the Chinese perceive review credibility holistically, whereas Germans tend to categorize its antecedents for evaluating them separately.

Theoretical models. The Elaboration Likelihood Model (ELM) and Information Adoption Model (IAM) are the most commonly used theoretical models in review quality research (Aghakhani et al. 2021; Choi, 2020; Elwalda et al. 2022; Park et al. 2021). In ELM, information is transmitted and received through the peripheral and central routes, thus changing consumers' psychological attitudes and achieving the purpose of persuasion (Petty and Cacioppo 1984). In IAM, factors determining review quality are categorized into two perspectives: the information source perspective and the information content perspective (Sussman and Siegal, 2003).

(1) From the perspective of information sources, review depth, review emotion, star rating, information quantity, persuasiveness, and source credibility are essential factors in the peripheral route that researchers often focus on. For example, Chong et al. (2018) have studied the quality of online reviews using the dual-process theory and found that the quality of online reviews largely determines their effectiveness. Wu et al. (2021) employed review depth, content bias, and voting deviation to assess review quality by the two-factor theory. The analysis shows that reviews with less variation have a more significant impact on review quality. Biswas et al. (2021) identified essential factors affecting online review quality, including review title, review emotion, star rating, and time features, and found that ethnic culture has a strong moderating effect. Mousavizadeh et al. (2022) examined the factors that influence the popularity and quality of online reviews. The results show that longer online reviews and those with extreme star ratings are more popular. Finally, Zhuang et al. (2023) employed regression analysis to develop a quality framework comprising social, textual, and evaluation features as indicators for predicting high-quality answers.
(2) From the information content perspective, title content, readability, review validity, semantic features, and language style are key research factors for central routes. For example, Zhou et al. (2020) used review data from Amazon.com to model the effects of content similarity and sentiment consistency on review quality. The results indicate that the similarity of title content has a positive impact on review quality. Wang et al. (2025) explored the effects of review length, readability, validity, polarity, and other characteristics on review quality. They found systematic differences between voted and unvoted reviews. Through experiments, Zhan et al. (2023) found that reviewers described their actions rather than reactions to utilitarian products, while they expressed enjoyment rather than responses to hedonic products. Xiang et al. (2017) believe that language features, such as readability, and semantic features, including semantic relationships in online text, can be used to measure the quality of reviews. Liu et al. (2019) found that the degree to which the language style of online reviews aligns with the target customer's language style directly impacts the quality of online reviews.

Technical models. Several scholars have investigated online review quality using various technical models.

(1) From a sentiment analysis perspective, Zhu et al. (2020) examined the differences in information quality among online reviews based on the information richness theory, emotional polarity, and product type. The results indicate that diverse information richness and product types have varying impacts on information quality across different positive and negative emotional polarities. Wang et al. (2024) extracted the statistical and topic semantic features of online reviews for NEV and clarified the quality evaluation indicators for online reviews. The research results have demonstrated that the statistical features and semantic topic features of online reviews on NEVs can serve as evaluation indicators to construct a quality evaluation system. He and Wang (2022) focus on evaluating NEV through online reviews. A data-driven method is developed to extract the attributes of the NEV. Sentiment analysis is conducted to discriminate the sentiment orientation of each review for each alternative under each attribute.
(2) From the machine learning perspective, Malik (2020) extracted language features from online review texts and constructed a review quality prediction model. The results show that the mixed features of the samples provided the best prediction accuracy. Ding et al. (2023) developed a combined forecasting model comprising a backpropagation neural network (BPNN), a recurrent neural network (RNN), and a long short-term memory (LSTM) neural network. The study aims to improve the prediction accuracy of product sales by developing an online review-driven combination forecasting model. Wang et al. (2023a) introduced an enhanced hybrid model, combining Enhanced Representation through Knowledge Integration and a deep Convolutional Neural Network. Liu and Feng (2022) employed the BERT-wwm-ext model structure, data mining, and deep learning to investigate the selection of NEVs. Also, they analyzed the positioning of domestic and foreign NEV brands, as well as their brand development, from a complex network perspective. Wang et al. (2023b) extracted consumer demands and topics using the LDA model and used a trained Word2vec tool to extend the consumer demand topics.

Methodology

Evaluation of review quality. In various studies on review quality, elements such as review titles (Zhou et al. 2020), review pictures (Li et al. 2022), review sentiment (Aakash et al. 2024), review semantics (Kim et al. 2024), and ratings (Lei et al. 2023) are essential factors that determine review quality. Inspired by this, the study attempts to construct an evaluation system for review quality based on the Yale Attitude Change Model (YACM). The YACM was founded in 1953 by Carl Hovland (Riley et al. 1954), a psychologist at Yale University in the United States. This model identifies factors that affect the evaluation of the credibility of online reviews, such as the information media and the information content.

Dimension of information media. The media of information is the carrier of information related to products and services. Online reviews are transmitted between publishers and readers through texts and pictures.

(1) Length of review title
Consumers tend to spend a high amount of time searching for review content when they are interested in a product or service. Consumers would like to find worthwhile information with less search costs (Salehan and Kim 2016). The review title provides a concise summary of the review content, giving readers a general understanding and serving as a clue to quickly grasp the content. Online forums encourage consumers to provide detailed product information from different dimensions, and review contents are generally lengthy. The length of online review titles varies and is more distinguishable, which can be used as an indicator to measure the review quality. Therefore, the "Length of review title" is defined as the quality evaluation indicator of online reviews.
(2) Number of review pictures
Pictures included in online reviews are an effective supplement to text content, capable of bringing visual impact to potential consumers. The fact that consumers are willing to spend time taking pictures and posting them along with reviews indicates the emphasis on online reviews. The overall quality of online reviews will be improved due to the decorative effect of pictures and the

consumers' attention (Liu and Park 2015). Pictures not only attract more attention through decorative function but also enhance potential consumers' understanding and cognition of products through informational function (Anindya and Panagiotis 2011). Therefore, the "Number of review pictures" is defined as an evaluation indicator for the quality of online reviews.

Dimension of information content. The information content refers to all constituent elements with actual expressive significance in each online review. It can be divided into structured content and unstructured content, and to a large extent, determines the review quality. Structured content mainly includes various evaluation scores; unstructured content mainly includes review text, semantic content, review readability, review sentiment, etc.

(3) Sentiment of review title
Sentiment refers to the attitude, opinion, or emotion expressed toward an entity, topic, or event (Liu, 2012). Potential consumers can infer the content of a review from the title, preliminarily judge its value, and decide whether to continue browsing. Review titles present reviewers' opinions, feelings, and preferences towards products or services, leaving a first impression with concise information (Ascaniis and Gretzel 2012). Studies have shown that emotional language in review titles is more likely to attract attention than neutral reviews (Ismagilova et al. 2019). The more emotions reviewers express in the titles, the easier it is to attract readers' attention. Therefore, "Sentiment of review title" is defined as an evaluation indicator for the quality of online reviews.

(4) Review positive sentiment
Review sentiment is regarded as one of the critical factors affecting review quality (Racherla and Friske 2012). Positive sentiment expresses favorable, approving, or optimistic attitudes (Liu 2012). A well-crafted review with a lot of positive sentiment will reduce the search cost of potential consumers and ease their anxiety. When most consumers are satisfied with a product or service, potential consumers tend to have a convergent psychology (Yi et al. 2020). "Review positive sentiment" is defined as a quality evaluation indicator for online reviews.

(5) Review negative sentiment
Negative sentiment conveys unfavorable, critical, or pessimistic opinions (Liu 2012). Negative reviews, which account for a small proportion, sometimes have more evaluative value than positive reviews (Hu et al. 2009). Consumers tend to rely more on negative information to avoid risks (Lee et al. 2017; Kwok and Xie 2016). People are more inclined to post negative reviews to expose merchants' destructive behaviors, thereby reminding others to avoid repeating the same mistakes (Yan et al. 2018). Negative evaluations formed among potential consumer groups are considered to have a higher reference value regardless of whether they reduce the possibility of consumption. In view of the above theories, "Review negative sentiment" is defined as a quality evaluation indicator for online reviews.

(6) Review semantics
Semantics is the study of meaning in language, encompassing word meanings (lexical semantics) and their combination in sentences (Saeed 2016). Studies by Kim et al. have shown that the amount of information semantics directly affects consumers' perception of products (Kim et al. 2017). If an online review includes semantic information that helps understand product quality, the review is considered to be more diagnostic of product performance (Filieri 2016). Online reviews rich in semantic content often have higher information relevance and greater information credibility, bringing consumers stable confidence. Based on the research above, this paper defines "Review semantics" as a quality evaluation indicator for online reviews.

(7) Review readability
Readability measures how easily written text can be understood by readers, influenced by factors like sentence length and vocabulary complexity (DuBay 2004). Suppose the semantic content represents the amount of information objectively contained in a review. In that case, readability refers to the amount of information that readers can obtain from the review through their subjective cognition (Liang et al. 2018). The readability of online reviews is directly related to the clarity of expression in the reviews. The more coherent the sentences are, the easier it is to increase the reading speed of potential consumers, enhance the reading impression, and thus improve the quality of online reviews (Ahmad and Laroche 2016). Readable reviews are more likely to be recognized as high-quality reviews and more likely to be disseminated (Liu and Park 2015). In view of the above, "Review readability" is defined as a quality evaluation indicator for online reviews.

(8) Rating consistency
Rating consistency refers to the stability of respondents' evaluations across repeated measurements or similar items (Krosnick and Berent 1993). If the consistency of product ratings is very high, it indicates that the reviewers' responses to the product attribute dimensions are consistent with those of other consumers. Correspondingly, reviews with low consistency are likely to cause cognitive dissonance among potential consumers (Choi and Leon 2020). Studies have found that review readers tend to consider opinions inconsistent with those of other reviewers as less credible and less helpful (Simon et al. 2015). Chua & Banerjee (2016) used attribution theory to explain that reviews with high rating consistency may be attributed to the experience of the product, while those with low rating consistency may be attributed to personal traits. Therefore, if consumers consistently rate a particular product attribute dimension highly, it means that their reviews are more useful and of higher quality. In view of the above, "Rating consistency" is defined as a quality evaluation indicator for online reviews.

(9) Performance score
Online forums allow reviewers to rate product performance. The score reflects the reviewers' level of affirmation towards the products' performance and also influences the review readers' judgment on the quality of online reviews. In view of the above, "Performance score" is defined as a quality evaluation indicator for online reviews.

The quality evaluation system for online reviews is shown in Table 1.

The weights and assignments of quality indicators. This paper first applies the entropy weight method to determine the weights of online review quality evaluation indicators. Then, the indicator weights are used to assign values to online reviews. The specific steps are shown below.

(1) Determine the evaluation object and construct a level matrix of online review data.

$$X = \left[x_{ij}\right] = \left[\begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix}\right] (i = 1, 2, \ldots, n; j = 1, 2, \ldots, m) \tag{1}$$

Where x_{ij} represents the value of the j-th evaluation indicator for the i-th online review.

(2) Data standardization processing. To eliminate the impact of dimensional differences between data on estimation results, the min-max normalization method is adopted for processing. The specific operation is that when the evaluation indicator value is positive,

$$y_{ij} = \frac{x_{ij} - \min(x_j)}{\max\left(x_j\right) - \min(x_j)} (i = 1, 2, \ldots, n; j = 1, 2, \ldots, m) \tag{2}$$

When the evaluation indicator value is negative,

$$y_{ij} = \frac{\max\left(x_j\right) - x_{ij}}{\max\left(x_j\right) - \min(x_j)} (i = 1, 2, \ldots, n; j = 1, 2, \ldots, m) \tag{3}$$

Among them, $\max\left(x_j\right)$ is the maximum value of the evaluation indicator, and $\min\left(x_j\right)$ is the minimum value of the evaluation indicator. After normalization, the matrix is generated as

$$\mathrm{Y} = \left[y_{ij}\right] = \left[\begin{pmatrix} y_{11} & \cdots & y_{1m} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{nm} \end{pmatrix}\right] (i = 1, 2, \ldots, n; j = 1, 2, \ldots, m) \tag{4}$$

(3) Using the entropy weight method to assign weights to evaluation indicators, calculate the probability matrix p, and use formula (5) to determine the value of each element p_{ij} in the matrix:

$$p_{ij} = \frac{y_{ij}}{\sum_{i=1}^{n} y_{ij}} (i = 1, 2, \ldots, n; j = 1, 2, \ldots, m) \tag{5}$$

Calculate the information entropy of each evaluation indicator. For the j-th indicator, the formula for its information entropy is:

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^{n} p_{ij} \ln\left(p_{ij}\right) (i = 1, 2, \ldots, n; j = 1, 2, \ldots, m) \tag{6}$$

The larger *the* e_j, the more significant the information entropy of the j-th indicator, indicating that it carries less information.

The information utility value d_j, is defined as

$$d_j = 1 - e_j \tag{7}$$

Normalize the information utility value to obtain the entropy weight of each quality evaluation indicator:

$$w_j = \frac{d_j}{\sum_{j=1}^{m} d_j} (j = 1, 2, \ldots, m) \tag{8}$$

Table 2 presents the information entropy, weight, and weight ranking of review quality evaluation indicators derived using the entropy weight method.

Combined with the indicator weights, the TOPSIS method is applied to assign a value to the review quality. First, a weighted matrix is established, and the fuzzy matter-element matrix is obtained through the correlation analysis.

$$\mathrm{Z} = \left[z_{ij}\right] = \left[\begin{pmatrix} z_{11} & \cdots & z_{1m} \\ \vdots & \ddots & \vdots \\ z_{n1} & \cdots & z_{nm} \end{pmatrix}\right] (i = 1, 2, \ldots, n; j = 1, 2, \ldots, m) \tag{9}$$

Among them, $z_{ij} = y_{ij} \times w_j$ (where w_j represents the indicator weight obtained by the entropy weight method).

The Positive Ideal Solution (PIS) and Negative Ideal Solution (NIS) of the model are determined according to Formulas 10 and 11:

$$Z^{+} = \left(z_1^{+}, z_2^{+} \ldots \ldots z_n^{+}\right) \tag{10}$$

$$Z^{-} = \left(z_1^{-}, z_2^{-} \ldots \ldots z_n^{-}\right) \tag{11}$$

Among them, $z_i^{+} = max\left\{z_{1j}, z_{2j,} \ldots \ldots z_{nj,}\right\}$, $z_i^{-} = min\left\{z_{1j}, z_{2j,} \ldots \ldots z_{nj}\right\}$

Subsequently, the Distance from PIS and Distance from NIS are calculated using Formulas 12 and 13.

$$D_i^{+} = \sqrt{\sum_j (z_{ij} - z_j^{+})^2} \tag{12}$$

$$D_i^{-} = \sqrt{\sum_j (z_{ij} - z_j^{-})^2} \tag{13}$$

Table 1 Quality evaluation system for online reviews.

Indicator dimension	Indicator	Indicator symbol
Dimension of information media	Length of review title	Q1
	Number of review pictures	Q2
Dimension of information content	Sentiment of review title	Q3
	Review positive sentiment	Q4
	Review negative sentiment	Q5
	Review semantics	Q6
	Review readability	Q7
	Rating consistency	Q8
	Performance score	Q9

Table 2 Indicator weights of the review quality evaluation system.

Indicators	Information entropy	Information utility value	Indicator symbol	Indicator weight	Rank
Length of review title	0.988	0.012	Q1	0.045	6
Number of review pictures	0.900	0.100	Q2	0.363	1
Sentiment of review title	0.990	0.010	Q3	0.034	8
Review positive sentiment	0.988	0.012	Q4	0.044	7
Review negative sentiment	0.952	0.048	Q5	0.181	2
Review semantics	0.992	0.008	Q6	0.029	9
Review readability	0.969	0.031	Q7	0.117	4
Rating consistency	0.963	0.037	Q8	0.139	3
Performance score	0.999	0.002	Q9	0.047	5

Finally, the relative closeness is computed via Formula 14.

$$C_i = \frac{D_i^-}{D_i^+ + D_i^-} \tag{14}$$

Measurement of quality indicators. The quality evaluation indicators for online reviews are measured as follows. Firstly, indicators such as "Length of review title" and "Number of review pictures" are extracted from the webpages. And "Performance score" is a value derived from ratings. Next, the snowNLP package (package version: v0.12.3) in the Python environment is applied to analyze the sentiment of the comments, and the indicators such as "Sentiment of review title," "Review positive sentiment," "Review negative sentiment," "Review semantics," and "Review readability" are calculated respectively.

SnowNLP is a Python library designed for processing and analyzing Chinese text data. Its functionality encompasses various natural language processing (NLP) tasks, including part-of-speech tagging, text segmentation, and keyword extraction. A core component of its analytical capability is sentiment analysis. The model calculates a sentiment polarity score for a given input text. This score is normalized to a continuous floating-point value within the range of 0 to 1. The polarity of this scale is defined such that a value closer to 1 indicates a more positive sentiment, whereas a value closer to 0 signifies a more negative sentiment.

Readability is a self-defined indicator based on features such as sentence length, word length, nouns, and neutral words extracted by SnowNLP. Among these features, average sentence length, average word length, and the proportion of nouns are negatively correlated with readability, while neutral words are positively correlated with readability. In line with the core principle of reverse processing for negatively correlated features and direct incorporation for positively correlated features, the readability score formula, normalized to the range of 0 to 1, is presented as follows:

$$R = w_1 \times \frac{1}{SL_{avg}} + w_2 \times \frac{1}{WL_{avg}} + w_3 \times \left(1 - N_r\right) + w_4 \times M_r \tag{15}$$

where R denotes the readability score (0 to 1), SL_{avg} represents the average sentence length, WL_{avg} stands for the average word length, N_r indicates the proportion of nouns (0 to 1), and M_r refers to the proportion of neutral words (0 to 1). w_1 is the weight of average sentence length (0.3), w_2 is the weight of average word length (0.25), w_3 is the weight of noun proportion (0.2), and w_4 is the weight of neutral word proportion (0.25).

As for the measurement of "Rating consistency," we first defined "Review discrete value," that is, the difference between the "comprehensive score" of each review and the average of all "comprehensive scores." Since the degree of dispersion of reviews and the degree of rating consistency are opposite concepts, the lower the degree of dispersion, the higher the degree of consistency. To maintain a positive proportional relationship, the higher the degree of consistency, the higher the rating consistency score should be. Therefore, the reciprocal value of "Review discrete value" was taken to obtain the variable of "Rating consistency."

Model of influencing factors of review quality. To study the influencing factors of review quality, it is necessary to consider the impact of review elements on review quality from the perspective of information content and pay attention to the impact of peripheral elements on review quality from the perspective of information sources. Relevant theories suggest that the time interval between consumers' post-purchase experience and their reviews puts consumers at a higher level of knowledge construction; the identity of reviewers affects consumers' trust and the generation of purchase intention; and review professionalism and reviewers' experience have a positive impact on the usefulness of online reviews (Huang and Huo 2022). Relevant studies show that consumers generally do not recognize that the technology of NEV is currently mature and believe that there are shortcomings in driving range and battery life; needs such as picking up and dropping off children and commuting to school are the main reasons for users to purchase NEV (Huang et al. 2016).

Based on the above theories and research, this paper analyzes feature mining results from NEV reviews, identifying key factors influencing review quality. These factors belong to information content features (e.g., driving mileage, time interval, number of car-buying purposes) and information source features (e.g., certified users, reviewer's expertise). A multivariate regression model is constructed to examine the impact of these features on review quality.

Influencing factors and hypotheses.

(1) Driving mileage
Driving mileage is vehicle-related information that online forums allow consumers to fill in when posting reviews, which helps readers understand consumers' previous usage of the vehicles. It is generally believed that the owner's understanding and feelings about the vehicles will deepen with the increase in vehicle experience time, and the description of the review content is more objective and profound. However, consumers' focus on new energy vehicles differs from fuel vehicles, such as evaluating battery performance, which is essential in consumers' overall car assessment. The survey by Li et al. (2020b) shows that respondents who have not used new energy vehicles show higher concerns about battery performance and battery safety. Surveys about online reviews show that consumers are less likely to recognize that the current technology of new energy vehicles has matured, which reflects the shortcomings of new energy vehicles in terms of range, power battery life, and stability (Li and Guo 2022). The issue of battery performance degradation has consistently impacted consumers' purchasing intentions and perceptions of new energy vehicles. Some consumers believe that the vehicle's speed, frequent acceleration, and braking will impact battery performance. Therefore, as the use time of new energy vehicles continues to accumulate, the owner's attention to new energy attributes, such as battery performance, will increase, and the resulting complaints will also be reflected in the emotions and content of online reviews. The new energy vehicle industry is still rapidly developing, and technical issues may impact consumers' experiences. Given the above, an increase in driving mileage may affect the quality of online reviews, and therefore, the following hypothesis is proposed:
H1: The Driving mileage negatively impacts the review quality.

(2) Time interval
The period from purchasing a product or participating in a service to writing a review, whether long or short, is referred to as the time interval. The temporal construal theory states that distant future situations are construed higher (i.e., using more abstract and central features) than near future situations. Accordingly, the theory suggests that the value associated with the high-level construal is enhanced over delay (Liberman and Trope 1998). The time interval may affect potential consumers' understanding of the online review content, influencing their

assessment of review quality. Observations indicate that the Time interval between product experience and posting reviews puts consumers at a higher level of construction, which helps to reduce the negativity in online reviews, making the reviews more specific and valuable (Huang et al. 2016). Combining the above analysis, a longer review time interval can effectively improve the knowledge construction level of review readers, so the following hypothesis is proposed:

H2: The Time interval positively impacts the review quality.

(3) Number of car-buying purposes

Different consumers have different car-buying purposes. Some models of consumers' car-buying purposes are dispersed, while others are concentrated. A survey-based study indicates that rigid travel demands—specifically, commuting to and from school as well as transporting children to and from school—constitute the primary motivations for users to purchase new energy vehicles. Specifically, over 70% of the surveyed users cited commuting to and from school as a key factor, while 33% identified transporting children to and from school as a relevant consideration (Wang et al. 2023c). Purposes such as "business trips," "commuting," "picking up and dropping off children," and "self-driving tours" objectively reflect consumers' positioning of car function requirements.

The purpose of buying a car affects the grade and type of car chosen by potential consumers to some extent. If the consumer's goal of buying a car is single, the discussion angle around the vehicle's use in the review content may be more concentrated. On the other hand, consumers purchasing a car for multiple purposes may discuss and compare various vehicle uses comprehensively in the review content. The consumer's description of the intention of using the vehicle enriches the information in online reviews. It enhances the richness of the information, improving the overall quality of the review content. Given the above analysis, this study proposes the following hypothesis:

H3: The Number of car-buying purposes positively impacts the review quality.

(4) Certified users

Users of online car forums get certified by posting car model information, owner data, and reviews. With the completion of personal information and the increase in page views, the level of certified users is expected to continue improving. From a motivational perspective, one of the reasons certified users strive to elevate their class is to achieve self-improvement and establish an expert image by posting reviews. According to IAM (Petty and Cacioppo 1984), although the peripheral route of information transmission and reception is irrelevant to the content of the information itself, it requires the information recipient to use simple heuristic cues to analyze the peripheral environmental factors of the information to achieve an attitude change, i.e., to assess the credibility of the information at a lower cognitive cost. The reliability of the reviewer is one of the key factors that review readers use to assess the credibility of the information. Qiu & Zhang (2024) found that the identity of the reviewers affects the trust and purchase intention of potential consumers. Therefore, whether the reviewer is a certified user can be an essential external supporting condition that potential consumers consider. Certified users may support external conditions, which become an external factor for review readers to consider. Given the theory mentioned above, this study suggests that if the user who posts an online review is certified, it may impact the overall quality of online reviews. Therefore, the following hypothesis is proposed:

H4: The Certified users positively impact the review quality.

(5) Reviewer's expertise

Expertise was measured by whether a reviewer had a platform ranking. This is a direct indication provided by the platform that the reviewer has contributed a large number of helpful reviews as voted by the community (Mudambi and Schuff 2010). Professional reviews can reduce the costs associated with information search for potential consumers, lower the motivation to verify the authenticity of information sources, and help accurately position product quality. Additionally, potential consumers seek advice from professional information sources when faced with complex decision-making dilemmas. For instance, Wang & Chen (2023) investigated the role of reviewer expertise in the perceived usefulness of online reviews, demonstrating that reviewers' professional background and experience significantly enhance the perceived credibility and practical value of their reviews. Through data analysis of online review websites, Zhang et al. (2016) found, after analyzing review data on qunar.com, that hotels with more expert reviews also receive higher scores. Furthermore, Huang et al. (2016) explored the impact of reviewers' characteristics on the usefulness of hotel online reviews. The results showed that the reviewer's expertise and experience have a positive impact on the usefulness of online hotel reviews, and the influence of the reviewer's expertise is even more significant. Based on the above analysis, this study proposes the following hypothesis:

H5: The Reviewer's expertise positively impacts the review quality.

Measurement of influencing factors. According to the hypothesis of the influencing factor model, this study identifies "Driving mileage," "Time interval," "Number of car-buying purposes," "Certified users," and "Reviewer's expertise" as five independent variables. "Driving mileage" refers to the kilometers the vehicle has travelled when consumers post reviews, which can be directly extracted from the webpages. "Time interval" refers to the number of days between when a consumer posts their review and the vehicle's purchase time, which can be directly extracted from the webpages. Regarding the acquisition of the "Number of car-buying purposes," the websites encourage reviewers to describe their purchase purpose in the form of short phrases, such as "commuting to work," "picking up and dropping off children," "long-distance driving," etc. The "Number of car-buying purposes" corresponding to each online review can be obtained by counting the number of purchase purposes. "Certified users" can be directly extracted from the reviewers' information pages.

"Reviewer's expertise" is assigned as follows: If an online review meets specific criteria, the review will be tagged with "Recommended." Meanwhile, according to the content of the reviews, "Recommended" reviews are divided into three levels of word of mouth (WOM): "Full-Level WOM," "Essence WOM," and "Recommended WOM." The grading criteria for 'Recommended' reviews are as follows: A 'Full-Level WOM' is awarded if the review content is detailed, has original insights, includes vehicle experience of more than 1,500 words, and is accompanied by vehicle photos. An "Essence WOM" is awarded if the review content is insightful, vividly written, persuasive, includes a vehicle experience of more than 1,000 words, and is accompanied by vehicle photos. A "Recommended WOM" is awarded if the review content is objective and fair, of reference value, includes vehicle experience of more than 500 words, and is accompanied by vehicle photos. The "Reviewer's expertise" is scored according

Table 3 Influencing factor variables and measurements.

Variable category	Variable names	Variable symbols	Variable measurements
Dependent variable	Review quality	Quality	Calculated through the quality evaluation system
Independent variable	Driving mileage	Mileage	Page capture
	Time interval	Interval	The difference between the review posting time and the car purchase time
	Number of car-buying purposes	Purpose	Statistics based on the content of online reviews
	Certified users	Certification	Certified users are assigned a value of 1, and non-certified users are assigned a value of 0
	Reviewer's expertise	Expertise	Full-Level WOM is scored 3, "Essence WOM" is scored 2, "Recommended WOM" is scored 1, and ordinary online reviews without a recommendation level are scored 0

Table 4 Correlation matrix and reliability measures.

	Review quality	Log(Driving mileage)	Log(Time interval)	Number of car-buying purposes	Certified users	Reviewer's expertise
Review quality	1.000	−0.007**	0.010**	0.246**	0.003**	0.250**
Log(Driving mileage)	−0.007**	1.000	0.543**	0.040**	−0.006**	0.022**
Log(Time interval)	0.010**	0.543**	1.000	0.033**	−0.005**	0.021**
Number of car-buying purposes	0.246**	0.040**	0.033**	1.000	−0.001**	0.261**
Certified users	0.003**	−0.006**	−0.005**	−0.001**	1.000	0.035*
Reviewer's expertise	0.250**	0.022**	0.021**	0.261**	0.035*	1.000

**Indicates significance at the 5% level.

to the recommended level of each online review: "Full-Level WOM" is scored 3, "Essence WOM" is scored 2, "Recommended WOM" is scored 1, and ordinary online reviews without a recommended level are scored 0.

Table 3 shows the descriptions and measurement methods of the dependent and independent variables.

Model and data. This study categorizes variables such as "Driving mileage," "Time interval," and "Number of car-buying purposes" into an information content perspective of IAM. It categorizes "Certified users" and "Reviewer's expertise" into an information source perspective of IAM. A regression equation is established, as shown in formula (16):

$$\begin{aligned} Quality = \beta_0 + \beta_1 Mileage + \beta_2 Interval + \beta_3 Purpose \\ + \beta_4 Certification + \beta_5 Expertise + \varepsilon \end{aligned} \tag{16}$$

Where β_0 is the constant, $\beta_1 \sim \beta_5$ are the parameters to be estimated, and ε is the error term.

Autohome.com.cn is the world's largest automotive website in terms of traffic, offering car price inquiries, test-drive evaluations, model recommendations, online inquiries, and other features. It is the fastest and most comprehensive Chinese automotive website. This article selects the review data from the "New Energy Vehicle" section on the website as the research object, collecting 6,446 reviews for "below CNY100,000 (USD 13,930)" models, 10,643 reviews for "CNY 100,000-200,000 (USD 13,930-27,860)" models, 11,354 reviews for "CNY 200,000-300,000 (USD 27,860-41,790)" models, and 6,149 reviews for "over CNY 300,000(USD 41,790)" models. Furthermore, after stop-word removal and word segmentation processing, the four groups of corpora were generated, ready for subsequent analysis.

Results

Regression analysis. Given that the values of the two independent variables, "Driving mileage" and "Time interval," are significant, they are subjected to a logarithmic transformation. The logarithmic transformation of independent variables originates from the data preprocessing principles in statistics and econometrics. Its core objectives are to satisfy model assumptions, optimize estimation effects, or enhance the interpretability of results. Logarithmic transformation does not alter the economic significance or logical relationship between variables but only improves computational stability by compressing the numerical range. It can eliminate the interference of measurement units on correlation calculations, making the relationships between variables more objective (Box and Cox 1964). It reduces the absolute value of the data without altering the nature and relationships of the data, making the data calculation smoother. However, it also weakens the issues of collinearity and heteroscedasticity in the model (Wooldridge 2019).

Initially, a correlation analysis of the variables is conducted, and the correlation matrix and reliability measures are shown in Table 4. There is a specific correlation between the variables, suggesting that these variables can be used to explain the quality of online reviews. The correlation coefficients between the independent variables are around 0.5 or less, initially meaning that there are no severe multicollinearity problems, and overall, the sample is suitable for subsequent research.

It has been verified that the variance inflation factor (VIF) values of all variables are less than 5, which further confirms the absence of an obvious multicollinearity problem and ensures the reliability of the subsequent regression analysis results. The regression results with confidence intervals are presented in Table 5, where the values in the table represent the standardized coefficients of the related explanatory variables after standardization and their corresponding significance levels.

Table 6 presents the results of the validity test of the regression model. As can be seen from the table, R^2 and adjusted R^2 in the model are 0.630, with $p < 0.01$, indicating that the model fits well in the case of a large sample size.

Table 7 presents the results of the Analysis of Variance (ANOVA) conducted for the proposed model. The F-value, which serves as a comprehensive metric for evaluating the

Table 5 Regression results with confidence intervals.

Variable type	Variable names	Standardized β and significance level
Independent variable	Log(Driving mileage)	−0.023***
	Log(Time interval)	0.160**
	Number of car-buying purposes	0.084***
	Certified users	0.003
	Reviewer's expertise	0.169***

** indicates significance at the 5% level, and *** indicates significance at the 1% level.

Table 6 Results of the model validity test.

	R	R^2	Adjusted R^2	Significance level
Model	0.652	0.630	0.630	0.000

Table 7 ANOVA analysis results.

		Degree of freedom	Mean square	F	Significance level
Model	Regression model	5	0.493	468.357	0.000
	Residuals	33183	0.001		
	Total	33183			

model's overall validity, is statistically significant. This indicates that the independent variables, taken together, exert a substantial impact on the dependent variable.

Findings and discussions. Based on the results of the regression analysis, this study conducts discussions from the perspectives of information content and information source, respectively.

Perspective of information content.

(1) "Time interval" significantly influences the review quality ($\beta = 0.160; p < 0.05$), supporting Hypothesis 2 (H2). "Number of car-buying purposes" significantly influences the review quality ($\beta = 0.084$; $p < 0.01$), supporting Hypothesis 3 (H3). According to temporal construal theory, consumers who post reviews immediately after experiencing the products tend to focus more on specific details (low-level construal). In contrast, those who delay posting reviews are more inclined to provide overall evaluations (high-level construal). Values associated with high-level construal increase with time delay, as readers perceive delayed reviews as more objective (Chen and Lurie 2013). The results indicate that the longer the interval between a consumer's feedbacks about a car product, the more pronounced the perceived effect of review quality for the reader. This finding is consistent with the perspective of temporal construal theory (Jürgen et al. 2023). Multi-functional reviews are more likely to be recognized because they cover a wider range of usage scenarios (such as charging convenience and space requirements), thereby meeting the needs of different consumers. Compared with the content features of online reviews, a larger "Number of car-buying purposes" may lead potential consumers to associate the car with more functions and uses, indirectly helping review readers better understand the review content and effectively promoting the quality of the content.

(2) "Driving mileage" significantly influences the review quality $\left(\beta = -0.023; p < 0.01\right)$, supporting Hypothesis 1 (H1). The negative regression coefficient suggests a negative correlation between the vehicle's mileage and the quality of online reviews. In the IAM, driving mileage is a content element of information. Combined with driving mileage, readers may change their attitudes or behavior patterns after reading the content of online reviews. By analyzing the review content, it is found that consumers with more kilometers tend to be more concerned about the limited range and the lack of charging stations. Battery lifespan and charging speed are the technical attributes that consumers are most concerned about (Pawel et al. 2023). These concerns may affect usage satisfaction and trigger negative emotions, and the negative reviews generated in such contexts may reduce the review quality. Therefore, higher driving mileage implies a potentially lower review quality, which makes it less likely to meet the psychological expectations of potential consumers about the product. Increases in both the time interval and the number of purposes exert a positive effect on review quality. In contrast to this positive impact, however, a higher driving mileage has an effect on review quality in the opposite direction.

Perspective of information source.

(1) "Certified users" do not significantly influence the review quality ($\beta = 0.003; p > 0.1$). Thus, Hypothesis 4 (H4) is not supported, indicating that whether a website user is certified does not correlate significantly with the review quality. The identity of an authenticated user has no substantial impact on review quality. Even with a large sample size, this variable still lacks a practical contribution to the model. As previously mentioned, many studies focusing on usefulness emphasize subjective feelings, such as consumers' emotional identification, while this paper focuses on the objective evaluation of review content. Although some studies have shown that the certification status of reviewers may affect consumers' purchase intention (Zhang and Wang 2021), the identity of reviewers may not have such a direct impact on objective review quality. Online review websites encourage users to post more reviews. Still, some websites lack robust regulatory mechanisms and fail to differentiate restrictions between certified and non-certified users posting reviews. Additionally, the presence of anonymous and spam posts makes it challenging for consumers to assess the reliability and credibility of online reviews based on user cues. Moreover, it cannot be ruled out that certified users might post some low-quality advertisements or even false posts for commercial benefits, leading to a generally low review quality.

(2) "Reviewer's expertise" significantly influences the review quality $\left(\beta = 0.169; p < 0.01\right)$, supporting Hypothesis 5 (H5). For every one standard deviation increase in the reviewer's expertise (such as upgrading from "ordinary review" to "recommended WOM"), review quality improves by 0.169 standard deviations. Professional labels (e.g., platform-certified "technical evaluation") enhance content persuasiveness through credibility endorsement. Especially in high-tech threshold fields like NEV, consumers rely more on expert opinions to reduce decision-making risks. Without relying on others for product information, reviewers need to invest some effort to write "recommended" level reviews. Whether praising the product or posting a negative review, professional reviews tend to be more objective, specific, and more readily accepted. Consumers are more

inclined to seek and follow advice from professional information sources when faced with decision-making dilemmas than ordinary online reviews.

In addition to the discussions from the two perspectives outlined above, this study also identifies that, in the evaluation system for online review quality, "Number of review pictures" carries significant weight. This high weight not only reflects the critical role that review pictures play in enhancing review quality, but also indicates the indicator's high variability and strong discriminative ability across reviews, which is the statistical basis of entropy weighting. As a core visual medium, pictures contribute to review quality through multidimensional mechanisms of information transmission and cognitive influence. From the perspective of information supplementation, pictures overcome the limitations of textual expression by directly presenting product details, usage effects, and contextual features, thereby effectively reducing information asymmetry in consumer decision-making. In terms of credibility construction, pictures serve as "visual evidence," enhancing the authenticity and objectivity of reviews and strengthening consumers' trust in the content. From a cognitive processing standpoint, the consistency between pictures and texts reduces consumers' cognitive load. As found in Ceylan et al.'s (2023) research, photos increase the helpfulness of a review. More importantly, though, greater similarity between photos and texts heightens review helpfulness more.

All hypothesis verification results are shown in Table 8.

Conclusion

Data analysis and mining of user-generated content in online environments is a hot issue shared by the industry and academia. Product information, service information, user experience, usage habits, and other relevant information in online reviews are of great value to potential consumers and product manufacturers. The findings highlighted theoretical contributions to the literature and have practical implications for managers and customers.

Theoretical contribution. First, this study pioneered the construction of a quality evaluation system for online reviews of new energy vehicles. The research on online reviews has thus been expanded from the original qualitative research of "good or bad" to the quantitative analysis of "how good or how bad." Unlike traditional questionnaires and field interviews used in previous studies, this paper uses online reviews as the primary data source. Based on the content of numerous reviews, this paper utilizes a model to determine the factors that influence the quality of online reviews. It helps validate the possibility of extracting review quality influences from massive online reviews and turning them into valuable business intelligence.

Second, our research provides an exciting contrast to the findings that "Certified users" will have a higher quality evaluation than uncertified users (Wang et al. 2023c; Petty and Cacioppo 1984). In contrast, this paper found that whether a website user is certified does not correlate significantly with the quality of online reviews. Some websites lack robust regulatory mechanisms and fail to differentiate restrictions between certified and non-certified users posting reviews. By adding the "Certified users" component as a precursor, the paper extends the previous literature, which has only elaborated on the impact of certified users on quality evaluation (Zhang and Wang, 2021).

Table 8 Hypothesis verification results.

Hypothesis	Results
H1: Driving mileage negatively impacts the review quality.	Support
H2: Time interval positively impacts the review quality.	Support
H3: The Number of car-buying purposes positively impacts the review quality.	Support
H4: Certified users positively impact the review quality.	Not support
H5: Reviewer's expertise positively impacts the review quality.	Support

Third, taking the quality of online reviews as the core concept, this paper expands the application scope of theoretical models such as the IAM and the ELM in the new scenario of automotive reviews, and provides a simple criterion for different consumer groups to perceive the value of online reviews. The methodology proposed in this study is a valuable attempt to conduct product research based on online reviews, and it also lays a solid foundation for other types of research based on online reviews (e.g., consumer preference analysis, product performance analysis, and product market analysis).

Managerial implications. First, as vehicle mileage increases, the quality of online reviews tends to decrease. Potential consumers who read online reviews tend to focus on evaluation information corresponding to fewer driving miles. They hope consumers who have already purchased cars can quickly provide valuable reference opinions. If the car product has been in use for a long time, it may indicate that less information is available for reference. Online review sites can improve the rules for posting reviews accordingly, such as setting up information filtering mechanisms and prioritizing the display of evaluation information with fewer miles driven.

Second, the period of online reviews positively impacts readers' information quality perception. The longer the interval between online reviews, the higher the quality the readers perceive. Although the analysis results show that the quality of online reviews improves as the consumer's review time lengthens, automobiles are consumer goods with a relatively fast replacement cycle. New energy vehicles are developing and updating faster than traditional fuel vehicles. Therefore, the content of online reviews needs to keep pace with product updates to resonate more easily with readers; otherwise, reviewers may lose valuable references due to memory loss.

Third, the Number of car-buying purposes objectively reflects consumers' positioning of car function requirements. Online forums can categorize online review content by car-buying purposes, guiding potential consumers with similar car usage requirements to browse corresponding grouped online reviews. Car manufacturers can also conduct targeted marketing activities based on consumers' positioning of car function requirements.

Fourth, most automobile review sites display reviews in chronological order, from the most recent to the oldest. While such a display mechanism ensures that the latest content is presented first, new reviews do not necessarily mean good reviews. Due to the high time cost of retrieval, most viewers only skim through some of the review content before making a decision. Therefore, through the quality evaluation mechanism, high-quality reviews are extracted from many reviews and prioritized to be pushed to viewers, which helps improve the efficiency of information utilization on the website. High-quality review content is the core resource of online review sites. Online review sites can combine reading habits and topic-mining results to filter out high-quality review content and prioritize and recommend it to review readers who need it.

Fifth, automobile companies must promptly address consumers' feedback on the use and experience of various automobile products, and identify the statistical patterns of online reviews and the semantic features of topics. Automotive companies need to establish a feedback mechanism with review websites to comprehensively analyze the quality of online reviews and their influencing factors from different dimensions. According to the review scores of different regions and models, they should flexibly adjust their production plans and marketing strategies. Automotive companies also need to take online reviews as a guide to improve product performance and service quality promptly, and ultimately realize increased sales volume and market share.

Limitations and future research. NEVs have a relatively short development history. Some models have been on the market for merely a year or even just half a year, which means the review data available for this study is limited. Such limited review data may compromise the overall quality of the regression model, and whether the model can be applied to other products still requires further verification. Additionally, the review data are mainly collected from mainstream domestic automotive vertical media platforms in China, which may diminish the promotional value of the research findings in countries and regions outside China.

Future research could focus on studies of influencing factors models in other industries and compare review quality evaluation models to discover similarities and differences. Future research could also segment consumer groups and focus on analyzing review quality evaluation models for specific consumer groups.

Data availability

Supplementary information. The online version contains supplementary material available at https://doi.org/10.6084/m9.figshare.29166875.

Received: 4 April 2025; Accepted: 25 February 2026;
Published online: 13 March 2026

References

Aakash S, Ali G, David L (2024) Looking for negativity! How review sentiment and review subjectivity form review helpfulness across cultures. J Consum Behav 23(3):1193–1207
Aghakhani N, Oh O, Gregg D, Karimi J (2021) Online review consistency matters: An elaboration likelihood model perspective. Inf Syst Front 23:1287–1301. https://doi.org/10.1007/s10796-020-10030-7
Ahmad N, Laroche M (2016) How do expressed emotions affect the helpfulness of a product review? Evidence from reviews using Latent Semantic Analysis. Int J Electron Commer 20(1):76–111
Anindya G, Panagiotis I (2011) Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. IEEE Trans Knowl Data Eng 23(10):1498–1512
Ascaniis S, Gretzel U (2012) What's in a travel review title? Information and communication technologies in tourism 2012, Helsingborg, Sweden, 494–504
Barbara K, Dale G, Gordon D (1997) Can humans detect errors in data? Impact of base rates, incentives, and goals. Mis Q 21:169–194
Biswas B, Sengupta P, Ganguly B (2021) Your reviews or mine? Exploring the determinants of perceived helpfulness of online reviews: A cross-cultural study. Electron Mark 32:1083–1102
Box G, Cox D (1964) An analysis of transformations. J R Stat Soc Ser B (Methodol) 26(2):211–252
Brand B, Kopplin C, Rausch T (2022) Cultural differences in processing online customer reviews: Holistic versus analytic thinkers. Electron Mark 32:1039–1060
Ceylan G, Diehl K, Proserpio D (2023) Words meet photos: When and why visual content increases review helpfulness. J Mark Res 61(4):002224372311697
Chen L, Wang Y (2022) Professional vs. user-generated auto reviews: A comparative study on information credibility and purchase intention. Int J Inf Manag 65:102502. https://doi.org/10.1016/j.ijinfomgt.2022.102502
Chen X, Li Y, Zhang H (2023) The impact of subsidy retreat on China's new energy vehicle industry: A market-driven perspective. J Clean Prod 385:135678
Chen Y, Lurie N (2013) Temporal contiguity and negativity bias in the impact of online word of mouth. J Mark Res 50(4):463–476
Choi H, Leon S (2020) An empirical investigation of online review helpfulness: A big data perspective. Decis Support Syst 139:113403
Choi S (2020) Antecedents of YouTube UGC tourism information adoption: An extended approach to information adoption model. J Tour Manag Res 24:681–708
Chong A, Khong K, Ma T, McCabe S (2018) Analyzing key influences of tourists' acceptance of online reviews in travel decisions. Internet Res 28:564–586
Chua A, Banerjee S (2016) Helpfulness of user-generated reviews as a function of review sentiment, product type, and information quality. Computers Hum Behav 54:547–554
Ding Y, Wu P, Zhao J, Zhou L (2023) Forecasting product sales using text mining: A case study in new energy vehicle. Electronic Commerce Research. https://doi.org/10.1007/s10660-023-09701-9
DuBay W (2004) The principles of readability. https://files.eric.ed.gov/fulltext/ED490073.pdf
Elwalda A, Erkan I, Rahman M (2022) Understanding mobile users' information adoption behavior: An extension of the information adoption model. J Enterp Inf Manag 35:1789–1811
Filieri R (2016) What makes an online consumer review trustworthy? Ann Tour Res: A Soc Sci J 58:46–64
He S, Wang Y (2022) Evaluating new energy vehicles by picture fuzzy sets based on sentiment analysis from online reviews. Artif Intell Rev 56:1–22
Hong H, Xu D, Wang A (2017) Understanding the determinants of online review helpfulness: A meta analytic investigation. Decis Support Syst 102:1–11
Hu H, Krishen A (2019) When is enough? Investigating product reviews and information overload from a consumer empowerment perspective. J Bus Res 100:27–37
Hu N, Pavlou P, Zhang J (2009) Overcoming the J-shaped distribution of product reviews. Commun Acm 52(10):144–147
Huang N, Burtch G, Hong K, Polman E (2016) Effects of multiple psychological distances on construal and consumer evaluation: A field study of online reviews. J Consum Psychol 26:474–482
Huang P, Huo H (2022) Influence of reviewer characteristics on the usefulness of hotel online reviews: A joint test based on big data and experimental methods. Resour Dev Mark 38(5):634–640
Ismagilova E, Slade E, Rana N, Dwivedi Y (2019) The effect of electronic Word of Mouth communications on intention to buy: A Meta-Analysis. Inf Syst Front 22:1203–1226
Jiang D, Zou J (2023) Is the emoji redundant? Effects of emojis on the perceived usefulness of online product reviews. J Ind Eng Eng Manag 37(1):66–77
Jürgen N, Dominik G, Dennis K (2023) Reviewing from a distance: Uncovering asymmetric moderations of spatial and temporal distance between sentiment negativity and rating. MIS Q 47(4):1709–1726
Kim J, Li X, Jin L, Kim J (2024) Attentive review semantics-aware recommendation model for rating prediction. Electronics 13(14):2815. https://doi.org/10.3390/electronics13142815
Kim S, Lee K, Shin S, Yang S (2017) Effects of tourism information quality in social media on destination image formation: The case of Sina Weibo. Inf Manag 54:687–702
Krosnick A, Berent K (1993) Comparisons of party identification and policy preferences. Public Opin Q 57(1):1–28
Kwok L, Xie L (2016) Factors contributing to the helpfulness of online hotel reviews Does manager response play a role? Int J Contemp Hospitality Manag 28(10):2156–2177
Lee K, Choi M, Park N (2020) The power of motion: How video reviews influence consumer engagement in the automotive industry. Computers Hum Behav 112:106478. https://doi.org/10.1016/j.chb.2020.106478
Lee M, Jeong M, Lee J (2017) Roles of negative emotions in customers' perceived helpfulness of hotel reviews on a user-generated review website a text mining approach. Int J Contemp Hospitality Manag 29(2):762–783
Lei J, Zhu C, Yang S, Wang J, Yu Y (2023) Influence of review properties in the usefulness analysis of consumer reviews: A review-based recommender system for rating prediction. Neural Process Lett 55:11035–11054
Li C, Cui G, He Y (2020a) The role of explanations and metadiscourse in management responses to anger-reviews versus anxiety-reviews: The mediation of sense-making. Int J Hospitality Manag 89:102560
Li D, Guo Y (2022) Study on consumers' perception and purchase decision of new energy vehicle. Hubei Soc Sci 8:55–65
Li H, Ji H, Liu H, Cai D, Gao H (2022) Is a picture worth a thousand words? Understanding the role of review photo sentiment and text-photo sentiment disparity using deep learning algorithms. Tour Manag 92:104559
Li L, Wang Z, Chen L, Wang Z (2020b) Consumer preferences for battery electric vehicles: A choice experimental survey in China. Transportation Res Part D: Transp Environ 78:102185
Liberman N, Trope Y (1998) The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. J Personal Soc Psychol 75(1):5–18
Liang S, Schuckert M, Law R (2018) How to improve the stated helpfulness of hotel reviews? A multilevel approach. Int J Contemp Hospitality Manag 31(2):953–977
Liu B (2012) Sentiment analysis and opinion mining. Springer Nature Switzerland AG 2012
Liu A, Xie Y, Zhang J (2019) It's not just what you say, but how you say it: The effect of language style matching on perceived quality of consumer reviews. J Interact Mark 46:70–86
Liu H, Feng L (2022) The study of new energy vehicle choice in China from the perspective of complex neural network. Front Phys 10:1015103

Liu Y, Teng S (2021) Consumer decision-making in the vehicle industry: The role of online reviews. J Retail Consum Serv 58:102317
Liu Z, Park S (2015) What makes a useful online review? Implication for travel product websites. Tour Manag 47:140–151
Lu B, Ma B, Cheng D, Yang J (2023) An investigation on impact of online review keywords on consumers' product consideration of clothing. J Theor Appl Electron Commer Res 18:187–205
Malik M (2020) Predicting users' review helpfulness: The role of significant review and reviewer characteristics. Soft Comput 24:13913–13928. https://doi.org/10.1007/s00500-020-04767-1
Mousavizadeh M, Koohikamali M, Salehan M, Kim D (2022) An investigation of peripheral and central cues of online customer review voting and helpfulness through the lens of elaboration likelihood model. Inf Syst Front 24:211–231
Mudambi M, Schuff D (2010) What makes a helpful online review? A study of customer reviews on Amazon.com. MIS Q 34(1):185–200
Nelson R, Todd P, Wixom B (2005) Antecedents of information and system quality: An empirical examination within the context of data warehousing. J Manag Inf Syst 21:199–236
Park J, Kim S, Lee H (2021) How technicality in automotive reviews shapes consumer trust: A text mining approach. J Consum Behav 40:245–260. https://doi.org/10.1002/cb.1932
Pawel B, Shuvam C, Beata C (2023) Consumer adoption of electric vehicles: A systematic literature review. Energies 16(1):205. https://doi.org/10.3390/en16010205
Petty E, Cacioppo T (1984) The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion. J Personal Soc Psychol 46:69–81
Qiu K, Zhang L (2024) How online reviews affect purchase intention: A meta-analysis across contextual and cultural factors. Data Inf Manag 8:100058. https://doi.org/10.1016/j.dim.2023.100058
Racherla P, Friske W (2012) Perceived usefulness of online consumer reviews: an exploratory investigation across three services categories. Electron Commer Res Appl 11(6):548–559
Riley W, Hovland I, Janis L (1954) Communication and persuasion: psychological studies of opinion change. Am Sociological Rev 19(3):355–365
Saeed I (2016) Semantics (4th ed.). Wiley-Blackwell
Salehan M, Kim D (2016) Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics. Decis Support Syst 81:30–40
Simon Q, Mario P, Iris V (2015) When consistency matters: the effect of valence consistency on review helpfulness. J Computer-Mediated Commun 20(2):136–152
Sun Y, Han X, Feng J (2019) Helpfulness of online reviews: Examining review informativeness and classification thresholds by search products and experience products. Decis Support Syst 124:113099
Sussman W, Siegal S (2003) Informational influence in organizations: An integrated approach to knowledge adoption. Inf Syst Res 14:47–65
Tsai C, Chen K, Hu Y, Chen W (2020) Improving text summarization of online hotel reviews with review helpfulness and sentiment. Tour Manag 80:104122. https://doi.org/10.1016/j.tourman.2020.104122
Wang L, Chen J (2023) The role of reviewer expertise in the perceived usefulness of online reviews. Tour Manag 85:104301
Wang M, You H, Ma H, Sun X (2023a) Sentiment analysis of online new energy vehicle reviews. Appl Sci 13:8176
Wang R, Strong D (1996) Beyond accuracy: What data quality means to data consumers. J Manag Inf Syst 12:5–34
Wang X, Cheng Y, Lv T, Cai R (2023b) Fuel vehicles or new energy vehicles? A study on the differentiation of vehicle consumer demand based on online reviews. Mark Intell Plan 41:1236–1251
Wang X, Lv T, Deng X (2024) Research on content analysis and quality evaluation for online reviews of new energy vehicles. Information Development. https://doi.org/10.1177/02666669241257186
Wang X, Wang J, Wang X (2023c) Characterization of new energy vehicle user behavior in the private sector. Automob Practical Technol 48:41–45
Wang Y, Han M, Li K (2025) What makes a review voted? A mixed-method analysis of review content and voter characteristics. Int J Hospitality Manag 129:104207
Wooldridge J (2019) Introductory econometrics: A modern approach (7th ed.) Thomson South-Western
Wu C, Mai F, Li X (2021) The effect of content depth and deviation on online review helpfulness: Evidence from the double-hurdle model. Inf Manag 58:103408
Xiang Z, Du Q, Ma Y, Fan W (2017) A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. Tour Manag 58:51–65
Yan Q, Zhou S, Wu S (2018) The influences of tourists' emotions on the selection of electronic word of mouth platforms. Tour Manag 66:348–363
Yi M, He X, Hao L (2020) Research on the compulsive purchase propensity of introverted members in online consumer communities: Based on the conformity theory. J Manag 33:49–60
Zhan M, Liu L, Wang X, Ju R (2023) The role of explained actions and reactions in the helpfulness of online reviews. Electronic Commerce Res. https://doi.org/10.1007/s10660-023-09752-y
Zhang H, Chen X, Li Y (2021) The subsidy retreat policy and its impact on China's new energy vehicle market. Energy Econ 99:105276
Zhang Y, Wang X (2021) The impact of reviewer identity on consumer trust and purchase intentions in online reviews: A meta-analysis. J Consum Behav 20(3):245–260
Zhang Z, Zhang Z, Yang Y (2016) The power of expert identity: How website-recognized expert reviews influence travelers' online rating behavior. Tour Manag 55:15–24
Zhao K, Stylianou A, Zheng Y (2017) Sources and impacts of social influence from online anonymous user reviews. Inf Manag 55:16–30
Zhou Y, Yang S, Li Y, Chen Y (2020) Does the review deserve more helpfulness when its title resembles the content? Locating helpful reviews by text mining. Inf Process Manag 57:102179
Zhu L, Li H, He W, Hong C (2020) What influences online reviews' perceived information quality? Perspectives on information richness, emotional polarity, and product type. Electron Libr 38:273–296
Zhuang W, Zeng Q, Zhang Y, Liu C, Fan W (2023) What makes user-generated content more helpful on social media platforms? Insights from creator interactivity perspective. Inf Process Manag 60:103201. https://doi.org/10.1016/j.ipm.2022.103201

Acknowledgements
This study is funded by Lianyungang "521" Project [Grant No. 2023].

Author contributions
Conceptualization: Xiaoguang Wang, Yijun Gao; Data curation: Lei Fan, Nan Li; Formal analysis: Xiaoguang Wang, Nan Li; Methodology: Yijun Gao, Nan Li; Software: Xiaoguang Wang, Lei Fan; Visualization: Yijun Gao, Lei Fan, Nan Li; Writing – original draft: Xiaoguang Wang, Nan Li; Writing – review & editing: Xiaoguang Wang, Yijun Gao, Nan Li.

Competing interests
The authors declare no competing interests.

Ethics approval
Since historical, publicly accessible social media data from Autohome.com.cn were used for the computational aspects of this research and it involved no interaction with the authors of the social media postings, this part did not constitute human subjects research and therefore ethics approval was not required.

Informed consent
Informed consent was not required for this secondary data analysis study.

Additional information
Correspondence and requests for materials should be addressed to Xiaoguang Wang.