

Article

Consumer Sentiment Involvement in Big Data Analytics and Its Impact on Product Design Innovation

Chang Liu ^{1,*} and Yueli Xu ²

¹ College of Design and Innovation, Tongji University, Shanghai 200092, China

² College of Environmental Science and Engineering, Donghua University, Shanghai 201620, China; moonlee.xu@gmail.com

* Correspondence: liuc@tongji.edu.cn

Abstract: With the development of Internet technology and the digital market in China, consumption through e-commerce platforms and giving online reviews after purchase have become mainstream. However, current market research still uses traditional methods of surveys and questionnaires, which largely influences the efficiency of relevant enterprises in finding key product issues and lowers new product performance. This paper conducts the aspect-based sentiment analysis (ABSA) method to study the impact of consumer sentiment involvement (CSI) on new product performance. We took the ceramic industry as a case study, and collected 3.22 million consumer responses for ABSA. A total of 22 performances of new products were analyzed for CSI hypothesis testing. We found that CSI big data analytics were positively related to new product performance and enterprise innovation. Our study contributes in three ways. First, we extend the concept of co-creation in open innovation studies into an intelligent data intensive context. The data-driven open innovation is conducive to product design and new product performance. Second, we enrich the empirical industry of ABSA sentiment analysis in the Chinese ceramic industry. Third, we contribute to the collaboration method to motivate consumers to participate, which means to make effective product reviews in this research.

Keywords: product design innovation; consumer sentiment involvement; digital market; big data; sentiment analysis; open innovation; China



Citation: Liu, C.; Xu, Y. Consumer Sentiment Involvement in Big Data Analytics and Its Impact on Product Design Innovation. *Sustainability* **2021**, *13*, 11821. <https://doi.org/10.3390/su132111821>

Academic Editors: Eloy Gil-Cordero and Juan Pedro Cabrera-Sánchez

Received: 31 July 2021

Accepted: 25 October 2021

Published: 26 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the rise of the Internet economy, online consumption has gradually become a significant economic growth force. The base of Internet users in China provides an advantageous platform for the digital economy and digital retailing development. China's Internet penetration rate reached 70% by the end of 2020, with a total digital economy of 3.92 billion RMB in the same year, accounting for 38.6% of total GDP [1,2]. On the premise that e-shopping is guaranteed, consumers tend to shop online and give feedback for products since the purchase risks are relatively low [3]. Their proactive reviews after purchase are valuable for market research in all industries and enterprises.

For most Chinese enterprises, the main methods of market research still use traditional surveys and questionnaires, which are hard to manage massive amounts of data on consumers' reviews on E-commerce platforms and social media. Moreover, the traditional methods cannot have a comprehensive evaluation of related products, resulting in many products using experience and feedback will not be included in surveys and questionnaires. This greatly affects the efficiency and accuracy of relevant enterprises in finding key product issues, and it is difficult to improve these key issues in their iterative products. Thus, the performance of these enterprises' new products is not satisfactory.

The importance of collaborating with consumers in the product design and innovation process has been recognized for many years [4]. Consumers play the roles of "data provider" and "data analyst", rather than the passive ones of the past [5,6]. The data-rich environment provides unprecedented opportunities for product innovation, which involves more and

more consumers in the design and innovation process [7]. Online commercial platform [8], social media [9,10], product feedback [11], and mobile phone [12] can all become channels to obtain and involve consumers in the innovation process.

Extant literature has claimed that consumers are a crucial part of the open innovation process, since they serve as not only the innovators but also beneficiaries from using related products [13,14]. However, the concept of consumer involvement with the use of big data is underexplored. The traditional method of gathering consumers' reviews has been insufficient to understand the needs of consumers in the Internet age. Individualized needs, public comments, and the ability to accurately obtain open data are preconditions of consumer-oriented open innovation in the current age [15,16].

We argue that consumer sentiment involvement (CSI) could improve the new product performance through big data analysis. One of the core characteristics of big data analysis is to understand consumer needs in the new products design [17,18]. Understanding consumers' emotions is of great value for marketing and financial market prediction in the business world [19]. Many scholars have studied that sentiment analysis of a large number of consumer reviews on e-commerce platforms can improve their satisfaction based on different models [20]. Consumers' sentimental expression of the product can intuitively reflect their feedback and demands for the product, whose simply positive or negative sentiments could represent both components and attributes of the entity [21,22]. Following this logic, we categorize CSI in big data analytics into three dimensions: *positive*, *neutral*, and *negative*, which, respectively, represent consumers' emotional feedback on the product.

This paper conducted the aspect-based sentiment analysis (ABSA) method and big data analysis to attain a comprehensive dataset of consumers' feedback. We took the ceramic industry as a representative case study, and collected 3.22 million online consumers' feedback on related ceramic products to locate the positive and negative parts of the related products from the consumer's side. Furthermore, we developed 22 new test projects for hypothesis testing, cooperated with firm D in Longquan city, China. We found that the consumer sentiment involvement in big data analytics was positively related to new product performance and enterprise innovation process.

We apply the consumer-involved big data analysis on the ceramic industry and its products in China, and argue that our findings contribute to the literature in three ways. First, we extend the concept of co-creation in open innovation studies into an intelligent data intensive context. We suggest that extracting consumers' sentiment polarities of their reviews on product use and design can help related companies to find major issues in product details. This data-driven and consumer-included open innovation is conducive to product design and new product performance. Second, we enrich the empirical industry of ABSA sentiment analysis. We apply this method to China's ceramic industry and its related products based on lexicon and rule set. Based on the existing lexica list, we add ceramic-specific words to enhance the scope and accuracy of this method. Third, we contribute to co-creation literature on how to motivate consumers to participate in the collaboration, which means to make effective product reviews and suggestions in this research. We investigate factors that influence consumers' willingness of giving valuable reviews and useful feedback on products, which are largely related to enterprises' timely responses and explanations of their comments. These findings also contribute to enterprises and industries related practitioners through improving their understandings of how to use big data analytics in the innovation process and managing online consumers' reviews with different needs in the Internet age.

2. Materials and Methods

We examine how CSI impacts new product performance. New product performance refers to the online commercial outcomes in related new products, which includes volume (product reviews) and valence (product ratings) of online reviews. The conceptual framework is displayed in Figure 1.

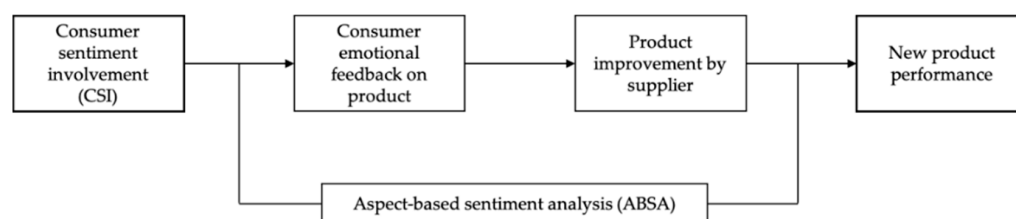


Figure 1. Conceptual framework.

CSI suggests that consumers' sentiment towards products is a direct and effective application of big data in the open innovation process. First, it is difficult for consumers to give professional and clear feedback on products when they comment on online shopping platforms; instead, it is easy for consumers to express their emotional inclinations towards products. Second, when their sentiment is categorized as *positive*, *neutral*, and *negative* towards related products, there is greater consumer engagement in data analytics on the product, which largely expands the sample size of product feedback research compared with traditional market research methods. Third, consumers bring innovations and form a virtuous circle, since they are not only users but also creators of the innovation production line. These creations form a new batch of products based on consumers' feedback of their using and experiencing needs.

The aspect-based sentiment analysis (ABSA) is conducted to collect consumers' feedback and analyze their sentiment towards the product. The result is used as the basis for suppliers to improve their related products, which will have a positive impact on the new product performance and the innovation process in our hypothesis.

ABSA is an important tool for the extraction of human emotional state information. It is mainly used for aspect extraction and sentiment classification of product reviews and sentiment classification of target-dependent Internet users [23]. Chinese text has its own characteristics that other languages do not have, the knowledge-based approach and the machine learning based approach are always used in monolingual approach for direct Chinese language [24]. Commonly used Chinese open-source sentiment dictionaries are: HowNet, the sentiment vocabulary ontology library of Dalian University of Technology, and the simplified Chinese sentiment polarity lexicon of Taiwan University (NTUSD) [25].

In this paper, we refer to the sentiment lexica including the sentiment vocabulary ontology library of Dalian University of Technology and HowNet. We refer to the modal dictionaries of sentiment dictionary developed by Yang and Wang [26]. The process of sentiment analysis is displayed in Figure 2.

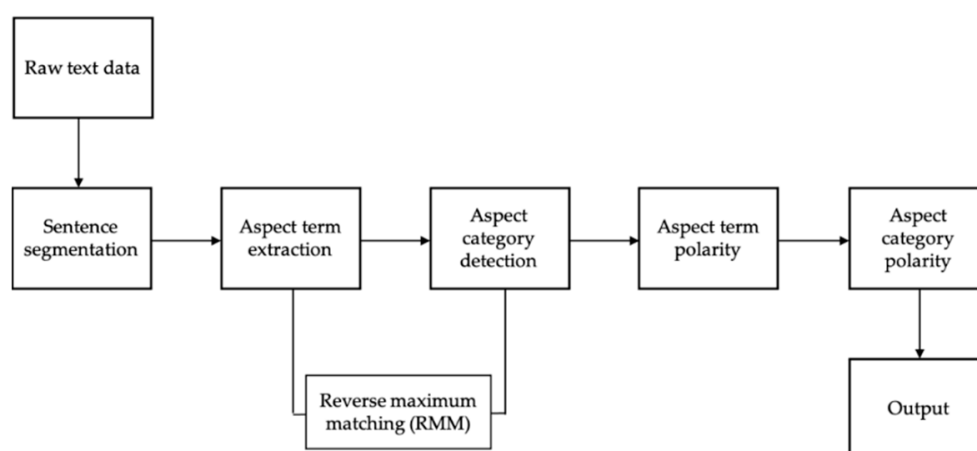


Figure 2. Method of ABSA.

The first step consists of collecting the raw data from Internet user-generated content in the targeted platforms. These data are unstructured and disorganized, expressed in diverse

ways with different vocabularies, sentence patterns, and writing styles, etc. Different from traditional manual analysis methods, the big data analysis of natural language processing allows these massive data to be filtered and useful information extracted.

The second step is to segment sentences into phrases and extract aspect term. The sentence segmentation uses common separators to cut into phrases such as comma, full stop, question mark, and exclamation mark. Due to the particularity of the Chinese language, a word often cannot express any meaning, we need to understand the complete emotional meaning through finding the maximum complete meaning of the clause. Therefore, we use Jieba as a tool to extract the aspect term and to create the aspect term dictionary. In the Chinese context, Jieba is the most commonly used in word graph scanning in academic and industry fields, which implements word graph scanning based on the Trie tree model to obtain all possibilities of the Chinese characters in the target clause [27]. Thus, dynamic programming is used to find the maximum probability path and finds the maximum segmentation combination based on word frequency, in order to establish a general keyword dictionary.

The third step is to detect the aspect category in hierarchy layers and keywords of the aspect category, which is used for long and complicated sentiment sentence to establish an aspect category and to give a classified evaluation and feedback of targeted sentences. In this process, the reverse maximum matching (RMM) method based on the extant dictionary is used to ensure the extracted aspect term expressing the most complete meaning of the processed text. Each time the last i characters (the threshold i is determined by the word segmentation) are taken as matching fields. If the matching fails, then one word will be removed from the beginning and the matching process continues. Here, when a keyword is successfully matched with the first layer and second layer, then the matching process stops. If a keyword corresponds to multiple second layers, only the earliest match will be retained. After searching all the sub-sentences, if none of the keywords match the first layer and the second layer, the first layer and the second layer will be “others”.

The fourth step is to classify the sentiment through aspect term polarity, which uses Jieba tool to obtain a sentiment dictionary. It first matches negative sentiment words in the text. If matched, the words would go to the negative sentiment words list. If not, it will continue to match the positive sentiment words. After it is matched, the words go to the positive sentiment words list. If still unmatched, then the sentiment will be judged as neutral. The longest sentiment words will always be matched first. For example, “the price is high”, “the price is too high”, and “the price is extremely high” are all negative sentiment words, and the priority match will be “the price is extremely high”. Some words are also set as reversal words. For instance, if “expensive” is a negative sentiment word, then “not” will convert the emotion into a positive polarity, which makes the final “not expensive” a positive sentiment sentence.

The final step is the preparation of output by aspect category polarity, which converts the unstructured text into meaningful information. The aspect term polarity combines and calculates based on a regular expression. Given a set of pre-defined aspect categories, which decides the polarity (positive, or negative, or neutral) into each aspect category. Each positive word was assigned a value of +1, and each negative emotion word was assigned a value of −1. It is assumed that the emotional value satisfies the principle of linear superposition and negative words will cause emotional reverse. Therefore, we judge the sentiment of the sentence based on the sum of positive and negative values.

3. Results

3.1. Case Study: Chinese Ceramic Industry

We use the Chinese ceramic industry as the representative case to explore the impacts of the big data analytics method on the new product performance and the innovation process. We choose the ceramic sector for two main reasons. On the one hand, the ceramic industry is one of the most traditional industries in China, which mainly uses traditional market research methods. This results in a great lagging development under the background of the rapid

development of the Internet in China. There are many types of ceramic products in China, and their sales are all over the country. It is impossible to collect market research on the shape, material, design, color, and other aspects of ceramic products timely and accurately through the traditional manual survey method. On the other hand, Chinese ceramic handicrafts still adopt the traditional closed innovation mode; that is, it mainly depends on the personal preferences of designers or related personnel to determine new products' design, which leads to the lack of innovative breakthroughs in products. In this process, the massive demands of consumers and users have been seriously ignored. Therefore, we focused on how to use the sentiment analysis of E-commerce big data to promote the new product performance and product design and innovation process in our case study.

We cooperated with ceramic firm D in Longquan city, and adopted big data analytics to involve consumers in the innovation process. We studied the impact of consumers' sentiment feedback on new ceramic products performance by ABSA method. The data source was consumers' comments from the public online platform, which covered the time range from January 1st 2017 to December 31st 2018. The total number of original online public comments was 3.22 million, which covered 2,395 online ceramic stores, and 77 thousand stock keeping units. The average length of product reviews is between 16 and 17 characters, excluding punctuation marks.

After the preliminary data sorting, we obtained 7.72 million NLP clauses records, and 7 major aspect categories: product, logistics, price, brand, packaging, service, and activity. The numbers of these categories are displayed in Table 1.

Table 1. Seven major aspect categories.

Aspect Category	Number of Records
Product	4,249,493
Logistics	1,020,448
Price	707,501
Brand	664,381
Packaging	517,493
Service	360,061
Activity	225,563

Second, the sentence segmentation cuts sentence with the common separators such as comma, full stop, question mark, and exclamation mark. For the sentiment sentence "The ceramic teapot is small in shape, good in the porcelain producing, exquisite and is excellently matched with the jasmine tea ceremony" (originally in Chinese: *chahu zaoxing xiaoqiao, cizhi hao, xiuzhen jingzhi, mailai pao moli huacha de, hen dapei.*), there were five sub-sentences: "the ceramic teapot is small in shape", "good in the porcelain producing", "exquisite", "jasmine tea ceremony", and "excellent matched". Then, we used aspect term of extraction to create the aspect term dictionary. Most words had been added in Chinese Jieba algorithm, but some industrial professional words may not be included, such as "glazing color" (originally in Chinese: *youse*) in the ceramic industry. These professional words were manually labeled and added to the general keyword dictionary.

Third, aspect category detection was used to detect the aspect category in hierarchy layers and the keywords of the aspect category, especially for long and complicated sentiment sentences. For the sentiment sentence "This ceramic teapot is bought as a gift for my husband. I like the ancient Chinese style which is very beautiful. The quality of the product is good and the price is fair. But, the logistics is too slow" (originally in Chinese: *zhege chahu shi mailai songgei laogong de, xihuan zhongguofeng de zaoxing, hen piaoliang, wumei jialian, dan meizhongbuzu de shi wuliu tai man le*). Two layers of this sentence are recognized, which are displayed in Table 2.

Table 2. The aspect category detection.

	1st Layer	2nd Layer	Keywords
1	Logistics	Speed of logistics	The logistics is too slow
2	Price	Price	The quality is good and price is fair
3	Product	Target user	Husband
4	Product	Appearance	Beautiful
5	Product	Design	Ancient Chinese style

Last, the aspect category polarity was used to decide the polarity (i.e., positive, negative, or neutral) of each aspect category. For the sentiment sentence “This ceramic cup is a gift for my bestie. It is the ancient Chinese design style and it also looks quite transparent. The product quality is good and the price is fair. The logistics is also very fast. This is a pleasant shopping experience and I am quite satisfied with it” (originally in Chinese: *songgei guimi de liwu, sheji poju zhongguofeng, chabei kanqilai ye tebie tongtou, wumei jialian, wuliu ye henkuai, hen manyi de yici gouwu*). The aspect and polarity collocation extraction attempted to extract the collocation <bestie, gift>, <ceramic cup, transparent>, <product quality, good>, <price, fair>, <logistics, fast>, <experience, pleasant and satisfied>. The sentiment recognition result is shown in Table 3.

Table 3. Result of the sentiment sentence analysis.

	Sentence Segmentation	Aspect Category Polarity	Aspect Term of Extraction	Sentiment Word	1st Layer	2nd Layer
1	A gift for my bestie	Positive	Gift, bestie, for	/	Product	Target user
2	The ancient Chinese design style	Positive	Design, ancient Chinese	/	Product	Design
3	The ceramic cups look transparent	Positive	Cups, look, transparent	transparent	Product	Appearance
4	The product quality is good and price is fair	Positive	Good quality, fair price	Product quality is good and price is fair	Price	Price
5	The logistics is also very fast	Positive	Logistics, also, very, fast	fast	Logistics	Speed of logistics
6	This is a pleasant shopping experience and I am quite satisfied with it	Positive	Experience, satisfied, pleasant, very	Pleasant and satisfied	Product	General preference

Since product innovation and product analysis was the main focus, we picked the product category and made similar subdivisions. We obtained 4.2 million product-related NLP clauses records, and we categorized them into 12 sub-categories of the product aspect with their sentiment rate, which could be found in Table 4.

Table 4. 12 sub-categories of the product aspect category and sentiment rate.

Category	Sub-Category	Number of Records	Positive Rate	Neutral Rate	Negative Rate
Product	General preference	992,874	91.2%	5.1%	3.7%
Product	Quality	854,547	83.8%	5.8%	10.4%
Product	Appearance	822,014	89.0%	7.1%	4.0%
Product	Workmanship	455,474	81.0%	11.0%	8.0%
Product	Shape	385,697	89.3%	6.0%	4.7%
Product	Target user	264,232	88.9%	6.8%	4.3%
Product	Specialty	170,379	76.6%	5.6%	17.9%
Product	Usage scenario	134,211	87.5%	7.5%	5.0%
Product	Material	127,964	79.0%	5.5%	15.5%
Product	Color	29,588	88.9%	7.1%	4.0%
Product	Design	10,918	87.8%	7.1%	5.1%
Product	Personalized customization	1595	85.8%	8.1%	6.1%

3.2. Precision, Recall, Accuracy and Inter-annotator Agreement

The precision, recall, F1, and accuracy of the result are calculated in this session. Two annotators (A and B), respectively, verify the same 1000 randomly sampled NLP results, and calculate category accuracy, positive sentiment accuracy, and negative sentiment accuracy. True positive (TP) is the number of NLP clauses which classified as “product” and manually verified as correct; true negative (TN) is the number of NLP clauses which classified as other categories and manually verified as correct; false positive (FP) is the number of NLP clauses which classified as “product” and manually verified as wrong; and false negative (FN) is the number of NLP clauses which classified as other categories and manually verified as wrong.

The results of the aspect category verification (Table 5) are:

Table 5. The aspect category verification.

Annotator	Category	TP	TN	FP	FN	Precision	Recall	F1-Score	Accuracy
A	Product	583	260	86	67	87.2%	89.8%	88.5%	84.7%
B	Product	601	254	83	62	87.9%	90.6%	89.2%	85.5%

The results of positive sentiment verification (Table 6) are:

Table 6. The positive sentiment verification.

Annotator	Sentiment	TP	TN	FP	FN	Precision	Recall	F1-Score	Accuracy
A	Positive	343	77	64	58	84.3%	85.5%	84.9%	77.5%
B	Positive	337	68	72	65	82.4%	83.8%	83.1%	74.7%

The results of negative sentiment verification (Table 7) are:

Table 7. The negative sentiment verification.

Annotator	Sentiment	TP	TN	FP	FN	Precision	Recall	F1-Score	Accuracy
A	Negative	288	71	65	34	81.6%	89.4%	85.3%	78.4%
B	Negative	293	69	68	28	81.2%	91.3%	85.9%	79.0%

The Pearson correlation coefficient is used to calculate the inter-annotator agreement between annotator A and B. Within 90% confidence interval, their Pearson correlation coefficient is shown in Table 8.

Table 8. The Pearson correlation coefficient of accuracy.

	Category Accuracy	Positive Sentiment Accuracy	Negative Sentiment Accuracy
Pearson Correlation Coefficient	0.97	0.93	0.98

3.3. Error Analysis

After a thorough analysis of the errors, the sources of which could be mainly classified into category detection mismatch, user error (including the spelling, grammar, and characters), indirect expression (which will mislead the terms of expression), and the algorithm. The distribution of the error sources is shown in Figure 3.

The category detection mismatch mainly stems from the different meanings of the same word in Chinese, which may mismatch the category and mislead the terms of expression. For instance, in the sentence “the product has a high value in beauty and shape” (originally in Chinese: *zhe kuan chanpin yanzhi hengao.*), the category is detected as in the *price* category, and the keyword is *value*, the result is *negative* in *price* due to the “high value.” However, the real category of the sentence should be in the *product* category instead of *price* category, and the sentiment is *positive*. Similarly, the indirect expression also may mislead the expression and the aspect category polarity analysis. These errors are

corrected by the manual verification of the result. In this study, the keyword dictionary and sentiment word dictionary are continuously updated and optimized to improve the accuracy of the results through manual verification.

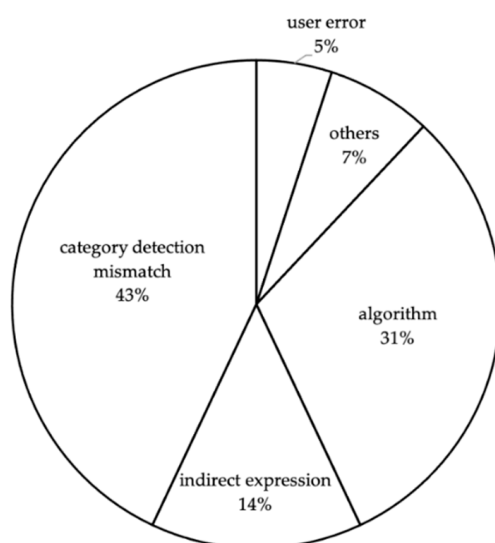


Figure 3. The distribution of the error sources.

3.4. Consumer Sentiment Involvement Innovation

We discussed with the cooperated ceramic firm D about the two aspects of the highest negative sentiments (i.e., product specialty and material), regarding the consumer sentiment analysis of product segmentation (see Table 4). After analyzing the original consumer comments and interviewing ceramic designers of firm D, we carried out two actions to integrate consumer sentiment feedback on ceramic products.

For ceramic material, we discussed with ceramic designers in firm D and decided to introduce the cross-border design to increase the material richness. On the one hand, all ceramic products of firm D were celadon (one of the sub-categories of ceramics glazed in the jade green), and their glaze color was relatively single, which made it easy for consumers to have aesthetic fatigue. On the other hand, the ceramic material was mainly clay; a high temperature was needed to set the shape during the production process, which greatly limited the morphological expression of products. Therefore, ceramic designers of firm D started to experiment with cross-border designs with handicraft designers of various materials. They could come from various industries, including woodcarving, colored glaze, metalwork, and textile, etc.

For product specialty, we discussed with the chief executive officer (CEO) of firm D about their strategic considerations. We decided to create a series of ceramic works and gradually build its intellectual property (IP). In Longquan city, the inheritance of ceramics was mainly taught by masters and apprentices, and ceramic works also respected ancient customs. During our interviews, some ceramic designers and craftsmen said that the different historical stages of celadon had standard styles (such as the curvature of the spout, the width of the cup, etc.) and the corresponding producing methods. During the apprenticeship, the closer they were to the standard style, the more being appreciated by their teachers. As a result, the ceramic firms in Longquan city were similar in designing and manufacturing ceramic products. The lack of specialty greatly affected the experience of consumers. Therefore, we used a series of products as an experiment to promote the specialty of firm D's ceramic products. When we were conducting this experiment, it was just before the Chinese New Year in 2019, which was the Pig Year on the Chinese lunar calendar. Firm D designed and launched a series of celadon products based on the image of a pig, including tea cans, teacups, and piggy banks, etc.

These new products were open to consumers in online stores in August 2019, after the process of design, proofing, modification, and production. We collected online data and consumers' feedback for four months from August to December 2019.

3.5. New Product Performance

New product performance is defined as the online commercial outcomes in related new products. We refer to the extant literature and use 6 items to measure the performance, which are (a) sales volume, (b) volume of page views, (c) volume of product reviews, (d) valence (average rating), (e) percentage of positive reviews, and (f) percentage of negative reviews [28].

Over four months, we collected data on 22 new products of firm D, including 10 cross-border design products and 12 series products. Table 9 provides the summary statistics, including the mean and standard deviation of variables (SD). On average, each cross-border design product received 15 reviews, and each series product received 30 reviews. Cross-border design products on average attracted more page views (1084.4), thus having greater dispersion than series products (743.42). The average ratings were similar, cross-border design products (4.89) received slightly higher valence than series products (4.88). The percentage of positive reviews for cross-border design products (81.70 percent) was more comparable with that for series products (78.75 percent), whereas series products received a greater percentage of negative reviews (2.42 percent) than cross-border design products (2.40 percent).

Table 9. Key descriptive statistics.

	Mean	SD	Minimum	Maximum
Cross-border design products ($n = 10$)				
Sales volume	77.30	48.33	4.00	152.00
Volume of page views	1084.40	1086.93	35.00	3675.00
Volume of reviews	15.40	10.91	0.00	34.00
Valence (average rating: 1–5)	4.89	0.09	4.70	5.00
Percentage of positive reviews ¹	81.70	28.29	0.00	100.00
Percentage of negative reviews ²	2.40	2.58	0.00	8.00
Series products ($n = 12$)				
Sales volume	222.67	354.51	6.00	1299.00
Volume of page views	743.42	1116.11	57.00	3572.00
Volume of reviews	30.33	43.38	0.00	157.00
Valence (average rating: 1–5)	4.88	0.09	4.70	5.00
Percentage of positive reviews ³	78.75	35.52	0.00	100.00
Percentage of negative reviews ⁴	2.42	3.12	0.00	9.00

^{1,2,3,4} The percentages of positive reviews and negative reviews do not add up to 100 percent, because the percentage of neutral ratings is not included.

Due to the limitation of time and the samples provided by firm D, we were unable to establish a metric model for the follow-up performance of new products for long-term tracking and adjustment. However, in terms of the performance of new products in these four months, the average percentage of positive reviews increased by 3.1% compared with that in the first half of the same year, and the number of firm D's online store visits increased by 16.7% in four months.

4. Discussion

In this research, we proposed the consumer sentiment involvement in big data analytics, and examined its impacts on the new product performance, taking the ceramic industry as a case study. Our findings showed that the consumer sentiment involvement facilitated new product performance in the e-commercial domain.

Although each online comment only represented consumers' individual feedback on products, it will have a huge impact on the creation and sales of new products if the number of consumers was large enough. The volume of page views signals the popularity of a product and had a positive relationship with the sales volume [29]. This helps those firms have sufficient preparations of manufacturing and stocking up, especially for products

requiring a certain production cycle such as ceramics. The ability to predict the product output in advance is of great help to firms and manufacturers. Valence and comment are direct feedback of consumers' experience of using the product, which also reflect the degree of consumer's involvement in related product innovation [30].

When consumers' sentiment is involved in the process of the new product of innovation, it not only is prone to focalizing on the shortcomings of previous products, but also integrates consumers' demands for products into the innovation process. Through the big data analytics on the online consumers' reviews, it is easy to locate the sub-category that urgently needs to be improved in each category based on the highest negative reviews. In our case study, specialty and material are two common weaknesses of ceramic products currently on the market. It is precise because of consumers' negative sentiment toward the material that we began to promote cross-border design cooperation, which achieved good performance during the four-month follow-up period.

For firms, consumer sentiment involvement innovations are more targeted and more efficient, especially in the current data-rich environment. Firms work closely with consumers and suppliers, which provides crucial sources to obtain new insights and knowledge from both sides [31,32]. The impact of consumers' feedback on new product design is positive, since the review contains their knowledge and capabilities [33]. Those firms' learning from consumers would push toward collaborative design and more open forms of innovation [34,35]. During this process, the firm's direct and indirect interaction with consumers would greatly affect the probability of consumers leaving valid and useful reviews. According to our four-month observation, some consumers left negative reviews because they had not known how to use the product properly. For this group of people, the online store's quick response and explanation could greatly improve the product's usage condition. When consumers make better use of the product, they tend to give follow-up detailed feedback and to establish a product-related community, which might attract more people to give comments and reviews after purchase.

There are several suggestions for future research. From the theoretical aspect, the extant open innovation literature infers that the ability of firm's knowledge integrates with outside knowledge determines the success of product performance. However, during our research, one of the biggest difficulties is how to make the innovation process of the enterprise better integrate with the rapidly changing market environment. We have shortened the time to process consumers' feedback through big data analytics and natural language processing, but it is still not fast enough for some short-term products that require immediate feedback. From the technical aspect, the aspect-sentiment analysis has a large volume of data with diversified data sources, including structured data and unstructured data such as text comments, images, and voices. Our study's analytics are based on the text comments information, whereas images and voice information have not been utilized yet. With the development of the 5G network, consumers' commenting methods also tend to be diversified, and many photos, emotions, voices, and pieces of video information also appear in their reviews. How to make better use of non-text information needs further study in future research.

5. Conclusions and Limitations

For China's traditional industries and enterprises, this study is of great significance. On the one hand, many traditional industries in China have a long history (some cultural and handicraft industries are even more than thousands of years old); these industries and their products tend to underline much on historic inheritance rather than the changing demands of consumers and the current market. Therefore, their sales are always not advantageous on e-commerce platforms, and many firms producing traditional products often need the government's financial or policy support for survival. We believe that these cultural treasures should not only exist in the content of cultural protection and museums; their products can also be of a modern style, fashionable, and cross-border designed. Thus, it not only

needs firms to transform their design, produce, and sale strategies, but also requires firms to co-create their products with the incorporation of consumers' knowledge and suggestions.

This study applies the aspect-based sentiment analysis to look into how consumers could be involved in influencing the development of the ceramic industry in China through open data platforms. There are three limitations regarding the study. First, the data noise cannot be fully wiped off, which is generated by accumulating a large number of user behaviors on the Internet. Thus, the accuracy could be improved in the future with more data correction and a manual screening process. Second, in the process of our cooperation with firm D, we only performed analysis and verification of 22 new products due to human and material constraints. The sample size of this part is relatively small, which might to some extent affect the judgment of consumers' feedback on new product performance. Third, tracking four-month new product research and observation, due to time constraints, might to some extent affect the sample size and the accuracy of consumer feedback on new products performance. In addition, we only successfully conducted a complete innovative cooperation with one firm in the field of cultural and creative industry with the inclusion of sentiment analysis due to research limitations. The applicability of this method to other companies and industries remains to be discussed in future research.

Author Contributions: Conceptualization, C.L.; methodology, C.L. and Y.X.; investigation, Y.X.; resources, C.L.; writing—original draft preparation, C.L.; writing—review and editing, Y.X.; visualization, Y.X.; supervision, C.L.; project administration, Y.X.; funding acquisition, C.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Shanghai Pujiang Program, grant number 2020PJC110.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data used in this study in part came from Taobao's public data, and in part from the data provided by firm D. Any questions or clarifications can be posed to: liuc@tongji.edu.cn.

Acknowledgments: This work was strongly supported by C.W., X.T. and F.W. from firm D in Longquan city, China.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. The 47th Statistical Report on Internet Development in China (in Chinese: Di 47 Ci Zhongguo Hulianwang Fazhan Zhuangkuang Tongji Baogao). Available online: http://www.cac.gov.cn/2021-02/03/c_1613923423079314.htm (accessed on 17 May 2021).
2. China's Digital Economy Development White Paper (in Chinese: Zhongguo Shuzi Jingji Fazhan Bai Pi Shu). Available online: http://www.caict.ac.cn/kxyj/qwfb/bps/202104/t20210423_374626.htm (accessed on 17 May 2021).
3. Lee, K.S.; Tan, S.J. E-retailing versus physical retailing: A theoretical model and empirical test of consumer choice. *J. Bus. Res.* **2003**, *56*, 877–885. [\[CrossRef\]](#)
4. Greer, C.R.; Lei, D. Collaborative innovation with customers: A review of the literature and suggestions for future research. *Int. J. Manag. Rev.* **2012**, *14*, 63–84. [\[CrossRef\]](#)
5. Zhang, H.; Xiao, Y. Customer involvement in big data analytics and its impact on B2B innovation. *Ind. Mark. Manag.* **2020**, *86*, 99–108. [\[CrossRef\]](#)
6. Liu, C. Political Power and the Industrial Development of Cultural Artifacts in China. *Lateral* **2020**. [\[CrossRef\]](#)
7. Bharadwaj, N.; Noble, C. Finding innovation in data rich environments. *J. Prod. Innov. Manag.* **2017**, *34*, 560–564. [\[CrossRef\]](#)
8. Liu, X.; Lin, Y.; Zhang, S.; Xie, Y. Frugal innovation in e-commerce: A case study of Taobao. In Proceedings of the 2016 IEEE International Conference on Management of Innovation and Technology (ICMIT), Bangkok, Thailand, 19–22 September 2016.
9. Mount, M.; Martinez, M.G. Social media: A tool for open innovation. *Calif. Manag. Rev.* **2014**, *56*, 124–143. [\[CrossRef\]](#)
10. Freixanet, J.; Braojos, J.; Rialp-Criado, A.; Rialp-Criado, J. Does international entrepreneurial orientation foster innovation performance? The mediating role of social media and open innovation. *Int. J. Entrep. Innov.* **2021**, *22*, 33–44.
11. Chesbrough, H.W. Bring open innovation to services. *MIT Sloan Manag. Rev.* **2011**, *52*, 85.
12. Yun, J.J.; Zhao, X.; Park, K.; Shi, L. Sustainability condition of open innovation: Dynamic growth of Alibaba from SME to large enterprise. *Sustainability* **2020**, *12*, 4379. [\[CrossRef\]](#)

13. Baldwin, C.; Von Hippel, E. Modeling a paradigm shift: From producer innovation to user and open collaborative innovation. *Organ. Sci.* **2011**, *22*, 1399–1417. [\[CrossRef\]](#)
14. Olson, E.; Bakke, G. Implementing the lead user method in a high technology firm: A longitudinal study of intentions versus actions. *J. Prod. Innov. Manag.* **2001**, *18*, 388–395. [\[CrossRef\]](#)
15. Von Hippel, E. Open user innovation. In *Handbook of the Economics of Innovation*, 1st ed.; Hall, B.H., Rosenberg, N., Eds.; North Holland: Amsterdam, The Netherlands, 2010; Volume 1, pp. 411–427.
16. Al-Hakim, L. Information quality factors affecting innovation process. *Int. J. Inf. Qual.* **2007**, *1*, 162–176.
17. Griffin, A.; Josephson, B.W.; Lilien, G.; Wiersema, F.; Bayus, B.; Chandy, R.; Dahan, E.; Gaskin, S.; Kohli, A.; Miller, C.; et al. Marketing's roles in innovation in business-to-business firms: Status, issues, and research agenda. *Mark. Lett.* **2013**, *24*, 323–337. [\[CrossRef\]](#)
18. Lilien, G.L. The B2B knowledge gap. *Int. J. Res. Mark.* **2016**, *33*, 543–556. [\[CrossRef\]](#)
19. Cambria, E.; Das, D.; Bandyopadhyay, S.; Feraco, A. Affective computing and sentiment analysis. In *A practical Guide to Sentiment Analysis*, 1st ed.; Cambria, E., Das, D., Bandyopadhyay, S., Feraco, A., Eds.; Springer: Cham, Switzerland, 2017.
20. Zeng, D.; Dai, Y.; Li, F.; Wang, J.; Sangaiah, A.K. Aspect based sentiment analysis by a linguistically regularized CNN with gated mechanism. *J. Intell. Fuzzy Syst.* **2019**, *36*, 3971–3980. [\[CrossRef\]](#)
21. Alessia, D.; Ferri, F.; Grifoni, P.; Guzzo, T. Approaches, tools and applications for sentiment analysis implementation. *Int. J. Comput. Appl.* **2015**, *125*, 3.
22. Liu, B. *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*, 2nd ed.; Springer: Berlin, Germany, 2011; pp. 136–198.
23. Do, H.H.; Prasad, P.W.; Maag, A.; Alsadoon, A. Deep learning for aspect-based sentiment analysis: A comparative review. *Expert Syst. Appl.* **2019**, *118*, 272–299. [\[CrossRef\]](#)
24. Peng, H.; Cambria, E.; Hussain, A. A review of sentiment analysis research in Chinese language. *Cognit. Comput.* **2017**, *9*, 423–435. [\[CrossRef\]](#)
25. Yang, L.; Li, Y.; Wang, J.; Sherratt, R.S. Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. *IEEE Access* **2020**, *8*, 23522–23530. [\[CrossRef\]](#)
26. Yang, L.; Wang, Y. Research of the Construction of Emotion Dictionary and Emotion Analysis in the Emotion Analysis of micro-blog (in Chinese: *Weibo Qinggan Fenxi de Qinggan Cidian Gouzao ji Fenxi Fangfa Yanjiu*). *Comp. Tech. Dev.* **2019**, *2*, 1–6.
27. Xing, B. Electronic mall shopping system based on Jieba word segmentation search and SSM framework (in Chinese: *Jiyu Jieba Fenci Sousuo yu SSM Kuangjia de Dianzi Shangcheng Gouwu Xitong*). *China Comp. Commun.* **2018**, *7*, 104–105.
28. Cui, G.; Lui, H.; Guo, X. The effect of online consumer reviews on new product sales. *Int. J. Electron Commer.* **2014**, *17*, 39–58. [\[CrossRef\]](#)
29. Duan, W.; Gu, B.; Whinston, A.B. Do online reviews matter? An empirical investigation of panel data. *Decis. Support Syst.* **2008**, *45*, 1007–1016. [\[CrossRef\]](#)
30. Park, D.H.; Lee, J.; Han, I. The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement. *Int. J. Electron. Commer.* **2007**, *11*, 125–148. [\[CrossRef\]](#)
31. Skaggs, B.C.; Youndt, M. Strategic positioning, human capital, and performance in service organizations: A customer interaction approach. *Strateg. Manag. J.* **2004**, *25*, 85–99. [\[CrossRef\]](#)
32. Mu, Z.; Zheng, Y.; Sun, H. Cooperative green technology innovation of an e-commerce sales channel in a two-stage supply chain. *Sustainability* **2021**, *13*, 7499. [\[CrossRef\]](#)
33. Franke, N.; Piller, F. Value creation by toolkits for user innovation and design: The case of the watch market. *J. Prod. Innov. Manag.* **2004**, *21*, 401–415. [\[CrossRef\]](#)
34. Chesbrough, H.W. *Open Innovation: The New Imperative for Creating and Profiting from Technology*; Harvard Business Press: Boston, MA, USA, 2003.
35. Bogers, M.; Afuah, A.; Bastian, B. Users as innovators: A review, critique, and future research directions. *J. Manag.* **2010**, *36*, 857–875. [\[CrossRef\]](#)