Customer Review Analysis of Online E-commerce Platforms
A Configurational Approach

Mezei, József; Davoodi, Laleh; Nikou, S.

Publication date
2024

Document Version
Final published version

Published in
Proceedings of the 57th Hawaii International Conference on System Sciences

Customer Review Analysis of Online E-commerce Platforms — A Configurational Approach

József Mezei
Faculty of Social Sciences,
Business and Economics, and Law
Åbo Akademi University
Jozsef.Mezei@abo.fi

Laleh Davoodi
Faculty of Social Sciences,
Business and Economics, and Law
Åbo Akademi University
Laleh.Davoodi@abo.fi

Shahrokh Nikou
Department of Design, Organisation
and Strategy, TU Delft
Department of Computer and
Systems Sciences, Stockholm University
S.N.Nikou@tudelft.nl

Abstract

In order to understand the antecedents of customer satisfaction, businesses can analytically utilise the growing amount of customer information. Unstructured text data can be used to uncover important information owing to developments in Natural Language Processing and text analytics approaches. In this paper, we focus on customer reviews posted on e-commerce shopping platforms. We perform manual data annotation to determine the sentiment of the review with respect to the most important aspects of the customer journey. The 14 extracted aspects are grouped into three categories that correspond to the stages of the customer's interaction with the e-commerce platform. We make use of a configurational approach, Fuzzy-set Qualitative Comparative Analysis, to understand how the sentiment with respect to the three stages combines to achieve positive customer satisfaction. The outcomes of the analysis show that all three stages of the customer journey play important roles in determining the final evaluation of a customer, leading to a positive or negative sentiment. The theoretical and practical implications are discussed.

Keywords: Customer journey, Customer reviews, e-commerce, FsQCA, Sentiment analysis

1. Introduction

While there is a rapid growth of online shopping [9], e-commerce and online platform owners are looking into how to transform the customers' shopping experience into a positive customer journey. According to [17], business-to-consumer (B2C) e-commerce refers to any businesses that provide goods, products, and services to general consumers over the Internet for their individual use. By the advancement of computer technologies, the amount of textual data available on the web is increasing continuously. Therefore, businesses have access to diverse sources of customer reviews and feedback that can be used to enhance decision-making processes and customer satisfaction. Artificial Intelligence (AI) has been shown to be one of the most efficient tools to analyse vast amounts of textual data. Recent literature shows that AI-based systems are transforming marketing decisions in the service industry and have great potential in customer experience improvement [3]. Understanding customer behaviour and what drives customer satisfaction has a major impact on a business's success not only in creating improved value creation but also gaining a competitive advantage in the market. Natural Language Processing (NLP), more specifically sentiment analysis techniques, help businesses to gauge customer's feelings and understand their requirements and subsequently generate more revenue and increase productivity.

To this end, the current research aims to evaluate and analysis customer reviews on online B2C shopping platforms given at different stages of the online shopping journey to understand what aspect(s) of online shopping leads to an overall positive or negative sentiment and increased customer satisfaction. Based on the literature and exploratory analysis of 3500 reviews, we identified 14 aspects (e.g. customer service, payment, app experience, and shipping), which cover all the important dimensions of customers' shopping online experience [16]. To extend the limited existing research on the relationship between online customer journey and customer satisfaction, we aim to answer the following research question (RQ): "How can we utilize online customer reviews to understand the impact of different stages of the online shopping customer journey on

URI: https://hdl.handle.net/10125/106564
978-0-9981331-7-1

customer satisfaction?". We answer the RQ by making use of a configurational thinking approach, namely Fuzzy-set Qualitative Comparative Analysis [21] on a dataset composed of 3500 online reviews of e-commerce platforms.

This research provides twofold contributions. First, we theoretically contribute to the literature on customer experience by scrutinising different stages of the customer journey model [16], and show that different aspects must be taken collectively into consideration to understand if the customer experience leads to high sentiment scores (positive evaluations) or a low sentiment scores (negative evaluations). Second, the findings of this research show how online reviews and comments provided by customers on online shopping platforms can be evaluated by making use of NLP and analysed through configurational thinking.

This paper is organised as follows. Section 2 discusses the literature review and briefly discusses the e-commerce customer journey, in particular as approached with a configurational analysis tool. Next, research methodology, data acquisition, preparation, and annotation are presented in Section 3. Section 4 provides the main results of the FsQCA analysis. Section 5 offers a discussion by elaborating on the results, and, finally, Section 6 concludes the research.

2. Literature review

In the rapidly expanding e-commerce sector, understanding the customer journey is critical for creating positive online experiences. To gain a comprehensive overview of customers' online shopping experiences, e-commerce websites make use of online evaluation criteria systems to understand customer experience over time. Evaluation systems offer valuable insights into these experiences, highlighting the need for platforms to integrate functions like IT, service operations, and payment systems for sustained customer satisfaction [16].

2.1. E-commerce customer journey

Despite the widely acknowledged benefits of e-commerce (e.g. time and cost savings), understanding the experience of customers when they shop online is one of the most crucial factors [2]. As such, many online shopping platforms offer forums where users can express and discuss their shopping experience. These discussions concern factors such as cost, shipping, packaging, payment, and delivery, and how these components of their buying experience meet expectations and impact their online shopping journey [1, 6]. Previous studies have explored the subject of customers' reviews using the comments and feedback they shared. For example, [28], applied machine learning techniques on customer reviews and sentiments to assess factors influencing the purchasing decision. The authors found that customers' buying intentions and seller characteristics are positively correlated.

In general, e-commerce customer experience can be assessed and examined through several individual aspects such as features of the website, the payment system, and the quality of shipping or app experience. As such, many authors have attempted to categorise these individual aspects into the higher-level of classifications that together impact e-commerce customers' shopping experience and their decision to provide a positive or a negative evaluation. For example, [6] in their research identified six categories of e-commerce website evaluation. In contrast, [4] has introduced three overarching categories as: (1) outlook, (2), operation, and (3) service. However, one of the most widely used conceptualisation of online customer experience is the model developed by [16]. The authors conceptualised the process model for the customer journey and experience into three categories: (i) pre-purchase stage, purchase stage, and (3) post-purchase stage. For example, the pre-purchase stage covers customer experience on e.g. search and need recognition, or post-purchase covers customer experience on e.g. consumption and usage. Building on [16], the current research has identified 14 general aspects, and has grouped them into three stages: (i) Stage 1: App experience, information, product availability, product features, pricing, and payment (ii) Stage 2: shipping, delivered product status, packaging, and item quality, and (iii) Stage 3: refund process, return process, trust, and customer service.

2.2. FsQCA in e-commerce research

While there is extensive academic research on different stages of customer journey in e-commerce and how they contribute to customer satisfaction, most of the existing studies apply traditional statistics-based methodologies. As it was pointed out recently by [15], there is a definite need for applying more advanced rule/configuration-based models in customer satisfaction research. In particular, the authors single out Fuzzy-set Qualitative Comparative Analysis (FsQCA) as a method with great promise in generating novel insights by focusing on the different combinations of antecedents of customer satisfaction, instead of an aggregation process. FsQCA [21] has become an increasingly used method in information systems research with applications in domains ranging from entrepreneurship to technology adoption. However, one particular area of interest that has not seen a large number of contributions is e-commerce. To the best of our knowledge, this article is the first that considers the complete purchase process in a model and makes use of a configurational approach to understand the most important shopping aspects leading to a satisfied customer.

One of the articles looking at this issue, albeit with a more limited scope, is the work by [7], in which the influence of motivations and barriers on online-shopping behaviour is analysed. By making use of survey data and FsQCA, the article combines seven motivator and three barrier variables. As the authors found, online shops should keep tabs on customers continuously to quickly adjust to changes, and an important tool for companies to achieve this goal is to make use of user-generated content in real-time, in particular online customer reviews. In another work, [19] utilises FsQCA to understand what combinations of variables promote online shopping behaviour. By combining cognitive and affective perceptions to explain high intention to purchase in an e-commerce environment, the authors identify nine distinct behavioural patterns that characterise different ways to achieve increased use of the services.

3. Methodology

In this section, we will present the research methodology and the dataset used in the analysis. We start with explaining the basic concepts of aspect-based sentiment analysis, and how our data was prepared. We present the selected aspects as the most important components of the e-commerce shopping process. Finally, we discuss the data analysis methodology used in the paper.

3.1. Aspect based sentiment analysis

As we have discussed, customer reviews can offer important insights to businesses to improve customer satisfaction. However, it is difficult to manually analyse large quantities of textual data. Hence, it is necessary to use NLP tools and techniques such as sentiment analysis (SA) to process text data. The goal of SA (sometimes termed as opinion mining) is to categorise users' emotions or opinions, and it has been widely used in many applications in business. SA is typically performed on three levels [12]: (1) Document level, (2) Sentence level, and (3) Entity and Aspect level. Aspect-based sentiment analysis (ABSA) seeks to identify the sentiment score towards a particular aspect in a text [14]. The users' comments may contain different aspects: *"Great quality items, fast shipping, 100% recommend Boozt"*. There are three aspect terms *"Item Quality", "Shipping ", and "Trust"* in this review, and all are associated with positive sentiments respectively. The results of ABSA allow businesses to acquire a better insight of their services.

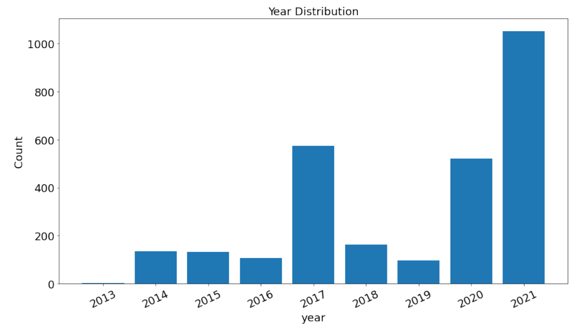

Figure 1. The distribution of reviews by year

3.2. Data acquisition and data annotation

To create our dataset, we retrieved English reviews from Trustpilot, a leading online review platform, in the time period of 2013-2021 (see Figure 1). For this study, we collected data from five online stores that did not have physical shops at the time of collecting the data. The online stores included in this research are the following: Zalando (974), Wish (673), Sheinside (543), Boozt (453), and Nelly (139). The reviewers are from 73 different countries. Britain (UK) with 1166 reviews, and the United States with 778 reviews, have the most share of reviews. Additionally to the text of the review, we also collected the additional numerical rating (on the scale from 1 to 5), provided by the customers.

In order to perform aspect extraction and manual data annotation, 3500 reviews were randomly selected. Thereafter, the text was cleaned and pre-processed for further analysis. We performed three steps to identify frequently mentioned candidate aspects in the reviews before doing manual annotation. We aimed to extract both previously identified aspects of the literature and aspects mentioned only in our corpus. First, we applied part-of-speech tagging for all the reviews. Afterwards, five bi-term combination types were constructed as suggested by [13]: (Noun, Adjective), (Noun, Verb), (Noun, Adverb), (Verb, Adjective), and (Verb, Adverb). Eventually, the aspects were chosen by counting the most frequent bi-term combinations e.g. *"Customer Service"* appeared directly 344 times in the reviews.

Table 1. The list of aspects categories and definitions.

	Aspect	Definition
1	Shipping	Quality of the delivery e.g. cost and timeline
2	Trust	Customers' general opinion about the store
3	Item quality	Products' quality
4	Customer service	Quality of customer's direct interaction with store's representative
5	Pricing	Price offerings, availability of discounts and campaigns
6	Product features	Quality of product's image and size guide provided on the website
7	Refund process	Refund speed and quality of handling refund issues
8	Return process	Speed, convenience, and cost of the return
9	App experience	User experience in interacting with the store's website
10	Delivered product status	Condition of delivered products e.g. broken, smelling
11	Information	Availability and quality of the information e.g. misleading ads
12	Packaging	Attractiveness and quality of the packaging
13	Payment	Quality of financial transaction
14	Product availability	Variety of offering products or brands

3.3. The final set of aspects

The final set of aspects to be used in our research was determined based on the presented exploratory analysis of the data and existing research on presenting important concepts related to a customer's interaction with e-commerce services [11, 23]. The aspects are listed in Table 1. Afterwards, we manually annotated 300 further reviews in order to cross-check the quality of the chosen aspects.

At the beginning of the purchase process, the customer starts to interact with the e-commerce platform and gathers information about products. Information (richness), as used; for example, by [24], refers to the availability (or lack) of the information that can help customers in making more informed purchase decisions; in turn, a more informed and better decision may result in a more satisfied customer. Product availability and product features [11] refer to a sufficiently wide selection and availability of product offerings and brands available on the platform, including different variations of products with varying features, e.g. different sizes and colours of clothes, etc. Pricing [23] as used in this analysis, concerns the perception of any information related to price offerings, including discounts and campaigns, and general price level.

Many different factors contribute to the general user experience when interacting with e-commerce platforms. This also includes factors such as website/application aesthetics, design, and convenience [5]. We term this aspect as App experience, and it is used when the review mentions related positive or negative issues. The last step of ordering a product is payment when a financial transaction is completed [5, 23]. While in previous research this was only considered as part of a general purchase process construct, based on the high frequency of the reviews mentioning specifically payment, we included it as a separate aspect.

The second stage of the process mainly concerns the delivery of the ordered products. The most frequently occurring aspect in the whole dataset is the core activity of this stage, shipping [11]. This includes any observations in the reviews related to the cost and timeliness of the delivery, or information about the package (order tracking). Based on the frequency of mentions, we separately considered specific components of shipping, namely packaging, that captures the attractiveness and quality of the packaging in which the product is delivered and delivered product status (termed as the delivery condition by [11]), that refers to the condition in which the products are delivered (e.g. broken, missing, etc.). After the ordered products are delivered, the customer evaluates the received product, i.e. assesses Item quality [23] with respect to expectations based on the information that was provided when the product was ordered.

If the customer is not satisfied with the received items, she/he can make use of the return process and request a refund of the purchase price [18]. Return is an essential component of e-commerce, and is one of the most significant differences compared to traditional commerce. Refund, in particular, and specifically the refund speed and effective handling of problems related to the refund through the e-commerce platform has also been found to play an important role in customer experience. An important aspect for many customers, as also highlighted by the frequent mentions, is customer service, i.e., the quality of engaging directly with the company, focusing on their competence, courtesy, and efficiency. Finally, we included an aspect we termed as trust, that incorporates any issues that are related to the possible future use of the e-commerce platform and the customers' general opinion on the e-commerce platform, e.g. recommendations and word-of-mouth [23].

In this research, the goal is to understand how different aspects of customers' interaction with an online e-commerce platform impact their satisfaction, as captured by the sentiment of the review. While each of the considered aspects offers valuable insights into an important dimension of the customer journey, in the analysis we have formed three higher-level categories corresponding to three main stages of a customer's interaction with the e-commerce platform:

- Stage 1: includes the six aspects (i) App experience, (ii) information, (iii) product availability, (iv)product features, (v) pricing, and (vi) payment. This stage captures every component of the process starting with the customer's search for information about products until the ordering and payment.

- Stage 2: includes the 4 aspects (i) shipping, (ii) delivered product status, (iii) packaging, and

(iv) item quality. This stage focuses on the delivery process and the customer's evaluation of the received product.

- Stage 3: includes the 4 aspects (i) refund process, (ii) return process, (iii) customer service, and (iv) trust. This stage measures the customer's possible interactions with the company after the order is received, including any indication of the opinion they formed on the future use of the platform's services.

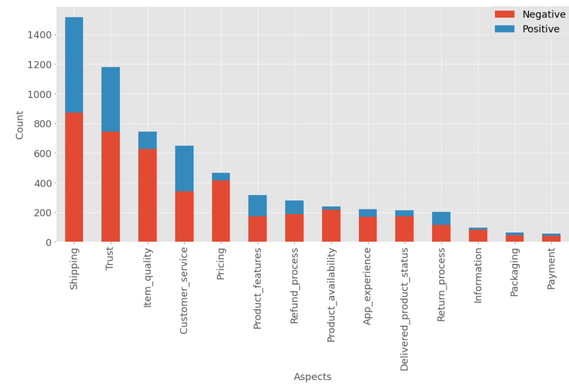

Figure 2. The distribution of 14 aspect categories

3.3.1. Data annotation To perform data annotation, we assigned a polarity to each aspect-review pair according to the emotion expressed in a review toward each particular aspect. Moreover, we defined the overall sentiment of the reviews individually. Two annotators (authors of this article) individually annotated each file and assigned a sentiment to each aspect and overall sentiment of each review. To mark the aspect's polarity, we employed two possible labels for each aspect: positive, and negative. However, to identify the overall sentiment of each review we used three labels to mark their sentiment: positive, negative, and mixed. After annotating each file, every single disagreement was discussed to reach the final agreement.

After performing manual annotation, we observed that there are a large number of disagreements between the rating (defined by the reviewer) and the actual sentiment of the review (assigned by the annotators). This indicates that the ratings may not be suitable to be utilised as a fully correct indication of overall customer sentiment[8]. Therefore in the analysis, we combined the user rating (provided with the review) and the overall sentiment value (assigned in the annotation process) as the outcome of interest. Finally, we selected reviews with at least one mentioned aspect. As presented in Figure 2 on the aspect level, the final dataset holds 2782 reviews, consisting of 3850 positive, and 2387 negative aspect polarities.

3.4. Fuzzy-set qualitative comparative analysis

Most of the existing research on customer satisfaction and its potential antecedents association makes use of different variants of regression analysis. There are three main issues with such techniques [20]: (i) the interaction among the variability of the relationship between variables, (ii) modeling only symmetric relationships, and (iii) the inability to capture non-linear relations among variables. In order to address these issues, a frequently used alternative is to use methods capturing various configurations of variables, with one prominent method being FsQCA [21]. FsQCA can help in establishing causal relationships among an outcome of interest and its hypothesised antecedents. In this paper, we have followed the general principles of performing FsQCA analysis as presented in [20]. As the first step of the analysis, the variables of the model need to be determined. As it was described, the data was annotated by specifying a sentiment for each aspect (positive or negative) and the overall sentiment (positive, negative, or mixed).

As the outcome variable of the analysis, we aim to capture the general satisfaction level of the customer. We combined two measurements for this: (i) the general sentiment assigned in the annotation process, and (ii) the numeric rating provided by the customer when creating the review. The general sentiment was coded as 0, 0.5, and 1 for negative, mixed, and positive sentiment, respectively. The user rating, originally an integer, was converted to the unit interval by assigning to the rating 1, 2, 3, 4, and 5, the values 0, 0.25, 0.5, 0.75, and 1, respectively. Finally, the variable that represents the outcome of interest in our analysis was computed as the average of the general sentiment and the customer's rating.

Additionally, we specified six condition variables, two for each of the three stages defined above: (i) a variable capturing whether there is any positive sentiment towards the stage mentioned in the review, and (ii) a variable capturing whether there is any negative sentiment towards the stage. For example, we will have two variables corresponding to 'Stage 1': 'Stage 1 Positive' and 'Stage 1 Negative'. The value for 'Stage 1 Positive' is then determined as: (i) 1, if at least one of the six constituent aspects is mentioned positively in the review, and (ii) 0 otherwise, i.e. when none of those aspects are mentioned positively. In contrast, 'Stage 1 Negative' is determined as: (i) 1, if at least one of the constituent aspects is mentioned negatively in the review, and (ii) 0 otherwise, i.e. when none of those aspects are mentioned negatively. This procedure resulted in a final dataset of 2292 reviews. In this study, the calibration process is straightforward for all the variables, as they take on values between 0 and 1. Furthermore, it is important to note that to justify the construction of the variables that, e.g. 'Stage 1 Positive' and 'Stage 1 Negative' are not simply the opposite of each other. As these sentiment scores combine valuations with respect to multiple aspects, it may occur that some of those aspects are valued positively and others negatively by the customer. For example, if the customer is satisfied with product availability but has problems with the payment process, both 'Stage 1 Positive' and 'Stage 1 Negative' variables will take on the value 1. It may also happen that the review does not make any observations about 'Stage 1', in which case both of the variables will have the value 0.

Table 2. Consistency of necessary conditions for both high and low values of sentiment (coverage value in parenthesis)

Conditions	High	Low
Stage 1 Positive	0.42 (0.93)	0.10 (0.07)
not Stage 1 Positive	0.57 (0.65)	0.90 (0.35)
Stage 1 Negative	0.45 (0.76)	0.43 (0.24)
not Stage 1 Negative	0.55 (0.74)	0.58 (0.26)
Stage 2 Positive	0.67 (0.95)	0.10 (0.05)
not Stage 2 Positive	0.33 (0.51)	0.90 (0.49)
Stage 2 Negative	0.72 (0.71)	0.87 (0.29)
not Stage 2 Negative	0.28 (0.87)	0.13 (0.14)
Stage 3 Positive	0.56 (0.94)	0.11 (0.06)
not Stage 3 Positive	0.44 (0.59)	0.89 (0.41)
Stage 3 Negative	0.57 (0.66)	0.88 (0.35)
not Stage 3 Negative	0.43 (0.91)	0.12 (0.09)

4. Results

In this section, the results of the FsQCA analysis are presented. We present configurations of conditions that consistently lead to either high sentiment scores (positive evaluations) or low sentiment scores (negative evaluations). We start by checking whether there are any individual conditions that are necessary for the outcome to be realised. Then, we determine configurations that are sufficient to reach either the outcome or its negation. The data collection and pre-processing of the review texts have been performed with various libraries of the Python programming language, while all the computations related to the use of FsQCA were performed with various packages of the statistical programming language R.

4.1. Necessity analysis

An important step in performing FsQCA is to run a necessity analysis and identify possible necessary conditions. For a condition to be necessary means that the outcome can almost never occur (in a fuzzy sense) without the condition being true. The results of the analysis are presented in Table 2. The calculations are performed for the presence (high values) and absence (low values) of the outcome. To identify which conditions are necessary and how relevant this relationship is, the measures of consistency and coverage are accounted for. In order to determine whether a variable is a necessary condition, consistency, and coverage measures are calculated. To estimate the strength of the relationship, consistency is calculated as a value between 0 and 1; as suggested by [22], values above 0.9 may indicate the presence of a necessary condition. To complement consistency by measuring the importance of the relations, coverage can be calculated; the higher the value of coverage is, the larger is the number of applicable cases.

For high (positive) sentiment scores, i.e. high levels of the outcome, the results show no necessary condition. This implies that we cannot single out any stage of the customer journey process as clearly more crucial than the others. In other words, there are many positive reviews that do not mention, e.g. 'Stage 1' positively. The highest consistency values can be observed with the two variables related to 'Stage 2', showing the importance of shipping and other related aspects in the customer journey. In contrast, we can find several variables reaching the 0.9 threshold value (or being very close to it) for low (negative) sentiment scores. The two variables that have a consistency value of exactly 0.9 are 'not Stage 1 Positive' and 'not Stage 2 Positive'. This implies that for a significant part of the data (0.49 and 0.35 coverage values, respectively), negative sentiment can only occur if the customers do not perceive any aspects related to these two categories positively. Additionally, the similar variable for the third category, 'not Stage 3 Positive', also achieved consistency 0.89, just slightly below the threshold, resulting in the same observation, i.e. there cannot be any positively rated aspects belonging to this stage in overall negative reviews.

4.2. Sufficiency analysis

In the following, we will present the results of the sufficiency analysis. Tables 3 and 4 show five sufficient configurations for both the presence and the absence of the outcome variable.

4.2.1. Sufficient configurations for high sentiment

The configurations for high (positive) sentiment can be seen in Table 3. It is important to notice that the 'Stage 1' (either Positive or Negative) category appears in all five solutions, while 'Stage 2' appears in four solutions. In contrast, the 'Stage 3' category appears in only two solutions. This general observation is indicative of 'Stage 3', i.e. the last stage of the e-commerce customer journey being less important in achieving a high level of sentiment, and in turn satisfied customers. This is highlighted by Solution 1, according to which a positive evaluation of the first two stages is sufficient to achieve high sentiment. Moreover, this is the solution with the highest coverage value, 0.30, meaning that it explains the sentiment of more reviews than any of the other solutions.

Solutions 2 and 3 present even stronger versions of Solution 1, albeit with lower coverage values, i.e. explaining fewer review sentiments. While Solution 1 states that a positive evaluation on both 'Stage 1' and 'Stage 2' is sufficient to achieve positive sentiment, according to Solution 2, it is sufficient to have only a positive evaluation on 'Stage 1' while not having any negative related to 'Stage 2'. Furthermore, Solution 3 states that not having any negative sentiment related to any aspects belonging to 'Stage 1' and 'Stage 2' is sufficient to achieve a positive overall sentiment. These three solutions highlight the above-mentioned importance of these two stages, as none of these configurations involve 'Stage 3', and their total combined coverage value is close to 0.6, i.e. explaining almost 60% of the reviews. In other words, these solutions state that for many customers, it does not matter what happens in the process after they receive the product ordered from the e-commerce platform. If everything goes smoothly without any problem until that point, the overall evaluation will almost always be positive.

The last two solutions in Table 3 include also 'Stage 3'. According to Solution 4, to achieve high sentiment, it is sufficient that the review has nothing negative about 'Stage 1' and 'Stage 3'. This solution characterises customers for whom it is not important what happens in 'Stage 2': if nothing goes wrong in the ordering process and with the services after the order has arrived, issues related to delivery will not generate negative overall sentiments for these customers. Finally, according to Solution 5, if the review positively mentions 'Stage 2' and 'Stage 3' and does not contain anything negative about 'Stage 1', the general sentiment will be positive. This highlights the importance of 'Stage 1': it is not sufficient to achieve positive sentiments just by having a positive evaluation of the last two stages of shopping, if a problem happens related to aspects in 'Stage 1', the overall sentiment is not guaranteed to be positive.

4.2.2. Sufficient configurations for low sentiment

The configurations for low (negative) sentiment can be seen in Table 4. When looking at the results, in particular the coverage values, we can see that Solutions 2 and 3 in Table 4 relate to a very low number of reviews. While in FsQCA analysis, we are not concerned in general with statistical significance, these numbers are very low considering general practice and comparable studies, meaning that these rules most likely express some patterns that are specific to the considered dataset. Future research with larger datasets or from different sources could approve or disprove the relevance of these configurations, as such in the discussion we focus on the remaining three solutions.

Solution 1 offers a quite straightforward path to negative sentiment: if there is no single aspect that the customer values positively, then the overall evaluation will be negative. According to the specifications of the dataset used, as there is no positively valued aspect, some of the aspects must have been mentioned negatively in the review, as we did not include completely neutral evaluations. In other words, this solution states that if there is a component of the purchase process that is not working, and there is no compensation in the form of at least one other positive aspect, the overall evaluation will be negative.

Solutions 4 and 5 highlight the importance of 'Stage 3' in affecting the general sentiment of customers. In both of these configurations, it is required that the reviews contain a negative evaluation towards at least one of the aspects belonging to this stage, e.g. refund and customer service, and there is no positive evaluation towards any of the related aspects. In addition to this, if either there is no positive statement about 'Stage 2' (Solution 4) or if there is something negative about 'Stage 1', the overall sentiment will be negative. In other words, while a negative evaluation of 'Stage 3' alone is not sufficient to reach a negative outcome, if it occurs with some problems related to at least one of the other two stages, that is enough to make the customer dissatisfied.

Table 3. Configurations for positive sentiment

Conditions	Solution 1	Solution 2	Solution 3	Solution 4	Solution 5
Stage 1 Positive	●	●			
Stage 1 Negative			○	○	○
Stage 2 Positive	●				●
Stage 2 Negative		○	○		
Stage 3 Positive					●
Stage 3 Negative				○	
Consistency	0.96	0.94	0.96	0.96	0.98
Coverage	0.30	0.11	0.16	0.23	0.13

Table 4. Configurations for negative sentiment

Conditions	Solution 1	Solution 2	Solution 3	Solution 4	Solution 5
Stage 1 Positive	○	○	●		
Stage 1 Negative					●
Stage 2 Positive	○	○		○	
Stage 2 Negative					
Stage 3 Positive	○		○	○	○
Stage 3 Negative		○	●	●	●
Consistency	0.96	0.86	0.81	0.97	0.94
Coverage	0.77	0.05	0.03	0.75	0.29

5. Discussion

In this article, we examined the process of a customer journey towards e-commerce platforms from searching for a product on the platform to providing a review. Based on the literature and empirical analysis of data from a leading review aggregator site, we identified a set of 14 aspects that capture the most important components of processes of end-to-end customer experience and their online shopping journey. Furthermore, we specified three higher-level stages of the process that aggregate related aspects together. The reviews used in the analysis were manually annotated to determine customer sentiment with respect to each of the 14 aspects, and these values were combined to obtain a sentiment for the three online shopping stages. Based on the sentiment scores that correspond to whether the customer was satisfied or not with the different aspects during online shopping, we propose to make use of a configurational approach to explain the overall sentiment of the reviews.

We identified configurations of positive and negative sentiments towards the three stages of the customer journey that are sufficient to achieve either a positive or negative overall sentiment. While the results of the annotation process have already shown that the individual aspects considered in the research are at the core of a customer's evaluation, the configurational analysis performed with FsQCA complements this observation and has shown that the three conceptualised stages play important roles, albeit in a different way.

As the solutions in Table 3 show, having a positive evaluation on the first two stages of the customer journey process is sufficient for most of the customers to be satisfied with the process as a whole. We may state that these functions play the role of a motivator, as good performance from the company on related aspects will likely make the customer satisfied, and continue using the platform. In addition, the configurations in Table 4 show that the third stage of the process (e.g. refund, return, customer service) plays an equally important role, but acts as an inhibitor. While positive evaluations about this stage in most cases do not result in a positive sentiment, a negative evaluation in combination with any issue at other stages in the customer journey process will result in a negative sentiment. The findings of the analysis provide a deeper understanding of the roles that different stages of the customer journey play in the satisfaction of customers. Our results indicate that companies need to pay attention to all stages of the process, but also focus on the individual stage according to the results. As the results show, the performance of the company in tasks related to the last stage of the shopping process is crucial as negative perception can easily result in an overall negative evaluation.

Our findings echo the broader literature on customer journeys in e-commerce settings, which have also emphasized the importance of different stages in shaping customer satisfaction [16, 25]. However, unlike previous studies that have often focused on isolated aspects, our configurational approach allows for a more nuanced understanding of how these aspects interact to influence overall sentiment. While there is a large number of studies focusing on the initial stages of the customer journey, our study highlights the pivotal role of the final stage, similarly to [26]. Negative evaluations at this stage, as shown in our results, can substantially affect overall customer sentiment, a finding that has not been sufficiently addressed in existing studies. Furthermore, the use of FsQCA in our research offers a complementary perspective to existing quantitative approaches commonly found in the literature. While studies based on correlation-analysis often quantify the impact of individual aspects, our configurational approach provides insights into the complex interplay between different stages, contributing a new layer of understanding to the field.

6. Conclusions

This research has several implications towards understanding the antecedents of customer satisfaction in e-commerce. First, we extended existing research on the relationship between different stages of the e-commerce customer journey process and customer satisfaction by looking at how different stages of the process combine together to result in positive or negative sentiments. Second, the current research is one of the first studies in e-commerce research that combines user-generated content in the form of customer reviews and makes use of configurational thinking to shed new light on the antecedents of customer sentiment. The results of the data annotation and analysis confirm the role of many important aspects frequently mentioned in the literature. Overall, our results show practical implications of the relevance of taking a holistic approach to understanding customers' interaction with e-commerce platforms. Our results may help e-commerce platform owners to refine their strategies and services by focusing on the core activities that enhance interactions with the customers. The multiple solutions identified in the analysis provide alternative paths for businesses to ensure positive customer sentiment and satisfaction in order for them to continue using the services provided by the company. To summarise the main contributions, as the main novelty for academic research, the present study showcases the application of FSQCA in e-commerce research, a method that allows for a complex interplay of variables, and takes into consideration (and validates) multiple stages of the purchase process. Regarding the managerial implications, the results offer actionable insights for e-commerce platforms looking to enhance customer interactions. Platform owners and vendors can employ the findings to refine their customer engagement strategies, focusing on variables that have been shown to contribute significantly to customer satisfaction and engagement.

We also need to acknowledge the limitations related to the current research. First, the reviews used in this paper may only reflect the evaluation for only the users of TrustPilot. Further research could focus on different data sources from different regions. Second, a more extensive systematic literature review combined with a larger dataset may uncover a wider set of aspects that need to be incorporated as important possible antecedents of customer satisfaction. Additionally, as FsQCA does not quantify the individual role of different dependent variables, future studies may combine configurational analysis with traditional statistical techniques to further confirm the significant role of various aspects of customer reviews. Such an approach would not only validate the findings of this study but also provide a nuanced understanding of the individual roles of different variables in customer interactions on e-commerce platforms. One significant aspect not considered in this study is the role of fake reviews in customer interactions on e-commerce platforms [10, 27]. These fake evaluations can distort the overall sentiment. While platforms like Trustpilot employ fraud detection systems, the complete elimination of fake reviews remains a challenge. Future research could investigate how the presence of fake reviews could potentially skew the findings presented in this paper.

References

[1] Agarwal, R., & Venkatesh, V. (2002). Assessing a firm's web presence: A heuristic evaluation procedure for the measurement of usability. *Information systems research*, *13*(2), 168–186.

[2] Ahmed, S. Y., Ali, B. J., & Top, C. (2021). Understanding the impact of trust, perceived risk, and perceived technology on the online shopping intentions: Case study in kurdistan region of iraq. *Journal of Contemporary Issues in Business and Government*, *27*(3), 2136–2153.

[3] Aldunate, Á., Maldonado, S., Vairetti, C., & Armelini, G. (2022). Understanding customer satisfaction via deep learning and natural language processing. *Expert Systems with Applications*, *209*, 118309.

[4] Arora, M. (2016). Selection of parameters of e-commerce websites using ahp. *Proceeding of 2nd Conference on Innovative Practices*

in Information Technology & Operations Management, 22–27.

[5] Blut, M. (2016). E-service quality: Development of a hierarchical model. *Journal of Retailing*, *92*(4), 500–517.

[6] Chae, M., Kim, J., Kim, H., & Ryu, H. (2002). Information quality for mobile internet services: A theoretical model with empirical validation. *Electronic markets*, *12*(1), 38–46.

[7] Chaparro-Peláez, J., Agudo-Peregrina, Á. F., & Pascual-Miguel, F. J. (2016). Conjoint analysis of drivers and inhibitors of e-commerce adoption. *Journal of Business Research*, *69*(4), 1277–1282.

[8] Davoodi, L., & Mezei, J. (2022). A comparative study of machine learning models for sentiment analysis: Customer reviews of e-commerce platforms. *35th Bled eConference Digital Restructuring and Human (Re) action*.

[9] Guthrie, C., Fosso-Wamba, S., & Arnaud, J. B. (2021). Online consumer resilience during a pandemic: An exploratory study of e-commerce behavior before, during and after a covid-19 lockdown. *Journal of Retailing and Consumer Services*, *61*, 102570.

[10] Hajek, P., Hikkerova, L., & Sahut, J.-M. (2023). Fake review detection in e-commerce platforms using aspect-based sentiment analysis. *Journal of Business Research*, *167*, 114143.

[11] Holloway, B. B., & Beatty, S. E. (2008). Satisfiers and dissatisfiers in the online environment: A critical incident assessment. *Journal of service research*, *10*(4), 347–364.

[12] Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, 168–177.

[13] Im, J., Song, T., Lee, Y., & Kim, J. (2019). Confirmatory aspect-based opinion mining processes. *arXiv preprint arXiv:1907.12850*.

[14] Karimi, A., Rossi, L., & Prati, A. (2021). Adversarial training for aspect-based sentiment analysis with bert. *2020 25th International Conference on Pattern Recognition (ICPR)*, 8797–8803.

[15] Krassadaki, E., Grigoroudis, E., & Zopounidis, C. (2021). Advanced rule-based approaches in customer satisfaction analysis: Recent development and future prospects of fsqca. In *Euro working group on dss* (pp. 345–378). Springer.

[16] Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, *80*(6), 69–96.

[17] Mangiaracina, R., Marchet, G., Perotti, S., & Tumino, A. (2015). A review of the environmental implications of b2c e-commerce: A logistics perspective. *International Journal of Physical Distribution & Logistics Management*.

[18] Martinez-López, F. J., Feng, C., Li, Y., & López-López, D. (2022). Using instant refunds to improve online return experiences. *Journal of Retailing and Consumer Services*, *68*, 103067.

[19] Pappas, I. O., Kourouthanassis, P. E., Giannakos, M. N., & Chrissikopoulos, V. (2016). Explaining online shopping behavior with fsqca: The role of cognitive and affective perceptions. *Journal of Business Research*, *69*(2), 794–803.

[20] Pappas, I. O., & Woodside, A. G. (2021). Fuzzy-set qualitative comparative analysis (fsqca): Guidelines for research practice in information systems and marketing. *International Journal of Information Management*, *58*, 102310.

[21] Ragin, C. C. (2000). *Fuzzy-set social science*. University of Chicago Press.

[22] Ragin, C. C. (2009). *Redesigning social inquiry: Fuzzy sets and beyond*. University of Chicago Press.

[23] Rita, P., Oliveira, T., & Farisa, A. (2019). The impact of e-service quality and customer satisfaction on customer behavior in online shopping. *Heliyon*, *5*(10), e02690.

[24] Salehi, F., Abdollahbeigi, B., Langroudi, A. C., & Salehi, F. (2012). The impact of website information convenience on e-commerce success of companies. *Procedia-Social and Behavioral Sciences*, *57*, 381–387.

[25] Terra, L., & Casais, B. (2021). Moments of truth in social commerce customer journey: A literature review. *Digital Marketing & eCommerce Conference*, 236–242.

[26] Vakulenko, Y., Shams, P., Hellström, D., & Hjort, K. (2019). Service innovation in e-commerce last mile delivery: Mapping the e-customer journey. *Journal of Business Research*, *101*, 461–468.

[27] Wu, Y., Ngai, E. W., Wu, P., & Wu, C. (2020). Fake online reviews: Literature review, synthesis, and directions for future research. *Decision Support Systems*, *132*, 113280.

[28] Zhong, M., Qu, X., Chen, Y., Liao, S., & Xiao, Q. (2021). Impact of factors of online deceptive reviews on customer purchase decision based on machine learning. *Journal of Healthcare Engineering*, *2021*.