Consumer Behavior Review
Revista Comportamento do Consumidor

Piva, B. & Marques, L. (2024). Helpful clicks: consumers' perceived diagnosticity of online book reviews. *Consumer Behavior Review*, *8*(1), e-258115.

DOI: https://doi.org/10.51359/2526-7884.2023.258115

ISSN: 2526-7884
Editor: Prof. Dr. Marconi Freitas da Costa
E-mail: cbr@ufpe.br

Evaluation: Double blind review
Received: April 07, 2023
Approved: August 25, 2023

HELPFUL CLICKS: CONSUMERS' PERCEIVED DIAGNOSTICITY OF ONLINE BOOK REVIEWS

Beatrice Piva[1]
ORCID: https://orcid.org/0000-0001-8102-1222
E-mail: beatrice.piva5@gmail.com

Lénia Marques[1]
ORCID: https://orcid.org/0000-0002-6360-9919
E-mail: marques@eshcc.eur.nl

[1]*Erasmus University Rotterdam, Rotterdam, Netherlands*

Abstract
This paper investigates online feedback systems and their perceived diagnosticity in the book industry, using a mixed methods approach, with data retrieved from Amazon. Several variables concur to give the overall perceived diagnosticity of customer reviews: score systems, written text and its themes, and reviews date. Even though reviews are positively biased, long negative reviews are perceived as the most diagnostic. The older the review, the higher the perceived diagnosticity, due to the combined effects of early bird bias and winner circle bias. By understanding the role that variables have on diagnosticity, the findings can be used by scholars, as well as by e-commerce and marketplace webmasters who seek ways to improve the feedback systems of the online platforms in or outside the book industry.
Keywords: diagnosticity; online feedback systems; book industry; cultural consumption; electronic word-of-mouth.

INTRODUCTION

Research shows that books, in both electronic and paper versions, still represent a dynamic market (Flood, 2019). The digital environment has grown in importance since the first infant steps of technology enabling readers to use an online environment to purchase both paper books and e-books. In 2018 digital purchases of paper books and e-books on Amazon.com won respectively 42% and 88.9% over purchases on any other American retail sector (Day & Gu, 2019), confirming the positive trend of previous years (Milliot, 2018). The COVID-19 pandemic exacerbated the worldwide transition into digital shopping not only on major retailers such as Amazon, but also on niche platforms like Bookshop.org (Flood, 2020).

B.I. (Before the Internet) readers used to buy books in physical shops, hence the buyer's choice was limited to books stocked in store or orders from paper catalogues. The sources of information available back then were mainly bound to the books on shelves or recommendations. The Internet has had an impact on several points of the book lover consumption journey (Ramrattan & Szenberg, 2016), starting with the information search process. Websites are established as an additional source of information, when not the main one, which eventually affects the decision-making process. E-commerce platforms provide information that was already available in physical shops, such as author, year of publication, number of pages, genre, price, sales rankings, and a short summary of the book. The novelty of online shopping lied in the possibility to find other consumers' opinions as a result of online reviews and ratings.

Every Web interface has a different ability "to convey customers relevant product information that helps them in understanding and evaluating the quality and performance of products sold online" (Mudambi & Schuff, 2010, p. 188). The degree of success of doing so can be determined by customers' helpfulness evaluation, i.e. how much this online information is considered helpful for prospective consumers. Consumers' perceived helpfulness of such information, i.e. diagnosticity, is still under-researched. The concept of diagnosticity covers the extent to which customers recognise the usefulness of the information carried by online reviews and ratings in order to conclude their transaction of buying an experience cultural good, such as a book.

Consumption of experience goods cannot be estimated before the act of consumption itself, which explains the value of existing consumers' judgement. The main objective of this exploratory study is to investigate what different aspects of online reviews affect consumers' perceived diagnosticity on book retail websites and how they relate to each other. Books are a key example to illustrate the functioning of diagnosticity of reviews for experience cultural goods. By using a mix-methods approach to online reviews, this paper aims to contribute to a better understanding of reviewers' perception of the information available for experience goods. Through aggregated data about the way star ratings, posting dates, and helpfulness votes affect prospects in making the decision to purchase, this study analyses the effects of new media information channels on (potential) consumers in online shopping platforms.

The findings of this paper can be useful to marketers, webmasters and entrepreneurs to develop a better customer reviews' section on their platforms, in order both to help customers to gain more valuable information about their purchase and to increase the conversion rate on their online business. Scholars and researchers can find advancements in previous insights, deepening their discussion of trends and contributing to the body of knowledge of media-related business.

After analysing the implications of online retail websites, this article moves on to reviewing online feedback systems, before delving into the specific theories of perceived diagnosticity. The section on perceived diagnosticity looks in depth into diagnosticity's different elements of star ratings, written reviews, helpfulness ratings and chronological bias.

LITERATURE REVIEW

Online retail websites

Before the use of the Internet became such a critical part of people's daily life, scholars were (perhaps naively) sceptical about consumers' lack of technological skills and their openness to online purchases since it required unveiling personal information to service providers (Mansell, 1999). This forecast was clearly disproved as years went by. In a (post-)pandemic world, where online shopping has become a familiar activity for many, marketplaces and e-commerce websites represent the main sales channel for many goods, including low-cost experience cultural goods, such as books. Not only do buyers use online shopping platforms on a daily basis, but they also engage in consumer-generated content such as online reviews for all kinds of goods (Ross, 2010; Sparks & Browning, 2011).

Availability 24/7, the easy access to information on the website, and their competitive prices made readers prefer online books purchases and use intermediary online retailers, such as Amazon websites or Waterstones (Milliot, 2018; Statista, 2022). Most data concerns mainly the English-speaking world, which means that in some countries the situation might be different. Latcovich and Smith (2001) considered that the Internet shows no significant price competition due to the effects of price dispersions and facility of searching for lower prices. Therefore the challenge for book retailers lies nowadays in the need for differentiation through services, discounts, and a better-designed purchase experience (Ramrattan & Szenberg, 2016). The better online book shops display additional information, the higher consumers' satisfaction and thus the probability of purchase on that website will be. E-commerce platforms with well-perceived diagnosticity performance also have more possibilities to attract new potential customers. This relationship between (helpful) reviews, consumers' willingness to pay, and, consequently, seller's ability to charge higher prices has been the object of study (Chen et al., 2006; Clemons & Gao, 2008).

The fact that consumers seek information in order to improve their satisfaction given by the decision-making process about purchasing a good is not new (Stigler, 1961). Prospective consumers decide to conclude the information-search process once the information search costs are too high or higher than the expected benefits (Gursoy & McCleary, 2004). The Internet, and particularly marketplaces and e-commerces, lower consumers' search costs facilitating the information-search process (Brynjolfsson et al., 2011; Baye et al., 2015), yet each online shop uses different tools leading sometimes to "*information overload situations in online environments, due to a large amount of information available"* (Alzate, Arce-Urriza, Cebollada, 2021, p. 1). Concerning the publishing sector, there are two main practices for collecting information: free sampling and online feedback systems. Free sampling represents proof of the quality of the book for potential consumers (Brynjolfsson & Smith, 2000; Latcovich & Smith, 2001), while online feedback systems involve star ratings and online reviews, serving as peer recommendations.

Online feedback systems

To understand online feedback systems, we need to look better at the role of consumers, as they are the major players in the platforms, providing new information through their comments and thereby influencing prospects. Customer reviews represent both the unit of analysis of customers' perceived diagnosticity of book retail websites and the output of the feedback systems employed by each website. Feedback is a form of prosocial behaviour since consumers who leave a public evaluation do not expect nor obtain any payoff, while booksellers fail to capitalise on satisfied consumers returning to the marketplace as they probably will not buy the exact same book again (Duan et al., 2008; Utz, 2009; Tadelis, 2015). Therefore, a website feedback system leaves to platforms' users - and not to experts - a huge part of the task of providing additional information, which cannot be controlled or verified (Chevalier & Mayzlin, 2006; Basuroy et al., 2020). Introducing amateurs' reviews may cause trust-related issues as it is difficult to establish reliable criteria to distinguish booklovers from experts in cyberspaces (Ross, 2010).

Online reviews feedback systems represent the digital version of the so-called *word-of-mouth* (WOM), which is the "oral, person-to-person communication between a receiver and a communicator whom the receiver perceives as non-commercial, regarding a brand, product or service" (Arndt, 1967). Electronic WOM (eWOM) communication is mediated by a device and shared via the Internet (Hennig-Thurau et al., 2004). EWOM was in fact critical for sales of all kinds of goods, especially for experience goods sales during the COVID-19 pandemic, when worldwide people had an unusual amount of time to dedicate to reading and leisure activities but no friends and acquaintances to share opinions with.

High sales for popular products means that more people consume the product, evaluate it, and increase eWOM (Duan et al., 2008); in the meantime, online reviews raise the awareness about the existence and the perceived quality of the product among consumers leading to even higher sales. A positive relationship occurs between the volume of eWOM reviews and sales of experience goods such as movies (Duan et al., 2008), events and tourism experiences (Williams et al., 2017), books and books market share (Chen et al., 2004; Chevalier & Mayzlin, 2006; Chen, 2008; Øystein et al., 2017). However, it is yet to be revealed the exact way how online reviews work and what aspects have the most effects.

As a social behaviour, feedback and review systems are potentially biased because of the heterogeneity of customers' tastes and preferences, especially for cultural goods such as books (Jaffry & Apostolakis, 2011). These systems can be inflated as the type of feedback system may lead to fear of retaliation or corrupted behaviours (Dellarocas & Wood, 2008; Mayzlin et al., 2014; Tadelis, 2015). Despite these considerations, consumers tend to trust and respond intuitively to feedback because they are more or less unconsciously affected by information asymmetry and loss-averse attitude (Kirmani & Rao, 2000; Li et al., 2016; Li & Shimizu, 2018). As for experience goods, customers prefer to take risks and trust additional information carried by anonymous online reviews to reduce the information asymmetry, as with touristic services and travel destinations (e.g. Vermeulen & Seegers, 2009; Sparks & Browning, 2011; Marques & Williams, 2019). Online reviews of books are perceived as reliable, trustworthy, and helpful because peers are supposed to be more honest since they belong to the same side of the transaction - they are both consumers (Li et al., 2011; Park & Nicolau, 2014) and supposedly do not have any other incentive but honesty and altruism (Duan et al., 2008). Reviews become such an important source of key additional information for potential customers that customers' online feedback directly seems to affect perceived diagnosticity.

Perceived diagnosticity

When looking at perceived diagnosticity, it is important to understand how previous and actual consumers communicate and how this affects future books purchase (see Figure 1).

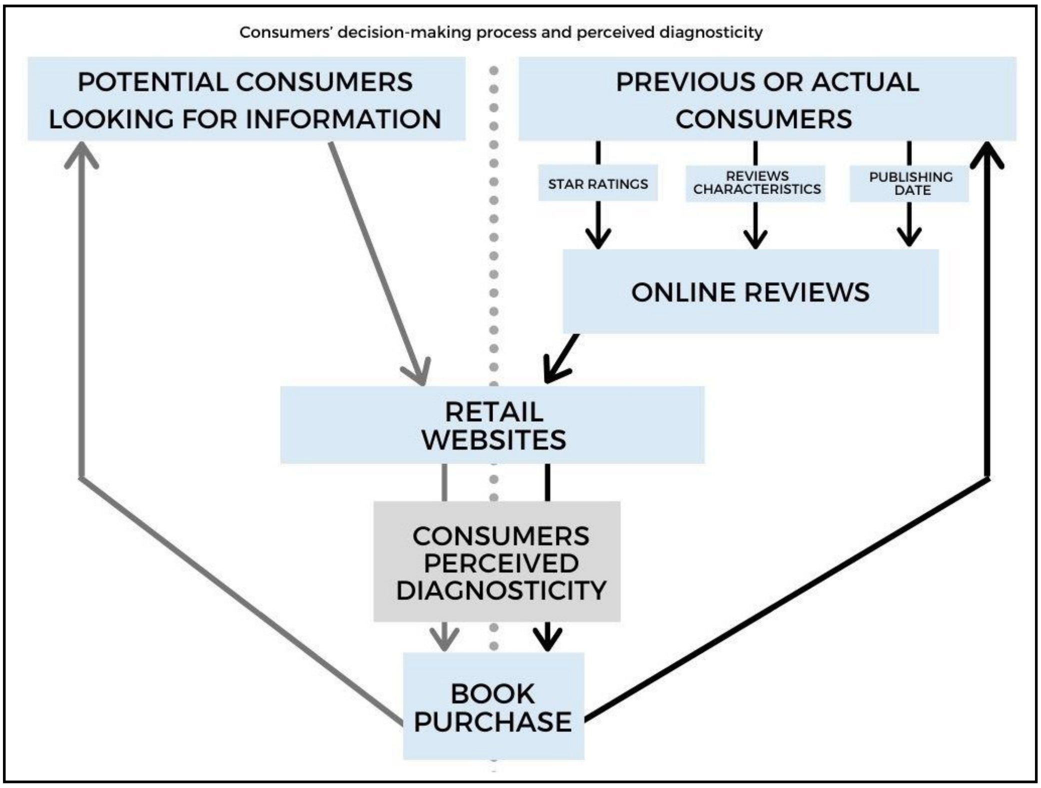

Figure 1. Consumers' decision-making process and perceived diagnosticity

According to Filieri (2015), diagnosticity is the degree of perceived usefulness of the information provided by a website according to consumers who rely on one specific website (and its consumers' reviews) to conclude their decision-making process. On Amazon, the perceived diagnosticity of prospective customers is linked to the perceived helpfulness of online reviews related to consumers' decision-making process (Mudambi & Schuff, 2010). For example, on Amazon books catalogue, the *helpfulness* votes feature facilitates assessing online reviews' diagnosticity. Helpfulness votes allow consumers to assess the degree of usefulness and novelty of the information provided in the review and its form. Several factors affect diagnosticity in online reviews, such as review rating, review depth, and review readability, reviewers' profile and product type (Chua & Banerjee, 2014). However, prospects' decision whether to trust or not the content depends on the quality of the information (Guoyin et al., 2021). In this study, the diagnosticity of book reviews in Amazon is demonstrated in helpfulness votes and it can be observed in star ratings, content (body) of written reviews, and posting date (see Figure 2).

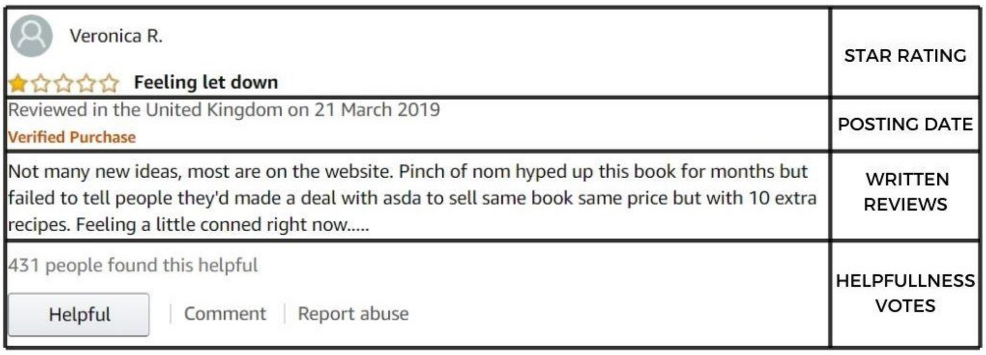

Figure 2. Example of customer review of the book Pinch of Nom from Amazon.co.uk and its corresponding categories

Star ratings

Online reviews usually comprise a score (star) system that expresses quantitatively an opinion which is qualitatively described in the body of the review/written comments (Tadelis, 2015). Star ratings convey consumers' satisfaction with the item and the delivery service offered by the sellers and the website. It summarises the valence of the review, which is the inclination of the review measured as negative (1-2 stars), positive (4-5 stars), or moderate (3 stars) (Duan et al., 2008).

High-starred and more reviewed books tend to have a higher market share (Chevalier and Mayzlin, 2006; Li and Shimizu, 2018) possibly as a consequence of the psychological phenomenon better known as *herd behaviour*: when all other factors are equal, consumers tend to choose books better rated and with more reviews (Deutsch and Gerard, 1955; Chen et al., 2004; Chen, 2008). Given the information asymmetry and the multitude of available products online, consumers tend to trust more popular brands because popularity is associated with good quality (Bikhchandani et al., 1992; Brynjolfsson and Smith, 2000; Chen, 2008). According to Chevalier and Mayzlin (2006), book reviews are positive on average since it is mostly satisfied consumers who post online reviews, while non-satisfied users leave the platform (Duan et al., 2008; Tadelis, 2015). By virtue of prospect theory, the marginal effects of positive and negative reviews decrease with the increase in their volume (Li and Shimizu, 2018). In other words, when reviews are positively biased, negative reviews tend to affect consumers' perceived diagnosticity more than positive ones.

Written reviews

Around 90% of prospective customers merely scan written reviews, while only 10% of consumers take the time to read reviews searching for information more deeply (Madu & Madu, 2002;

Raju & Joseph, 2017; Sunil et al., 2018). However, not all reviews mention crucial themes and accurate information that are relevant to assess product quality (Raju & Joseph, 2017). For instance, “fabulous”, “great book”, “not worth it” represent subjective opinions, yet positive or negative adjectives are some of the most cited words in online book reviews. Particularly, emotional content is effective on perceived credibility and diagnosticity in unpleasant reviews (Guo et al., 2020).

Helpfulness ratings

Since not all reviews are equal, Amazon websites allow customers to rate not only books sold on the retail website but also peer reviews according to their helpfulness in providing useful information to potential consumers. This feedback system simplifies the research for fruitful information among all the data provided by reviews so that reviews helpfulness votes represent customers' reviews diagnosticity (Li et. al, 2013; Mudambi & Schuff, 2010).

As Rietsche and his colleagues (2019) expose with some detail, there are different studies pointing out in sometimes opposite directions regarding review extremity (valence). On the one hand, moderate reviews can be considered more significant because they may be a signal of objective assessment (Chua & Banerjee, 2014; Mudambi & Schuff, 2010). On the other hand, negative reviews are deemed more meaningful to consumers because of *negativity bias*, as known as the tendency of considering negative feedback as more diagnostic because reviewers are regarded as more intelligent, knowledgeable, and critic about the product (Folkes & Sears, 1977; Chevalier & Mayzlin, 2006; Kim et al., 2008; Wu et al., 2011). Regarding the depth of the information carried by reviews, it includes relevance, accuracy, and comprehensiveness of the reviews, the themes covered and the language adopted. The degree of review depth affects the perceived diagnosticity of information, which also raises consumers’ confidence in completing the decision-making process (Mudambi & Schuff, 2010).

Chronological bias

Usually books receive hundreds or thousands of online reviews. However, human beings are not able to handle all the available sources of information due to limited processing capacity (Alzate, Arce-Urriza, Cebollada, 2021). Amazon comments are organised in several pages and users can decide to read the most recent reviews or “top reviews” first. Old reviews are perceived as more diagnostic than new reviews due to *winner circle bias* and *early bird bias*. The winner circle bias has been derived from the so-called Matthew effect and it represents the trend where more rated reviews obtain even more helpfulness votes because they capture the readers’ attention reducing their objectivity regarding the actual content of the review (Liu et al., 2007). Early bird bias happens when the first reviews posted receive more helpfulness votes because those reviews are available to potential consumers for a longer period of time (Liu et al., 2007). Eventually, these two biases influence each other and work together (Li et al., 2013).

In brief, previous studies have looked into different aspects of diagnosticity mainly separately. This exploratory study therefore provides insights on how a set of review features (stars ratings, posting dates, helpfulness votes, body of review) work together and affect prospective customers in making the decision to purchase a book on an online platform.

METHODOLOGY

In order to address these different aspects of diagnosticity (stars ratings, posting dates, helpfulness votes, body of review), a mixed-methods approach was used to analyse book reviews from Amazon. Python software was employed to scrap the set of reviews from the website during a relatively short period of time (less than a day) to limit timeline biases. The data set consists of all the reviews posted on Amazon.co.uk on the page of the first five English bestsellers of the ranking on 13th May 2019. To present effective suggestions to website managers, this paper adheres as much as possible to real life circumstances, analysing all reviews for top bestsellers on the most worldwide spread marketplace for books, no matter the genre or the date of publishing. The only exception was a limitation to the reviews of Book 5, since Amazon caps the visualisation of all available reviews at five thousand. Therefore, when reviews exceed that number, which was the case of Book 5, they are not accessible, and therefore not included in the data set.

Books in the sample include both fiction and non-fiction. Three books have thousands of reviews, while the other two books have less than one hundred reviews each. The final sample consists of 12,670 online reviews. Book 1 is a slimming recipe book while Book 2 is a novel for teenagers. Book 3 and Book 5 are medical-related (fictionalised) biographies. Book 4 is a guide for house decluttering written by a notorious English influencer. Details on the data collection can be found in Table 1.

Table 1
English Bestsellers on Amazon.co.uk on 13th May 2019

	Ranking Position	Title	Numbers of Reviews	Release date
Book 1	1	Pinch of Nom	4544	03/21/19
Book 2	2	You Got This	19	05/02/19
Book 3	3	Confessions of a Menopausal Woman	70	06/28/18
Book 4	4	Hinch Yourself Happy	3054	04/04/19
Book 5	5	This Is Going to Hurt	5745	04/19/18

After being scraped in Python, a .csv file was created, and the reviews were organised in Excel. Excel and SPSS 26 were used for statistical analysis. The reviews included star ratings, title, date, body of the text, and helpfulness votes, as shown in Figure 2.

Before delving into the quantitative analysis, and to seek to understand the content and context of the review itself in a way that it would shed light on the quantitative findings, a qualitative analysis was performed taking into consideration the 100 most popular reviews or the total reviews for each book when there were less than 100. First, reviews were inductively coded to allow the most important and frequent themes to emerge while getting to saturation point. The themes that emerged from this analysis were *delivery services, reviewers' feelings, book's content, information about the author(s), disappointment and desire of returning the book*, and *positive characteristics of the book*. It is important to note that polarisation can occur in the number of helpfulness votes, as only a few reviews receive the majority of the votes. This is explained by early birds bias and winner circle bias.

Second, the five most helpful reviews for each book have also been clustered to understand the topics that make popular reviews so meaningful. This was followed by a clustering process on Python based on the total of words included in the review, regardless of the fact that the topic had a positive or negative valence in the sentence. The ten most common words for each book have been listed and visually displayed using word cloud visualisation tools. Remarkably, most of the common words were positive or negative adjectives that do not provide objective information, but do point out to an emotional and valuation type of reaction (*lovely looking, fantastic, disappointing, not great*).

For the quantitative analysis, certain parameters and criteria were used. Star ratings were considered as a proxy for diagnosticity so to calculate the average number of reviews that have the same valence and their helpfulness votes mean; hence to confirm the occurrence of positivity or negativity bias or both. Since content inquiry of the qualitative analysis involves also the number of words used to communicate the message, the text has been studied considering review length. According to Madu and Madu (2002), consumers want to easily find information, therefore it was expected that shorter reviews would be considered more helpful. Word count tools were employed to measure overall review length (not considering the title), and then the number of words was related to the number of helpfulness votes to find a hypothetical trend.

The reviews' date was the last feature of the analysis as the research focuses on the correlation between the date reviews were posted and the number of helpfulness votes they received. The older the review, the higher the number of helpfulness votes (Liu et al., 2007). Nevertheless, the correlation between the date the reviews are posted and the perceived diagnosticity may be problematic for experience goods such as books. Some reviews are posted exactly on the date the book was published. For experience goods, such as books, customers should consume before assessing the quality; either the customer has to read the book previously (or very quickly), or the review is essentially about personal expectations and delivery services. The reviews were divided according to the date they were posted in order to assess whether reviewers had posted more reviews on the days immediately following the release date of the book or later. Having divided the number of helpfulness votes per day, this step sought to understand whether early bird reviews tend to receive more or fewer helpfulness votes than late reviews, helping to determine whether winner-circle bias occurs.

To complete the analysis of the impact of reviews on the consumer's diagnosticity perception, the interaction amongst these parameters was performed with multiple linear regression. This way, we could consider how much the independent variables (in this case star ratings, review length, and the book number) affect the final dependent variable, which is the number of diagnosticity votes for each review. Due to the significant difference in the number of reviews per book, the regression was run firstly including data from all five books and secondly excluding Books 2 and 3. No significant discrepancies were detected.

FINDINGS AND DISCUSSION

From the sample of books in this article, it can be noted that they are listed according to their position in the sales ranking (see Table 1). However, the number of reviews for these books changes significantly. Book 1, Book 4, and Book 5 have more than three thousand reviews each, while Book 2 and Book 3 gathered less than one hundred reviews. Moreover, Books 2 and 3 got a higher rank position compared to Book 4 and 5, contradicting the theory proposed by Chevalier and Mayzlin (2006), who stated that more reviewed books also have higher sales. For Book 2, this might be explained because Amazon gave readers the possibility to pre-order it, which meant that consumers bought books before reading any reviews. Book 3 may appear in the top five positions because Amazon considered all copies sold, both in the ebook and audiobook versions (firstly released on 28th June 2018) and in the paper version (available from 4th April 2019).

Star Ratings and Helpfulness Votes

It was expected to find reviews that are positive on average, considering the sample includes only bestsellers. Extreme negative reviews were also supposed to be the most diagnostic for their disruptive knowledge. Overall, univocal results seem to confirm previous studies concerning the relationships between star ratings and diagnosticity. Reviews for each book are positive on average (see Table 2). In this case, they are extremely positive as the average is 4.63 stars per review, confirming the previous findings by Chevalier and Mayzlin (2006) that online reviews are positive on average, being affected by a positive bias.

Table 2
Stars Rating and Helpfulness Votes for Top Five Bestsellers Books on Amazon.co.uk (on 13th May 2019)

	Average Stars Rating	Average HV* Per Review	Average HV 1 Star Reviews	Average HV 5 Stars Reviews
Book 1	4.79	3.06	15.75	2.85
Book 2	4.63	3.05	1	3.24
Book 3	4.31	4.89	28	3.16
Book 4	4.61	4.98	49.09	1.39
Book 5	4.83	1.17	10.73	1.09

Average	4.63	3.43	20.91	2.35

Note. *HV: Helpfulness Votes

Each review has 3.43 helpfulness votes on average, but the means vary across books being higher than average for books with fewer reviews. Nonetheless, it is not possible to assert that reviews' helpfulness affects sales among bestsellers. Books with the lowest means for star ratings also show the highest average of helpfulness votes per review, reinforcing the theory about negativity bias (Wu et al., 2011). The highest average of helpfulness votes belongs to 1-star reviews. Overall, extremely positive or negative reviews, such as 5-stars and 1-star reviews, received most helpfulness votes, confirming the study by Forman et al. (2008). Even taking into account the deviancy given by the poor data of Book 2, the findings support the hypothesis that for experience cultural goods such as books, extremely negative reviews are perceived as the most diagnostic ones (Folkes and Sears, 1977; Chevalier and Mayzlin, 2006; Kim et al., 2008; Wu et al., 2011).

Reviews content and Helpfulness Votes

From the thematic analysis of the reviews, clear patterns emerged. Certain themes occur in many reviews and different reviewers write similar or identical words to express the same theme about different books. The following themes emerged:

1. *delivery services*
2. *reviewers' feelings*
3. *book's content*
4. *information about the author(s)*
5. *disappointment and desire of returning the book*
6. *positive characteristics of the book*

These themes are recurrent in all the reviews, including those that did not receive any helpfulness vote. Most reviews cover usually two to three themes. According to Madu and Madu (2002) and Raju and Joseph (2017), customers look for online reviews to acquire more information about the goods they are willing to consume, especially for experience goods such as books. Therefore, themes related to book content and writers were expected to be the most significant topics. When analysing the themes ratio in the 25 most helpful reviews (with 5 top reviews per book), books' content and information regarding the author(s) were debated in almost all of them (see Figure 3). The authors were mentioned not only with their names and third-singular person pronouns, but many reviewers referred to the authors as if they were writing directly to them and not *about* them using the pronoun "you". This aspect has also been confirmed by the word analysis. This use of "you" can be explained by a certain emotional connection to the book, to the author and also a preferred form of active dialogue, which, in turn, unconsciously, leads to more cognitive and emotional involvement from the consumer (reader of reviews).

Half of the reviews discussed positive attributes of the books, such as the usefulness of the tips or the quality of the photographs. Many reviewers also disclosed their feelings or their stories about books, which is in line with the identity-disclosure theories (Forman et al., 2008; Mudambi & Schuff, 2010). Even if self-disclosure was not the most common theme, it is covered in the most useful review for each book, meaning that consumers highly value personal information about the reviewer over other additional facts. This again reinforces the emotional connection that might be involved in understanding diagnosticity. Lastly, some reviewers wrote about their experience on the delivery service of Amazon.co.uk. These reviews concern the delivery itself or the conditions of the item. In particular, this theme was found in the most helpful reviews of three of the five books in the sample. Overall, only 5 reviews out of 25 had a negative valence, supporting once again the occurrence of positivity bias (Chevalier & Mayzlin, 2006).

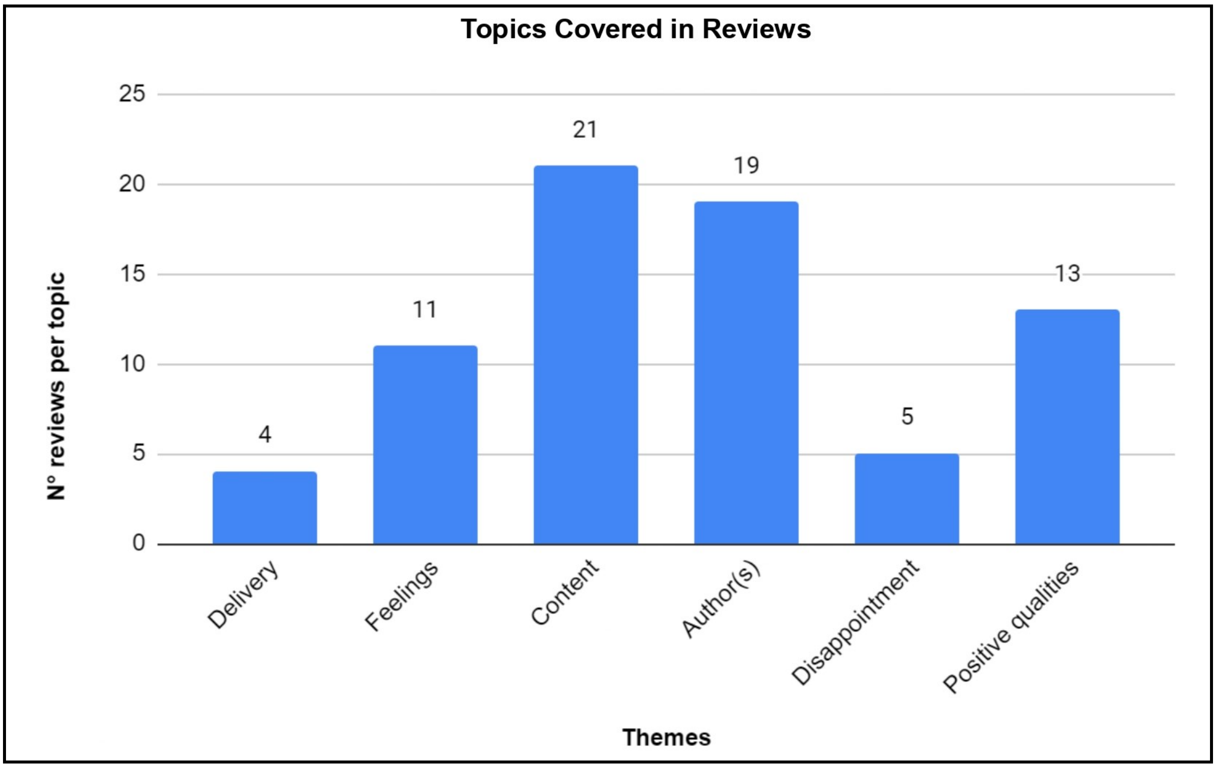

Figure 3. Topics Covered in Reviews. Ratio in the 25 Most Helpful Reviews (Five Top Reviews per Book)

On another level, the most used words in the book reviews are also quite illustrative (see Table 3). The most common words for all books are "read" and "book", which is not surprising. However, it is noteworthy that "read" is not one of the ten most common words in reviews for Book 1, which might be explained by the practical nature of the text (a recipe book).

Table 3
Ten Most Used Words in Reviews per Book

Book 1 - Pinch of Nom by C. Allinson, K. Allinson, K. Featherstone		***Book 2 - You Got This by Bryony Gordon***		***Book 3 - Confessions of a Menopausal Woman by Andrea McLean***		***Book 4 - Hinch Yourself Happy by Sophie Hinchliffe***		***Book 5 - This is Going To Hurt by (doctor) Adam Kay***	
Reviews	**4544**	**Reviews**	**19**	**Reviews**	**70**	**Reviews**	**3540**	**Reviews**	**5000***
Word	**Freq.**	**Word**	**Freq.**	**Word**	**Freq.**	**Word**	**Freq.**	**Word**	**Freq.**
book	5593	book	29	book	87	book	4302	read	3907
recipes	4585	read	28	read	54	Hinch	2437	book	3835
great	2210	Bryony	19	menopause	39	mrs	2234	funny	2436
love	1610	daughter	11	Andrea	32	love	1945	NHS	2236
easy	1587	year	8	July	22	read	1874	doctor	1016
follow	974	love	7	good	21	cleaning	1053	great	1014
amazing	939	amazing	6	life	19	amazing	968	brilliant	913
wait	858	girl	6	August	18	tips	907	doctors	896

fantastic	744	normal	6	women	18	Sophie	897	life	794
nom	734	chapter	5	honest	17	great	722	hilarious	779

Note. *The overall number of reviews for Book 5 is 5745, but Amazon.co.uk allows users to browse up to 5000 reviews, so the table analyses only 5000 reviews.

Another common word is the name, surname, or profession of the author (see *Bryony, Andrea, Sophie, Hinch, doctor*). This finding is common throughout the books in the sample, and it confirms reviewers' tendency to address the review directly to the authors or to refer to them as if they were acquaintances. The rest of the list includes adjectives or verbs that indicate personal opinions and feelings about the book, which are usually enthusiastic, such as great, love, amazing, honest, good, and funny. Despite the positive valence of these adjectives or verbs, they are not really descriptive as they do not provide a clear reason why reviewers liked or disliked the book.

Reviews Length and Helpfulness Votes

Madu and Madu (2002) posit that most consumers seek immediate information about the quality of the item, thus short reviews were expected to be more diagnostic than long reviews as the former go directly to the point. However, examining the length of all the online customers' reviews included in the sample together with the helpfulness votes, there is not a clear pattern (see Table 4).

Table 4
Reviews' Length/Helpfulness Votes per Book

	Average length	Average helpful. votes	Average length negative* reviews	Average helpful. votes negative* reviews	Average length positive* reviews	Average helpful. votes positive* reviews
Book 1	29,48	3,06	49,76	10,93	28,44	2,78
Book 2	67,79	3,05	49,00	1,50	70,29	3,24
Book 3	53,64	4,89	155,22	17,33	30,89	2,59
Book 4	48,41	4,98	60,82	40,16	46,61	1,38
Book 5	34,22	1,17	58,86	6,90	33,50	1,04

Note. *negative reviews got 1 to 2 stars as a rating; positive reviews got 4 to 5 stars.

Book reviews are 46 words long on average. Reviews of books that are reviewed negatively are longer on average (60 words), whilst reviews of highly reviewed books are shorter on average (37 words). Despite the number of reviews for each book, the reviews perceived as more useful are outliers as they are significantly longer than the average. The fact that outliers are the most voted might also be significant in that consumers look for detailed information where they can understand the rationale behind a certain review that they consider helpful, which was later confirmed by the multiple linear regression analysis.

Reviews Date and Helpfulness Votes

The date reviews were posted may affect their perceived diagnosticity, star ratings, and content. The reviews were divided according to the date they were posted, with the expectation that reviews posted in the early days were considered more helpful, enhancing perceived diagnosticity. Most of the reviews were posted during the week that followed the publishing date, and from there on, the comments decreased. However, the decrease is not steady because there are usually peaks two or three weeks after the publishing date, as illustrated by Book 4 (see Figure 4).

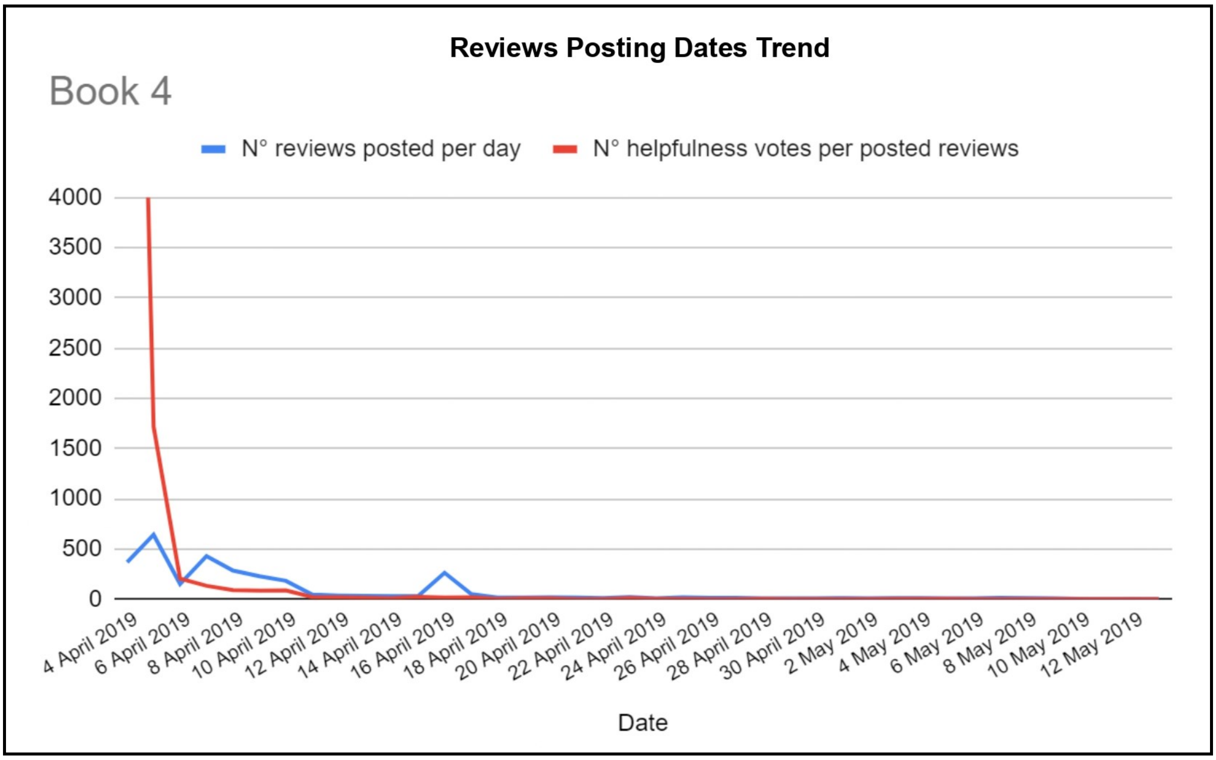

Figure 4. Reviews Posting Dates for Hinch Yourself Happy by Sophie Hinchliffe (Book 4)

As in Figure 4, it can be observed that the trend of helpfulness votes does not perfectly follow the posting date line. The helpfulness votes line looks like an exponential function with a positive coefficient so it exhibits a steady and sharp decrease. Reviews posted immediately after the publishing date reach the highest number of helpfulness votes, then the line drops so that late reviews gather few (or no) helpfulness votes, evidencing the early bird effect (Li et al., 2013). Nevertheless, older reviews may be affected also by winner circle bias (Li et al., 2013). Since the first reviews were read by more customers who looked for information about the book, they acquired more votes. Having more votes than new reviews, they also received more attention because Amazon displays top reviews in a higher position in the reviews section of the website in the default option, starting a vicious/virtuous circle which makes them gather even more attention (Madu and Madu, 2002).

Relating factors towards diagnosticity

In order to understand the role different factors play in the perceived diagnosticity of consumers, a multiple linear regression analysis was performed. The regression results reinforce the idea that star ratings, review length, and the book ranking influence the number of helpfulness votes that reviews receive. According to the results of the regression, a possible function of the dependent variable Helpfulness Votes (H) can be drawn:

$$\mathbf{H = -0{,}416B - 5{,}202R + 0{,}56L + 26{,}876}$$

Considering B as the independent variable for book ranking, R the number of star ratings, L the length of reviews and 26,876 the constant.

Two models were tested where the dependent variable is the number of helpfulness votes, and the independent variables the following:

- book ranking - relating to the the position of the book in the sales ranking on the moment of data collection;

- star ratings - relating to how many stars (from 1 to 5) were given by the consumers;
- length of the review - the number of words used in the written section of the review, i.e what consumers have really written.

In Model 1, all five books of the sample were included (see Table 5). Although the factors explain only a marginal percentage of the helpfulness votes (6,4%), significant relationships are found between all the three independent variables and the dependent variable (see Table 6). In Model 2, the books with fewer reviews (Books 2 and 3) were taken out of the statistical analysis to avoid bias. However, the differences were very marginal and the levels of significance the same. Therefore all books can be considered as in Model 1.

Table 5
Model 1

Model	R	R square	Adjusted R square	Standard error
1	,252[a]	0.064	0.063	19.657

Note. [a]. Predictors: (constant), Length Review, Book Ranking, Star Rating

Table 6
Model 1 Coefficients[a]

Model		Unstandardised Coefficients		Standardised Coefficients		
		B	Standard Error	Beta	t Stat	Significance
1	(Constant)	26.876	1.112		24.170	0.000
	Book Ranking	-0.416	0.099	-0.036	-4.215	0.000
	Star Rating	-5.202	0.216	-0.209	-24.119	0.000
	Length Review	0.056	0.004	0.117	13.420	0.000

Note. [a]Dependent variable (H): number of Helpfulness Votes.

Books in higher-ranking positions that have high star ratings gather the least helpfulness votes. In fact, the multiple linear regression analysis confirms that most helpful reviews belong to the lowest ranked books and show the lowest star rating. This is not surprising as a negative relationship between helpfulness votes and star ratings was expected (see section 5.1). This can be explained because, when in doubt, consumers appreciate the lowest scores of reviews as the most useful to make a (better) decision due to negativity bias effects. As a consumer expresses clearly while pointing the finger at what they consider a bad practice by Amazon: "If Amazon deletes (sic) bad reviews, what's the point in reviews?" (Book 4, 4 April 2019).

Contrary to star rating, the length of the review has a positive relationship with the helpfulness votes. The longer the review, the more relevant it seems to be for the consumer. For example, in Book 3 the most helpful review (136 votes) has one-star rating and is comparatively very long as it is composed of 398 words. The negative score is clearly explained by the set of reasons why the book is not interesting for this reader:

"This book should have been called "Diary of a privileged middle-aged woman"..(...) Don't expect anything revelatory in this book (...) She doesn't live in the real world. (...) OMG. It is suggestions like that which make people simply employ a man instead. Really!!!..And I am sorry Andrea (...) (you) struck me as a bit of a moaner. (...) As for the price of this book... rip off (...)"

Since people look for useful and clear arguments for whether or not to buy the book, the review above has been identified as highly informative. The reviewer explains their dislike based on certain reasons more than on simple adjectives like "not great" or "disappointing". This way such a review can be perceived as helping new consumers to avoid what they later might consider a mistake, also relating to price and expectation. These factors all together contribute therefore to different degrees of perceived diagnosticity.

CONCLUSIONS

This study explores the concept of perceived diagnosticity in the book sector. By using reviews posted on Amazon and extracting different variables, it sheds light on aspects which influence the customers perception of useful information in online reviews and corresponding ratings.

According to previous studies (e.g. Bonabeau, 2004), book star ratings are positively biased, as a consequence of herd behaviour in the consumption of cultural goods. The fact that satisfied customers return to the platform to write a review is also a factor to take into consideration, as non-satisfied consumers do not have the same motivation to spend more time sharing their views (Tadelis, 2015). Even though there are only a few online negative reviews that carry criticism about the quality of the product, those reviews reduce the quality uncertainty (Kim et al., 2008; Wu et al., 2011). The most useful reviews do not concern the positive characteristics of books, but the negative ones, validating the coexistent negative bias too.

Since books are cultural experience goods, consumers' often search for further information because of the natural quality uncertainty (Raju & Joseph, 2017), which is also confirmed by the present research. Firstly, most of the reviews concern the books' content. Although customers can find a description of the books on the website, they rely more on the information provided by peers, confirming the thesis by Park and Nicolau (2014). Secondly, self-disclosure about feelings is proved once again to be one of the consumers' most-valued themes (Mudambi & Schuff, 2010). The written text on the review appears to have a great influence on consumers' perceived diagnosticity. However, the reviews' length can be a proxy for reviews' helpfulness. Therefore, what consumers value the most is the content of the reviews and their quality, which can be related to the quantity of the information provided. This could be reinforced by webmasters and marketplace business developers by, for example, pre-set specific questions to improve information quality and the reviews' perceived diagnosticity. Often there are price-related comments, and it could be insightful to investigate how many of the potential consumers clicking on the helpfulness votes do effectively (not) purchase the items.

The enthusiasm of readers who have immediately read the book (influencing herd behaviour) explains the peak of reviews during the first week of sales. Reviews posted in the first days after the publishing date receive a high number of helpfulness votes due to the early bird effect and the winner circle bias (Liu et al., 2007). E-commerce platforms may encourage customers to write comments a few weeks after their purchase to give them time to read the book and only then help them to recall and compose a review. Although, the industry and e-commerce platforms know that once a book becomes a bestseller, herd behaviour will lead to even more sales. Thus online sellers publish books that are similar to the bestsellers in order to sell more, leading to less differentiation for consumers and a certain uniformisation. A discussion to have in the future also in regards to diagnosticity relates to the role algorithms play and their influence on different factors, such as helpfulness votes, star ratings or text of reviews.

Expanding the analysis to a wider sample, including bestsellers and bottom-sellers, could be beneficial to understand better the book sector as a whole, as very often the focus is on top sellers, top reviews, or at the most star ratings. Since the present sample includes only bestsellers and not items from the last part of the long tail, it was not possible to generalise results. The researchers also acknowledge that the effects of star ratings, themes, length and date when those reviews are posted may differ according to genres or position in the ranking.

With the COVID-19 pandemic, the use of e-commerce platforms only increased. Although there is no reason to believe that there would be a change in the results, further research into the effects of the pandemic on the perceived diagnosticity of book reviews could be done.

These findings can be useful for marketplaces and e-commerce websites managers who have the opportunity to develop more effective feedback systems thanks to the understanding of the role and effect of different diagnosticity aspects and the influence that these have in the perception of consumers. Besides, the present research contributes to a deeper understanding of diagnosticity as a concept applied to cultural goods in the digital environment.

Research ethics statement

This article is the authors' own original work, which has not been previously published elsewhere.

Author contribution statement

Beatrice Piva wrote the starting version of the article, setting the literature review. She was responsible for gathering data and performed most of the data analysis. Lénia Marques ensured the coherence of the research process from ideation till publication, re-writing and revising the paper, in particular methodology, findings and conclusions.

Funding

There was no funding for this research.

Disclosure statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank Sara Cardi Cigoli for supporting with an early revision of the paper.

References

Alzate, M., Arce-Urriza M., & Cebollada J. (2021). Online Reviews and Product Sales: The Role of Review Visibility. *Journal of Theor. Appl. Electron. Commer. Res*. 16, 638–669.

Arndt, J. (1967). *Word-of-mouth advertising: a review of the literature*. New York: The Advertising Research Foundation Inc.

Basuroy, S., Abraham, R.S., et al. (2020). Is everybody an expert? An investigation into the impact of professional versus user reviews on movie revenues. *Journal of Cultural Economics* 44, 57–96.
Baye, M.R., De los Santos, B., & Wildenbeest, M.R. (2015). Searching for Physical and Digital Media: The Evolution of Platforms for Finding Books. In: Goldfarb A, Greenstein SM, Tucker CE (Ed.), *Economic Analysis of the Digital Economy*. University of Chicago Press, pp.137-165.
Bikhchandani, S., Hirschleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change in informational cascades. *Journal of Political Economy* 100(5):992–1026.
Bonabeau, E. (2004). The perils of the imitation age, *Harvard Business Review* 82(6):99–104.
Brynjolfsson, E., & Smith, M. D. (2000). Frictionless Commerce? A Comparison of Internet and Conventional Retailers. *Management Science* 46(4): 563-585.
Brynjolfsson, E., Hu, Y. J., & Simester, D. (2011). Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales. *Management Science* 57(8): 1373-1386.
Chevalier, J. A., & Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research* 43(3): 345-354.
Chen, P., Wu, S., & Yoon, J. (2004). *The Impact of Online Recommendations and Consumer Feedback on Sales*. Proceedings of the International Conference on Information Systems. Report no: ICIS 2004-58. Available at: http://aisel.aisnet.org/icis2004/58.
Chen, P., Dhanasobhon, S., & Smith, M. D. (2006). All Reviews are Not Created Equal: The Disaggregate Impact of Reviews and Reviewers at Amazon.com. Available at: SSRN's eLibrary: http://ssrn.com/abstract=918083.
Chen, Y. (2008). Herd behavior in purchasing books online. *Computers in Human Behavior* 24(5): 1977-1992.
Chua, A. Y. K., & Banerjee, S. (2014). *Developing a theory of diagnosticity for online reviews.* Proceedings of The International MultiConference of Engineers and Computer Scientists (IMECS). Report no: IMECS 2014-1. Available at: https://dr.ntu.edu.sg/handle/10220/43861.
Clemons, E. K., & Gao, G. (2008). Consumer informedness and diverse consumer purchasing behaviors: Traditional mass-market, trading down, and trading out into the long tail. *Electronic Commerce Research and Applications* 7(1): 3–17.
Day, M. & Gu, J. (2019). The Enormous Numbers Behind Amazon's Market Reach. Bloomberg. Available at: https://www.bloomberg.com/graphics/2019-amazon-reach-across-markets/
Dellarocas, C., & Wood, C. A. (2008). The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias. *Management Science* 54(3): 460-476.
Deutsch, M., & Gerard, H. (1955). A study of normative and informational social influences upon individual judgement. *Journal of Abnormal Social Psychology* 51(3): 629–636.
Duan, W., Gu, B., & Whinston, A. B. (2008). The dynamics of online word-of-mouth and product sales - An empirical investigation of the movie industry. *Journal of Retailing* 84(2): 233–242.
Filieri, R. (2015). What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. *Journal of Business Research* 68(6): 1261-1270.
Flood, A. (2019). Leading the entertainment pack: UK print book sales rise again. *The Guardian*. Available at: https://www.theguardian.com/books/2019/jan/03/leading-the-entertainment-pack-uk-print-book-sales-rise-again.
Flood, A. (2020). 'This is revolutionary': new online bookshop unites indies to rival Amazon. Available at: https://www.theguardian.com/books/2020/nov/02/this-is-revolutionary-new-online-bookshop-unites-indies-to-rival-amazon
Folkes, V. S., & Sears, D. O. (1977). Does Everybody Like a Liker?. *Journal of Experimental Social Psychology* 13(6): 505-519.
Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the Relationship between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets. *Information Systems Research* 19(3): 291–313.
Guo, J., Wang, X., & Wu, Y. (2020). Positive emotion bias: Role of emotional content from online customer reviews in purchase decisions. *Journal of Retailing and Consumer services* 52(C).

Guoyin, J., Fen, L., et al. (2021). Effects of information quality on information adoption on social media review platforms: moderating role of perceived risk. *Data Science and Management* t 1. 13–22
Gursoy, D., & McCleary, K. W. (2004). An Integrative Model Of Tourists' Information Search Behavior. *Annals of Tourism Research* 31(2): 353–373.
Hennig-Thurau, T., Gwinner, K. P., Walsh, G. et al. (2004). Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the Internet?. *Journal of Interactive Marketing* 18(1): 38–52.
Jaffry, S., & Apostolakis, A. (2011). Evaluating individual preferences for the British Museum. *Journal of Cultural Economics*, 35, 49–75.
Kim, D. J., Ferrin, D. L., & Rao, H. R. (2008). A Trust-Based Consumer Decision-Making Model in Electronic Commerce: The Role of Trust, Perceived Risk, and Their Antecedents. *Decision Support Systems* 44(2): 544–564.
Kirmani, A., & Rao, A. R. (2000). No pain, no gain: A critical review of the literature on signaling unobservable product quality. *Journal of Marketing* 64(2): 66-79.
Latcovich, S., & Smith, H. (2001). Pricing, Sunk Costs, and Market Structure Online: Evidence from Book Retailing. *Oxford Review of Economic Policy* 17(2): 217- 234.
Li, M., Huang, L., Tan, C., et al. (2011). *Assessing The Helpfulness Of Online Product Review: A Progressive Experimental Approach*. Pacific Asia Conference on Information Systems. Report no: PACIS 2011-111. Available at: http://aisel.aisnet.org/pacis2011/111.
Li, M., Huang, L., Tan, C. et al. (2013). Helpfulness of Online Product Reviews as Seen by Consumers: Source and Content Features. *International Journal of Electronic Commerce* 17(4): 101–136.
Li, Z., & Shimizu, A. (2018). Impact of Online Customer Reviews on Sales Outcomes: An Empirical Study Based on Prospect Theory. *The Reviews of Socionetwork Strategies* 12(2): 135-151.
Liu, J., Cao, Y., Lin, C. Y. et al. (2007). *Low-quality product review detection in opinion summarization.* Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLPCoNLL). Report no: 2007. Available at: https://www.aclweb.org/anthology/D07-1035.
Madu, C. N. & Madu, A. A. (2002). Dimensions of e-quality. *International Journal of Quality and Reliability Management*. 19(3): 246–258.
Mansell, R. (1999). New media competition and access: the scarcity-abundance dialectic. *New Media & Society* 1(2): 155-182.
Marques, L., & Williams, N. (2019). Networked Hospitality and Placemaking in the Sharing Economy. *Revista Turismo em Análise*, *30*(3), 516-538.
Mayzlin, D., Dover, Y., & Chevalier, J., (2014). Promotional reviews: An empirical investigation of online review manipulation. *The American Economic Review* 104(8): 2421-2455.
Milliot, J. (2018). Sales of Print Books Increased, Slightly, in 2017, *Publisher Weekly.com,* 4 January, 2018.
Mudambi, M. S. & Schuff, D. (2010). What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. *Management Information Systems Quarterly* 34(1): 185-200.
Park, S., & Nicolau, J. (2014). Asymmetric Effects of Online Consumer Reviews. *Annals of Tourism Research* 50(C): 67-83.
Raju, G. A. & Joseph, D. (2017). An Empirical Investigation of Online Review Diagnosticity. *International Journal of Business Information Systems* 25(3): 319-335.
Ramrattan, L., & Szenberg, M. (2016). *Revolutions in Book Publishing.* Palgrave MacMillan.
Rietsche, R., Frei, D., Stoeckli, E. et al. (2019). *Not All Reviews Are Equal - A Literature Review On Online Review Helpfulness.* In: Proceedings of the 27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Sweden, June 8-14, 2019. ISBN 978-1-7336325-0-8 Research Papers. Available at: https://aisel.aisnet.org/ecis2019_rp/58
Ross, P. (2010). Is There an Expertise of Production? The Case of New Media Producers. *New Media & Society* 13(6): 912–928
Sparks, B. A. & Browning, V. (2011). The impact of online reviews on hotel booking intentions and perception of trust. *Tourism Management* 32(6): 1310–1323.
Stigler, G. (1961). The Economics of Information. *Journal of Political Economy* 69(3): 213–25.

Statista (2022). Book purchases by store brand in the UK. Available at: https://www.statista.com/forecasts/997836/book-purchases-by-store-brand-in-the-uk

Tadelis, S. (2015). The Economics of Reputation and Feedback Systems in E-Commerce Marketplace. *IEEE Internet Computing* 20(1): 12- 19.

Utz, S. (2009). Egoboo vs. altruism: the role of reputation in online consumer communities. *New Media & Society* 11(3): 357-374.

Vermeulen, I. E. & Seegers, D. (2009). Tried and tested: the impact of online hotel reviews on consumer consideration. *Tourism Management* 30(1): 123–127.

Williams, N. L., Inversini, A., Ferdinand, N. et al. (2017). Destination eWOM: A macro and meso network approach?. *Annals of tourism research 64*: 87-101.

Wu, P. F., Van der Heijden, H., & Korfiatis, N. T. (2011). *The Influences of Negativity and Review Quality on the Helpfulness of Online Reviews*. In: Thirty Second International Conference on Information Systems, Shanghai 2011. Available at: http://ssrn.com/abstract=1937664.