University of Antwerp

Faculty of Business and Economics

Department of Marketing

# *Online reviews and how to manage them:* Effects of eWOM and Webcare on consumer responses and business performance

**Ana Isabel L. LOPES**

Proefschrift voorgedragen tot het behalen van de graad van Doctor in de Toegepaste Economische Wetenschappen

Universiteit Antwerpen

**Promotors:**

**Prof. Dr. Nathalie Dens**

**Prof. Dr. Patrick De Pelsmacker**

Antwerp, 2021

ISBN: 9789057286940
Deposit number: D/2021/12.293/15

# Doctoral Jury

**Prof. Dr. Patrick De Pelsmacker (promotor)**

University of Antwerp

**Prof. Dr. Nathalie Dens (promotor)**

University of Antwerp

**Prof. Dr. Karolien Poels (chair)**

University of Antwerp

**Prof. Dr. Martin Eisend**

European University Viadrina

**Prof. Dr. Edward C. Malthouse**

Northwestern University

**Prof. Dr. Ewa Masłowska**

University of Illinois

**Prof. Dr. Charlotte van Hooijdonk**

Utrecht University

# Acknowledgements

So many people deserve their place in this section. For a person that said for years that “doing a PhD is not for me”, I had quite a happy journey! The reason for that is that I am a fortunate person that had the chance of meeting many great friends, colleagues and mentors during my PhD journey. I will do my best to live up to them all in the next paragraphs.

First and foremost, I thank Prof. Dr. Nathalie Dens and Prof. Dr. Patrick De Pelsmacker for all the guidance, support, feedback and opportunities they provided me from the beginning. I am grateful that they chose me for the position on the Online Reviews project that gave me an enriching doctoral experience. Throughout the years, we had many meetings and talks and in all of them I learned something new. Maybe it was because I knew too little but, certainly, it was because you both know so much. Even during my moments of stubbornness (and I know I had a few…) it was enlightening for me to learn your perspectives. It is incredibly fortunate of me to end up with two respectable, approachable and well-humored supervisors as the two of you. Being supervised by you, Patrick and Nathalie, was a fantastic experience for me and I will forever be grateful for it! I truly hope that we can keep working together for many years and that if, one day, I supervise PhD students I can do it as well as you did with me.

I would also like to thank the other members of my doctoral committee, Prof. Dr. Karolien Poels and Prof. Dr. Martin Eisend, for the interest and constructive feedback on my project over the years. I learned a lot from our interactions and I hope that I have many chances of sharing my work with you over the years. I am also very thankful to the other members of my doctoral jury, Prof. Dr. Charlotte Hooijdonk, Prof. Dr. Ewa Maslowska and Prof. Dr. Edward Malthouse, for their availability and interest in being part of this journey. It is a true privilege to have a panel of academics that I respect so much evaluating my work.

A special word of appreciation has to be devoted to Prof. Malthouse. Since 2019 you have been an incredible mentor from whom I learned more than I thought I could in 10 years. I will always be grateful for the time you devoted working with me and teaching me without ever making me feel like my questions or issues were an inconvenience. I respect you for both being an incredible academic and a fantastic person. I would also like to acknowledge the Spiegel

Research Center and its members for agreeing in having me a visiting scholar and for welcoming me so well during my stays.

As I mentioned in the beginning, so many people helped me to be here today and to be writing these words. One of the persons that helped me the most was Mr. Christian Holthof. Mr. Holthof facilitated the access to the data that was used in the last chapter of this dissertation and was of invaluable help to explain to me the hospitality industry in general and the Antwerp context in particular. Thanks to him, I had the chance to get in contact with the Antwerp Hotel Association (AHA), another important actor in the story of this dissertation. I am thankful to AHA for their interest in my research and for facilitating the contacts of hotel managers that helped gather the booking data for the last chapter of this thesis. I cannot identify the hotels and managers that collaborated with me but I am very thankful to all. Finally, I have to thank specially Mr. Filip Goorden for his help in contacting all the managers and for inviting me to present my work at the 2019 AHA meeting.

I am so thankful to all (yes, to all!) the people from the Marketing Department of the University of Antwerp. To Prof. Annouk Lievens, I thank all I have learned from working with her and watching her organize and teach her course. What an example of rigor and empathy! To Prof. Ingrid Moons, I thank the constant good energy. To Yana, Cristian, Freya, Kristien and Bram, I foremost thank you for your friendship. You are fantastic colleagues and amazing people that, from start to end, made this journey easier. I would have been lucky to find one like you, but I must have done something right in my previous life to find you all! I would write a special message for each one of you, but that might get too personal and weird for this context… I will let you know later, when I can drink again and the wine allows me to speak 'the language of emotion'. As if this was not enough, other people that left along the way or got on board later on were also very important in this journey. Liesbet, Carmem, Elke, Maxime, Anne, Clara, Jana, Lorens, Lore, Wolfgang, Georg: thank you for the nice talks that allowed me to 'get the load off', for the laughs and for everything. Special thanks to Anne, who helped me with the intercoder reliability for one of the projects in this dissertation. I hope one day I can reciprocate the favor! Thank you Mona, Nele and Elisa for all the support as secretaries of the department and for always answering patiently to the same questions I ask you all the time! And thank you to all the others that in other instances and for shorter periods made my PhD journey so joyful!

There are many people that through their competence and extreme niceness made this PhD journey an easy ride. These people are all of those that work at the administration of Faculty of Business and Economics and that I had the pleasure to meet and work with. Thank you! I will not write up names, as I am too afraid of forgetting someone, and that would be certainly unfair!

I am also thankful to my many friends and family members in Portugal that were always so supportive during the PhD process. Aos meus queridos pais, António e Lurdes, agradeço todas as oportunidades que me deram na vida. De onde vim não é fácil chegar onde estou. Aqui estou por mérito meu e graças a vocês que, apesar de nem sempre compreenderem as minhas escolhas, sempre me deixaram ser livre. Mesmo quando eu sei que custou um bocado: estar longe é difícil mas sempre é mais fácil quando nos temos uns aos outros no coração. To Marcos and Tânia, my parents' other children, my gratitude. You 'hold the boat' while I venture around here. As for my friends from 'home', it is difficult to list and thank them all, which in itself is an excellent indication. Catarina, Marta, Sónia, Daniela, Deborah, Pedro, Fábio… only we know how much the long and frequent conversations we have (often daily) keep us sane. Thank you for everything!

To Moriarty, thank you very meowch. Although some people might say that cats cannot read, they certainly never met Professor Moriarty's intellectual portent. A cat, a friend, a companion, the boss. Life would be much duller without him. To João, I am thankful for everything. It is thanks to having you in my life that I am who I am today, happy and fulfilled. So that my thanks to you do not become a second thesis, I will only refer to your contributions to this achievement of mine. In fact, it is because of you that I decided to go for a PhD. It was by seeing how much you love research and the panache with which you are committed to your projects that I decided to give the academy a chance. It was thanks to how much you make me believe in myself, because you also believe in me, that I knew I was capable. Whenever I lacked strength or motivation, you were always there to help me and remember me that I could do it. Besides all the emotional support, I am also very thankful to all the support you gave me with my coding tasks and for teaching me everything I know about BERT. From the bottom of my heart, thank you. To Sara, who came to turn my life upside down, which is how life is worth living! I hope that when you come to read this, many years from now, you will know that you were also part of this moment, because you already lived in me, in my womb and in my heart, where you will always live.

# Table of contents

# List of Tables

# List of Figures

# 1. Introduction

*"One thing I love about customers is that they are divinely discontent. Their expectations are never static – they go up. It's human nature. We didn't ascend from our hunter-gatherer days by being satisfied. People have a voracious appetite for a better way, and yesterday's 'wow' quickly becomes today's 'ordinary'. I see that cycle of improvement happening at a faster rate than ever before. It may be because customers have such easy access to more information than ever before – in only a few seconds and with a couple taps on their phones, customers can read reviews, compare prices from multiple retailers, see whether something's in stock, find out how fast it will ship or be available for pick-up, and more. These examples are from retail, but I sense that the same customer empowerment phenomenon is happening broadly across everything we do at Amazon and most other industries as well. You cannot rest on your laurels in this world. Customers won't have it."*

— Jeff Bezos[1]

# Context and relevance

Traditional marketing frameworks, such as the hierarchy-of-effects models, have long dominated the conceptualization of the mechanisms by means of which consumers can be persuaded to become (loyal) customers of brands. In those models, it is traditionally assumed that potential customers go through a number of decision making stages before they buy a product and become loyal customers (De Pelsmacker et al., 2021). The basic idea is that marketers should 'guide' potential consumers through these stages by delivering the right message and activation in each stage of the process. In other words, marketers should steer potential consumers through what is called the 'Marketing Funnel': building brand awareness, brand knowledge, brand attitude, consideration, familiarity, consideration, trial, and, finally, loyalty (Stankevich, 2017). The origin of these frameworks is the AIDA model, developed in 1902 (Attention, Interest, Desire, Action). Later on, alternative conceptualizations have been proposed that provide a framework for the consecutive steps in a marketing (communication) plan that reflect this logic, such as the Lavidge and Steiner (1961) model and the DAGMAR model (De Pelsmacker et al., 2021). In today's increasingly online marketing context, this basic idea has been conceptualized as ToFU-MoFu-BoFU, which has become a leading principle of many online advertising campaigns (De Pelsmacker et al., 2021). When consumers are at the top of the funnel (ToFU), they get to know the product or the company. In the middle of the funnel (MoFU), awareness becomes interest; consumers take action to contact the company and start to consider buying the product or the brand. Finally, at the bottom of the funnel

[1] In a 2019 letter to shareowners marking the 20th anniversary of Amazon: https://www.sec.gov/Archives/edgar/data/1018724/000119312518121161/d456916dex991.htm

(BoFU), buyers become (loyal) customers or even advocates of the product or the brand. In each stage, marketers provide incentives for customers to move towards the bottom of the funnel.

However, in the online context, brand communication has changed fundamentally in two ways. First, contrary to traditional offline marketing, (potential) consumers online take an active part in brand communication in every stage of the funnel by engaging with branded content. By doing so, consumers generate what is called 'earned media': they react to these branded messages by liking, sharing, and creating brand-related messages (consumer-generated content) for free, leading to organic reach of other consumers and thereby influencing others in all stages of the funnel. Second, marketers now actively engage with consumers by interacting with them in every stage of the funnel (two-way interaction). Brands are now built jointly by marketers and consumers. This new paradigm is picked up in models such as the Consumer Engagement Engine that also consider today's digital ecosystem, describing it as working like an engine, where brands and consumers synergistically interact with each other (Collinger et al., 2011). Another metaphor used to illustrate these brand-dialogue behaviors is the "pinball machine" (Hennig-Thurau et al., 2013; Hennig-Thurau et al., 2010). Before consumers had access to new media that allow them to broadly disseminate their views on products and brands, marketers' activity resembled playing bowling: they would use mass media (alley), throw a ball (message) towards pins (consumers), hoping to touch as many as possible. Nowadays, this interaction is more like a pinball game: marketers drop a ball (message) in the pinball machine (new media), and try to keep it in the game as long as possible by operating the flippers (consumers). However, like the ball in a pinball machine, this message can be thrown back and forth by bumpers, kickers, and slingshots (or consumers portraying their own views on the product or brand), changing the direction of the message constantly. This 'pinball game' shows that the new marketplace rewards participatory, sincere, and less directive marketing styles (Deighton & Kornfeld, 2009), reinforcing the importance of understanding and managing earned media.

One particular form through which consumers actively engage in communication about a brand and create earned media is word-of-mouth (WOM). Since the first time "word-of-mouth" was mentioned in an academic publication in the '50s by Brooks Jr (1957), much research has been conducted on the topic, with more recent efforts focusing on **eWOM (electronic Word of Mouth)** (Dellarocas, 2003; Hennig-Thurau et al., 2004). A recent conceptualization of eWOM defines it as "consumer-generated, consumption-related communication that employs digital tools and is directed primarily to other consumers" (Rosario et al., 2020, p. 427). Online reviews, i.e., online product evaluations by users or experts, are commonly considered a form of eWOM (e.g., Zhang et al., 2010). Therefore, in this thesis, the terms eWOM and online reviews are used interchangeably.

Online reviews are a form of consumer engagement, including consumer opinions voiced through online review sites (e.g., TripAdvisor), (micro-) blogging platforms (e.g., Twitter), video sharing sites (e.g., YouTube), social networking sites (e.g., Facebook) or even on the seller's website (e.g., Amazon). Sharing our opinion about a product, service or brand is part of our experience as consumers. In 2018 alone, Google received 30.1 million reviews, Booking.com generated 28.3 million reviews, and TripAdvisor 11.3 million (WiT, 2019). Yelp counted nearly 150 million business reviews, attracting 186 million users per month on various devices (Capoccia, 2018). Online reviews are, therefore, an essential part of businesses' day-to-day interactions with consumers. Previous research shows that online reviews are gaining territory in influencing consumers' decisions (Keller, 2007). A survey of young U.K. adults shows that, when asked about whom they would trust when choosing a new brand or product, people trust first recommendations by friends (42.8%), followed by online reviews (30.7%), with a minority of the respondents (9.1%) claiming to trust ads for this purpose (Majors, 2020). The reason for this is that eWOM is seen as more reliable than marketer-initiated communications because it is perceived as having passed through an 'unbiased filter' that is the other consumers' experiences (Allsop et al., 2007). The influence of online reviews on businesses is, therefore, undeniable. Previous research finds that eWOM can be integrated into traditional marketing frameworks, for instance, that it influences consumers at every stage of the marketing funnel (Colicev et al., 2019).

The participatory environment where consumers talk publicly online about their experiences shifts the marketing focus from one-way communication to two-way communication. Indeed, digital media empower consumers to communicate easily with and about organizations and brands, creating the need for organizations to manage these communications (Deighton & Kornfeld, 2009; Kozinets et al., 2010; Vargo & Lusch, 2004; Williams & Buttle, 2011). These brand-dialogue behaviors are part of the customer engagement ecosystem, where each action causes a reaction of not only the intended recipient of the message but the whole ecosystem (Maslowska et al., 2016; Santini et al., 2020). The management of these communications is called webcare. **Webcare** is defined as the act of engaging in online communication to address client feedback (Edwards & de Kool, 2015) and has been shown to affect consumers' attitudes, intentions and behavior, and company performance (e.g., Sheng et al., 2019; Xie et al., 2016). While much research on webcare focuses on mitigating the effects of NWOM (negative eWOM) (Dens et al., 2015; Van Noort & Willemsen, 2012), webcare is, in fact, an integrative organizational tool combining customer care, public relations, and marketing that can broadly be used to increase consumer engagement (Edwards & de Kool, 2015; Schamari & Schaefers, 2015; Van Noort et al., 2015). Therefore, webcare is a crucial form of consumer feedback management and a crucial component of a brand marketing strategy. A response - or webcare - strategy refers to the characteristics of the answer businesses provide to reply to online reviews (e.g., apologize for mistakes; always reply within 24 hours, …).

eWOM is thus a dynamic process that involves many aspects, from the product or service being reviewed, to the person writing the review, the bystanders reading the reviews, the brands that manage the reviews by responding to them, and again the reviewers and bystanders that read the brand responses. The organizing framework proposed by Rosario et al. (2020), and seen in Figure 1.1, portraits the roles of the sender (or reviewer) and receiver (or bystander) of eWOM.

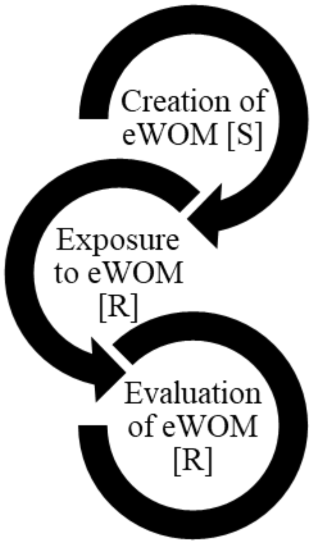



*Figure 1.1. Organizing framework*

*Notes: [S] denotes eWOM sender; [R] denotes eWOM receiver; adapted from Rosario et al. (2020)*

eWOM and webcare have gotten the attention of academics and practitioners for its influence on consumers and, consequently, on businesses. Therefore, these topics require extensive analysis and reflection by both parties (researchers and marketers) to ensure that its effects are well understood. Considering the relevance of eWOM and webcare, this thesis will have two foci:

1. **the influence of different eWOM characteristics on consumers' responses; and**
2. **the effects of webcare strategies on consumers' responses and business performance.**

Over the next section of this introduction, we will present relevant literature leading to the concrete research objectives for each of the two overarching foci on this thesis.

## The effects of eWOM characteristics on consumers' responses

Scholars and practitioners alike consider eWOM important for its influence on consumer decision-making (Verma & Yadav, 2021). Reports show that 93% of consumers say that online reviews influenced their purchase decision (Podium, 2017). In fact, previous research shows that interpersonal influence is a determinant of eWOM production and consumption (Park et al., 2011) and that, in turn, eWOM influences business performance (Xun & Guo, 2017). eWOM, and online reviews in particular, can instill trust (Evans et al., 2020), influence consumers' attitudes (Casado-Díaz et al., 2020), drive purchase decisions (Maslowska, Malthouse, & Viswanathan, 2017), impact sales (Li et al., 2020) and can also influence post-purchase evaluations (Liu, Jayawardhena, Osburg, et al., 2019). Online reviews are significant sources of information with estimations that they influence as much as 20–50% of consumer's online buying choices (Mathwick & Mosteller, 2017).

In light of the relevance of eWOM and online reviews for businesses, many scholars have mentioned a need for more research to better understand how several eWOM characteristics influence consumer's responses (e.g., Hair & Bond, 2018; Vermeer et al., 2019; Wang et al., 2015). There are several reasons behind the need for more research on eWOM, despite abundant research that already exists on the topic (see Rosario et al., 2020; Verma & Yadav, 2021 for systematic reviews on what was already published on eWOM). Perhaps the most important one is related to the diversity of content, format, and shape that eWOM and online reviews can assume. Online reviews can differ in many ways depending on the platform, from structure (i.e., some might demand that reviewer leaves a message besides only rating a product or service) to access (i.e., only customers with verified purchases can leave reviews). Besides, the content of the review, such as the review text, depends on a number of aspects endogenous (e.g., motivation to post as studied by Mathwick and Mosteller (2017)) and exogenous to the reviewer (e.g., the presence of other review elements, such as the overall star rating, in the moment of writing a review by Askalidis et al. (2017)).

It is crucial to understand how review readers process and respond to these different types of reviews. The way they do depends on a vast panoply of aspects, either related to the review (e.g., message content or style as studied by Schindler and Bickart (2012)) or the reader (e.g., country of origin as studied by Fong and Burton (2008)). This diverse range of angles from which eWOM can be considered justifies the need for more studies on the topic. Understanding how consumers respond to reviews with varying characteristics is especially important given the rise of automation of online Customer Relationship Management (CRM) (i.e., chatbots), which is nowadays a reality or an aspiration for many brands (Li et al., 2021; Liebrecht & van Hooijdonk, 2019). Identifying relevant and influential eWOM, based on online review characteristics and how consumers respond to them, is therefore imperative for brands aiming to implement strategies, manual or automated, to manage the online relationship with their consumers. Therefore, **the first main objective of this thesis is to find which online review cues are important to determine the credibility and helpfulness of a review and how the composition and content of a set of reviews affect consumers' responses.** The first two empirical chapters of the thesis address this objective.

### *Which review cues determine review usefulness and credibility?*

Previous research has explored the influence of various online review characteristics or cues on consumer perceptions of review helpfulness (e.g., Li et al., 2019; Yang et al., 2017) and credibility (e.g., Moran & Muzellec, 2017). It shows an increasing volume of fake online reviews (Y. Wu et al., 2020), which results in review skepticism (Zhang et al., 2016). Therefore, it is essential to look into the characteristics that make consumers perceive a review as helpful and credible for consumers. This is the objective of chapter 2.

Previous research has explored consumer responses towards, and more particularly the perception of usefulness and credibility of, amongst others, argument strength (Clare et al., 2018; Filieri, Hofacker, et al., 2018), review sidedness (e.g., Mayweg-Paus & Jucks, 2018; Pentina et al., 2018), writing quality (e.g., Clare et al., 2018; Schindler & Bickart, 2012), number of arguments (e.g., Chua & Banerjee, 2015; Thomas et al., 2019), number of reviews (e.g., Thomas et al., 2019), helpfulness votes (e.g., Cheung & Lee, 2008) and summary review ratings (e.g., Chevalier & Mayzlin, 2006; De Pelsmacker, Van Tilburg, et al., 2018).

However, the literature on these review characteristics often does not present a clear direction for the influence of all the cues on our outcome variables in each other's presence, the way that people experience them in practice. Also, most previous research is experimental and only studies a limited number of cues simultaneously. By studying them all together, we enhance the realism of scenarios presented to the readers. Therefore, the first research question that we aim to answer in chapter 2 is:

*RQ1: What is the relative importance of argument strength, sidedness of the message, writing quality, number of arguments, number of reviews, rated review usefulness, and summary review rating, for perceived review credibility and usefulness?*

The way these online review cues are processed may depend on the level of involvement with the reviewed product. We use the Elaboration Likelihood Model (ELM) (Petty & Cacioppo, 1984; Richard E. Petty & John T. Cacioppo, 1986) to explore how relatively important online review cues are for lowly and highly involved consumers. The ELM posits that highly involved consumers will process a message centrally, using central cues, while lowly involved consumers will process it peripherally, using peripheral cues (Baek et al., 2012; Filieri, Hofacker, et al., 2018; Thomas et al., 2019). Central cues are related to the content of the message (Baek et al., 2012), in the present study, argument strength, and sidedness of the message. Peripheral cues in an online review are non-content factors that do not require a lot of processing effort. Writing quality, the number of reviews, rated review usefulness, and summary review rating can be considered as peripheral cues. The role of the number of arguments is unclear: reviews with more arguments can be processed centrally as they provide more opportunities for consumers to elaborate on the message (Kim et al., 2018); however, readers may use multiple arguments as an indication of the amount of available information (i.e., a peripheral cue) (Richard E. Petty & John T. Cacioppo, 1986). The second research question of this chapter is:

*RQ2: What is the relative importance of the writing quality, number of reviews, rated review usefulness and summary review rating, number of arguments, argument strength and sidedness for highly involved individuals and lowly involved ones when evaluating review usefulness and credibility?*

Considering that we aimed to study the relative importance of the review cues in the presence of each other, a conjoint analysis was used for this study. Using a balanced orthogonal design, we generated eight cards that correspond to individual reviews and measured how consumers rate their usefulness and credibility. This study provides a comprehensive test of how consumers perceive online reviews, as it is the first to simultaneously investigate a large set of cues using conjoint analysis. Using a conjoint analysis allows for the implicit valuation (utility) of the individual cues, revealing the cues' relative importance for usefulness and credibility in a setting that comes close to a real-life context (Janssens et al., 2008; Levy, 1995; Rhee et al., 2016). Besides, insights of the ELM are used to understand how the relative importance of cues differs depending on the level of review readers' product category involvement.

### *The effect of review set valence, review attribute importance and review argument repetition on purchase intention*

Chapter 3 builds upon chapter 2 by studying cues related to the review arguments, which we found to be the most important review element. In this chapter, we look into the role of the content and number of the arguments by studying how varying reviews' argument importance, and repetition influence consumers' intentions, depending on the ratio of positive and negative reviews available.

One of the commonly studied topics in previous literature on eWOM is the role of review valence (for a meta-analysis, see Purnawirawan et al., 2015). The valence of a set of reviews can be either positive, neutral, or negative depending on the ratio of positive and negative individual reviews (Purnawirawan, De Pelsmacker, et al., 2012). The effect of review set valance on review readers can be explained by the bandwagon effect (Lee et al., 2018; Sundar et al., 2008), which states that people tend to make choices based on a perceived trend without making judgments about the trend. Previous literature shows that negative reviews are commonly seen as more influential (e.g., Brunner et al., 2019; Lee et al., 2009). Therefore, products or services with a majority of negative reviews containing a minority of positive reviews would always be negatively evaluated. This effect is explained by the negativity bias, which claims that negative events are more salient and more influential than positive events (Rozin & Royzman, 2001). However, previous research showed that there are nuances to the negativity bias (e.g., Hair & Bond, 2018; Wu, 2013) which means that negatively valenced sets of reviews might not always lead to negative evaluations. The goal of chapter 3 is to explore two factors that might attenuate the influence predominantly negative review sets, namely the importance of arguments used and review argument repetition.

Relevant reviews, i.e., reviews about important attributes, are perceived by review readers as more diagnostic (Filieri, Hofacker, et al., 2018). Since negative reviews are considered more influential than positive ones, one would expect that readers negatively evaluate neutral and negative review sets. However, it is not clear what would be the effect of having negative review sets where the positive reviews are about important attributes and the negative reviews about less important attributes, unbalancing the anticipated stronger influence of negative reviews. As far as we know, no previous research has yet addressed the research gap of how predominantly negative sets of reviews are processed when the positive reviews in the set are about important attributes and the negative reviews are about less important attributes. Therefore, the first research question we aim to answer in this chapter is:

*RQ1: How do varying ratios of a majority of negative reviews about less important attributes and a minority of positive reviews about more important attributes influence consumers staying intention at a hotel, and where is the 'tipping point' at which a number of positive reviews in a predominantly negative review set leads to a positive hotel booking intention?*

Another aspect that can influence the perceived positivity or negativity of a set of reviews is argument repetition. The effects of repeated message exposure, containing the same or different arguments, are frequently studied in advertising (Chang, 2009), but the effect of repetition in the context of online reviews is unexplored. There are arguments in favor of repeating a message to increase its believability (e.g., Dechêne et al., 2010) and others in favor of diversifying the message to increase information utility (e.g., Zhang et al., 2014). According to the truth effect, repeating arguments (i.e., reviews pertaining to the same attributes as the other reviews in the set) increases participants' subjective judgments of a statement's truth (Dechêne et al., 2010; Roggeveen & Johar, 2002). Previous research found that message repetition increases persuasion (Cacioppo & Petty, 1989). In light of this effect, we would predict a positive effect of attribute repetition on hotel staying intention. However, the repetition-variation hypothesis in advertising states that providing different arguments increases persuasion by increasing issue-relevant thoughts or by serving as a simple acceptance cue (Calder et al., 1974; Petty & Cacioppo, 1984). The persuasion literature also shows that messages with more arguments are more persuasive as they provide confidence in decision-making (Srivastava & Kalro, 2019). Increasing the number of arguments across a set of reviews will increase the amount of available information to make their judgment, as expected based on the accessibility-diagnosticity theory (Herr et al., 1991). Information diagnosticity refers to the ability of the information in online reviews to enable readers to learn and evaluate the quality and performance of services before purchasing them (Filieri, Hofacker, et al., 2018). The greater the information diagnosticity of reviews, the higher will be the influence on purchase intentions (Filieri, 2015; Herr et al., 1991). Therefore, there are divergent views on how repeating the arguments in reviews and using different arguments would affect hotel staying intention. The second research question of this chapter is:

*RQ2: How does having (multiple) positive reviews about the same attribute versus positive reviews about different attributes moderate the ratio of positive reviews on readers' intention to stay at a hotel?*

Considering these research questions, chapter 3 comprises a 4 (ratio of positive reviews about important attributes to negative reviews about less important attributes) x 2 (attribute repetition vs. different attributes for the positive reviews) between-subjects full factorial design experiment. We tested the effects of different valence ratios on the intention to stay in a hotel and how argument repetition moderates this effect. This study sheds light on the nuances of relevant theories and effects commonly used in the eWOM literature, such as the negativity bias, the bandwagon effect, the truth effect, and the repetition-variation hypothesis.

## The effects of webcare on consumer responses and business performance

Having studied how several review characteristics are perceived by review readers, the second part of the thesis focus on another very relevant aspect of eWOM: the effects of webcare – managerial responses - on consumer responses and business performance. As mentioned, brands frequently have to act to mitigate the negative effect of negative reviews in particular. Previous studies on complaint handling show that perceived justice (with how the complaint was handled and with the services and service provider as a whole) plays a role in restoring customer satisfaction (for a meta-analysis on this topic, see Gelbrich & Roschk, 2011). Similarly to perceived justice, fairness perceptions are highly relevant when managing online reviews, especially negative ones. Therefore, webcare should stimulate perceptions of fairness to reviewers and bystanders, making the question of how to respond to eWOM an important one (Van Noort et al., 2015). However, it is very common that businesses struggle to know which guidelines to implement. In 2020, news outlets were flooded with the story of a man who was sued and faced jail time in Thailand after posting negative reviews of a hotel he stayed in[2]. This resulted in a public backlash against the hotel, with TripAdvisor - the platform where the reviews were originally posted - issuing warnings to users about the Thai hotel that legally prosecuted the reviewer[3]. The hotel ended up dropping charges against the reviewer but demanded that he issues a public apology for his comments[4]. Although this is a drastic example of management of online reviews, it portraits the hardship and the possible nefarious consequences of having to deal with consumer feedback.

With the increasing volume of eWOM, it is crucial for organizations to know how to invest their efforts in webcare to achieve positive business results (Schamari & Schaefers, 2015; Williams & Buttle, 2011). In fact, when surveying consumers about their review habits, 96% of reviewers claim to read businesses' responses to their reviews – with 40% saying they 'always' read the responses (Murphy, 2020). However, organizations frequently struggle to know which webcare strategies to employ to yield the best results (Van Noort et al., 2015). **The second main objective of this thesis is to build up a framework for webcare and to explore the effects of the different webcare strategies on business performance**.

### *The effectiveness of webcare: A literature review and research agenda*

Over the last decades, dozens of papers were published on a track of eWOM research that studies the effects of webcare. Previous studies on webcare tackle different aspects of responding to eWOM. Studies have focused on the effects of responding versus not responding to online reviews (e.g., Proserpio & Zervas, 2017; Van Noort & Willemsen, 2012; Wang & Chaudhry, 2018), who responds to the review (e.g., Wang & Chaudhry, 2018), the timing of this response (e.g., Stevens et al., 2018), and what is being said, for instance in what tone of voice (e.g., Van Noort & Willemsen, 2012) or the response length (e.g., Javornik et al., 2020). It will also matter whether the original review was positive or negative (Dens et al., 2015). Most previous research on webcare focuses on responses to negative WOM, mainly classifying

[2] https://www.bbc.com/news/world-asia-54335789, accessed April 19th 2021.
[3] https://www.bbc.com/news/world-asia-54914768, accessed April 19th 2021.
[4] https://www.theguardian.com/world/2020/oct/09/thailand-hotel-american-guest-jail-bad-reviews, accessed April 19th 2021.

them as either accommodative – complaisant and comprising corrective action, compensation and/or mortification – or defensive – denial and evasion of responsibility (Einwiller & Steilen, 2015). One might say that the Thai hotel in the example in the introduction that sued the client that wrote a bad review is the most extreme example of defensive behavior.

Overall, previous research is often inconclusive about adequate webcare strategies and the extent to which they lead to positive outcomes. Therefore, the research questions of this chapter are:

*RQ1: What are the webcare strategies that yield the most positive (and negative) results for businesses?*

*RQ2: Which webcare strategies show inconsistent outcomes and are under-researched requiring further studies?*

In order to answer these research questions, a literature review was conducted. The review comprises 70 papers published under the keywords of *managerial responses to reviews*, *webcare*, *service failure* and *service recovery*, *complaint handling* and *complaint recovery*, *online communities* and *online firestorms*, *response strategies to online reviews*, *service intervention*, *reputation management* and *customer care*, from 2000 until 2020, and that are related to managerial responses to eWOM or online reviews. The literature review proposes a comprehensive framework of webcare strategies and discussing findings from previous studies. This allows us to (1) identify potential generalizations from the findings of previous research, (2) discuss possible explanations for inconsistencies that need to be further explored, and (3) identify the under-researched areas with respect to the managerial responses to online reviews. Such framework would lead to a better understanding of the current state of knowledge of the effect of webcare strategies, the knowledge gaps and how to close them and provide guidance to managers seeking to establish guidelines to manage and respond to electronic word-of-mouth.

### *Is webcare good for business? The effect of managerial response strategies to online reviews on business performance*

Chapter 4 identified the webcare strategies more commonly used in practice and what previous research found about their effects. Building upon the findings from chapter 4, chapter 5 focuses on solving the contradictory findings and lack of research of multiple webcare strategies.

As previously mentioned, it is crucial for organizations to know how to invest their efforts in webcare to achieve positive business results (Schamari & Schaefers, 2015). Webcare is not only read by those who have written the review but also by bystanders who use the information to decide if they will book some hotel. According to the social learning theory, individuals learn from observing others' behaviors and/or the consequences of those behaviors (Bandura & McClelland, 1977). Bystanders learn from online reviews and managerial responses to them. This explains why bystanders are motivated to read webcare (Schamari & Schaefers, 2015), to help them build their attitude towards the brand (Weitzl & Hutzinger, 2017), develop trust perceptions (Ku et al., 2021), and make purchase decisions (Kim et al., 2016). The fifth chapter builds upon these insights to understand how different webcare strategies influence actual behavior by focusing on the bookings that hotels receive. Most previous research on webcare measures attitudes or purchase intentions, which does not necessarily reflect or predict actual buying behavior (Morwitz et al., 2007). Also, the results of previous studies are often inconsistent. These discrepant findings require further research, as they seem to indicate that there are aspects other than merely responding that determine the effects of webcare, for instance, the strategy used for responding. Therefore, this study aims to answer the question:

*RQ1: How do specific webcare strategies affect hotel bookings?*

More specifically, in chapter 5, we look into the effects of providing webcare, regardless of its content, on hotel bookings, building upon previous literature (e.g., Bhandari & Rodgers, 2018; Chen et al., 2019) to study the consequences of providing webcare versus not providing webcare. Looking into specific strategies, we first test the effects on future hotel bookings of having different intervenient responding to reviews (e.g., staff members versus managers) and using different styles (i.e., signing with own name versus with the hotel name). Second, we build upon previous research (e.g., Grégoire et al., 2015) and look into the effect of changing the channel of conversation to a private channel on future bookings. In this chapter, we also study strategies commonly referred to in the literature (e.g., Van Noort & Willemsen, 2012; Xie et al., 2017) that are related to the style of the response, such as the use of tailoring (versus generic) webcare and how using a conversational human tone affects future bookings. Considering that most previous research shows the importance of addressing negative reviews, this chapter also looks into the effects of several strategies typically directed at negative reviews. Although previous research finds that strategies such as showing gratitude and asking for more information are frequently used in practice (Einwiller & Steilen, 2015), there is a lack of research on how they affect business performance. Therefore, we study the effects of these strategies on hotel bookings. We then look into some commonly researched strategies whose effects on business performance are not yet clear, namely, the use of apologies (e.g., van Hooijdonk & Liebrecht, 2021), the offering of compensation (e.g., Rose & Blodgett, 2016), or the use of defensive statements (e.g., Weitzl & Hutzinger, 2019). Chapter 5 also tests several leading machine learning classifiers (support vector machines, boosted trees, random forests, naïve Bayes, and BERT - bidirectional encoder representations from transformers) in order to identify which of these classifiers performs best in classifying the different webcare strategies. Therefore, this chapter solves another gap in terms of the operationalization of studies on the topic of webcare.

## Structure of the dissertation

The remainder of this dissertation consists of three empirical chapters, a literature review, and a concluding chapter, as shown in Figure 1.2.

The concluding chapter summarizes the findings, discusses theoretical contributions and implications for practitioners, and offers suggestions for further research.

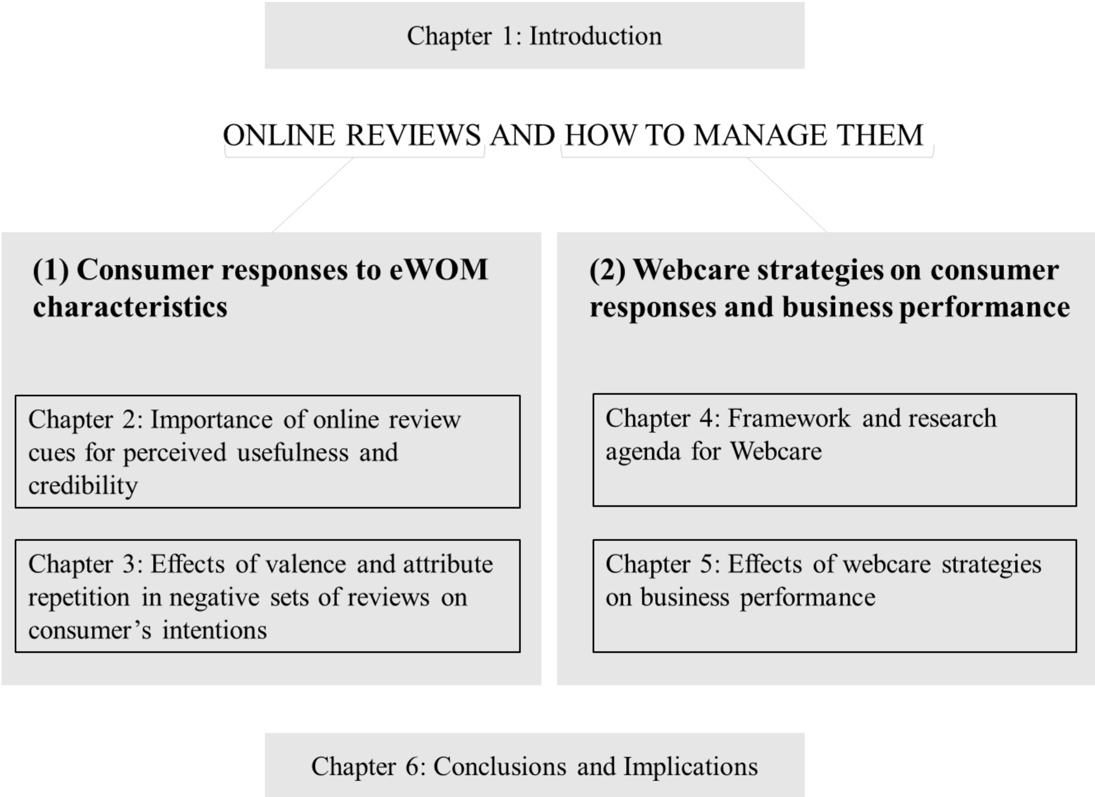



*Figure 1.2. Structure of the dissertation*

## Contributions

This dissertation contributes to the extant literature on eWOM and webcare in several ways. First, chapter 2 provides insights into how different review cues are perceived when presented together. Although they are commonly found in literature, they are usually studied in isolation or combined with just one or two other cues (e.g., De Pelsmacker, Van Tilburg, et al., 2018). By employing a conjoint analysis, this study allows us to study the relative importance of seven commonly used and studies reviews cues in the presence of each other. Managers can use the insights of our study to guide reviewers and optimize the usefulness and credibility of online reviews on their e-commerce sites.

Second, chapter 3 focuses on review arguments by challenging the commonly established conception that negativity will always overweight positivity, as expected by the negativity bias (Rozin & Royzman, 2001). This study nuances the bandwagon effect, negativity bias, and truth effect by studying if and how that negative review sets can be positively evaluated. This study also has implications for practitioners. Receiving little or no negative reviews is often considered an ideal scenario. However, brands frequently receive negative reviews, and our research can help managers to understand how these might undermine (or not) their business and how they can avoid this.

Third, chapter 4 provides a much-needed systematic overview of the literature of the past 20 years on webcare strategies, proposing a comprehensive framework of the different strategies and their possible brand outcomes and identifying consistencies, inconsistencies, and research gaps. This framework also guides practitioners based on what previous research tends to find leads to positive outcomes.

Fourth, chapter 5 builds upon some of the research questions raised in chapter 4 and on the effects found in previous literature (e.g., Dens et al., 2015; Sheng, 2019) and tests how different webcare strategies affect business performance in a hotel context, providing clear guidance for academics and practitioners on this topic. First, this study uses a machine learning approach to explore the effect of specific webcare strategies on actual bookings, while controlling for other factors that can affect business in the hospitality industry, such as seasonality. The second contribution is that this is the first comprehensive study that simultaneously examines the effects of several webcare strategies common in practice, for which prior research documents contradictory findings, on business performance, which is also under-researched as an outcome. Finally, this study also contributes to advances in webcare by developing automated machine learning tools and methods for coding webcare responses by testing several machine learning text classifiers.

# 2. Which cues influence the perceived usefulness and credibility of an online review? A conjoint analysis[5,6]

[5] This manuscript is published in Online Information Review:

Lopes, A. I., Dens, N., De Pelsmacker, P., & De Keyzer, F. (2020). Which cues influence the perceived usefulness and credibility of an online review? A conjoint analysis. *Online Information Review*, 45(1), 1-20. doi:10.1108/OIR-09-2019-0287

[6] An earlier version of this chapter was presented at the 2019 International Conference on Research in Advertising, Krems an der Donau, Austria

# Abstract

This article aims to assess the relative importance of the argument strength, argument sidedness, writing quality, number of arguments, rated review usefulness, summary review rating, and number of reviews in determining the perceived usefulness and credibility of an online review. Additionally, we use insights from the Elaboration Likelihood Model (ELM) to explore the effect of consumers' product category involvement on the cues' relative importance. A conjoint analysis (N= 287) is used to study the relative importance of the seven previously mentioned attributes. A balanced orthogonal design generated eight cards that correspond to individual reviews. Respondents scored all eight cards in random order for perceived usefulness and credibility. Overall, argument strength is the most important cue, while summary review rating and the number of reviews are the least important for perceived review usefulness and credibility. The number of arguments is more important for people who are more highly involved with the product while writing quality and rated review usefulness are relatively more important for the low involvement group. This study provides a comprehensive test of how consumers perceive online reviews, as it the first to our knowledge to simultaneously investigate a large set of cues using conjoint analysis. This method allows for the implicit valuation (utility) of the individual cues, revealing the cues' relative importance, in a setting that comes close to a real-life context. Besides, insights of the Elaboration Likelihood Model (ELM) are used to understand how the relative importance of cues differs depending on the level of review readers' product category involvement.

# Introduction

Online reviews, i.e. online product evaluations by users or experts, are significant sources of information for customers that influence as much as 20–50% of their online buying choices (Mathwick & Mosteller, 2017). Online reviews have attracted considerable attention from both marketers and academics (e.g., De Keyzer et al., 2017; Li et al., 2020; Tang et al., 2019). Online reviews can instill trust (Evans et al., 2020) and influence consumers' attitudes (Casado-Díaz et al., 2020). They impact sales (Li et al., 2020) and can also influence post-purchase evaluations (Liu, Jayawardhena, Osburg, et al., 2019).

Reviews may include different cues that help evaluate the product or service, such as the number of arguments, argument strength, argument sidedness (inclusion of positive and/or negative information), writing quality, rated review usefulness, summary review rating, and the number of available reviews. The Limited Capacity Model of Motivated Mediated Messages Processing (LC4MP) (Lang, 2000) and the cognitive load theory (Sweller, 1988) state that people have a limited cognitive capacity and cannot process more than a limited amount of information in a short time. Gottschalk and Mafael (2017) found that consumers selectively process online review cues. Therefore, it is important to understand which review cues have the greatest impact on review credibility and usefulness for consumers. Credibility and usefulness are commonly studied consumer responses because of the impact they exert on product and brand evaluations (Craciun & Moore, 2019). As consumers will want to avoid manipulated or biased online reviews, review credibility is an important determinant that affects whether consumers are persuaded by a reviewer's opinion (Grewal & Stephen, 2019). If a review is considered credible, the containing information is considered more valuable, is more often believed and accepted by the reader, and affects attitudes and behaviors (Thomas et al., 2019). Review usefulness refers to a measure of perceived value in the decision-making process (Siering et al., 2018) and is one of the most important determinants of information adoption (Ventre & Kolbe, 2020; Wang & Li, 2019; Wang et al., 2019). Review usefulness helps consumers deal with information overloads and facilitates their decision-making (Li et al., 2019). Therefore, it is important to understand what makes a review useful and how to extend customer access to such reviews.

Previous studies have explored the influence of various online review characteristics or cues (such as review sentiment, star rating, readability, length, and posting date) on consumer perceptions of review helpfulness (e.g., Li et al., 2019; Yang et al., 2017) and credibility (e.g., Moran & Muzellec, 2017). Much of this research is experimental, which allows causal relationships to be tested with a high degree of internal validity. At the same time, experiments suffer from the limitation that they can only manipulate and test a limited number of review characteristics at a time. Other studies draw upon the observation of textual details from databases of reviews obtained from websites such as Amazon, which use voting mechanisms asking readers about the extent to which a review was helpful (Hong et al., 2017). However, helpful voting mechanisms can be easily manipulated. In addition, the results of previous studies on the determinants of review credibility and helpfulness seem to contradict each other at times. Therefore, the first objective of this study is to investigate, by means of conjoint analysis, the relative importance of the number, strength, and sidedness of the arguments, review writing quality, the summary review rating, the rated review usefulness, the number of reviews, on consumers' perceptions of review usefulness and credibility. By using conjoint analysis as a methodology to study a diverse range of cues in a single comprehensive study, we present a new perspective on how these cues can influence the perceived credibility and helpfulness when presented together. Previous research recommends conjoint analysis as a useful method to understand the combined effects of multiple attributes and precisely analyze the relative importance of these attributes (i.e., Baek et al., 2006; Rhee et al., 2016). As stated by Levy (1995), a conjoint analysis is relevant to predict overall consumer preferences by considering the aggregated utility scores of a product. As such, it is widely used as a marketing research tool to predict consumer choices among multiple (product) attributes (Baek et al., 2006). This study exposes consumers to multi-attribute stimuli as they would see them in real life, based on which consumers establish credibility and usefulness assessments, enhancing the external validity of our results (Janssens et al., 2008). Yang et al. (2017) conducted a conjoint analysis of six heuristic review attributes related to the reviewer (e.g., reviewer location) or review itself (e.g., review length) to test how these affect review helpfulness. Their study, however, does not look into the effects of the review text (such as the number of arguments or argument strength), where text has previously been proven to be very important for review readers (De Pelsmacker, Dens, et al., 2018).

Besides looking at the relative importance of the selected review cues in general, we use the insights of the Elaboration Likelihood Model (ELM) to understand how these cues' relative importance differs depending on the level of review readers' product category involvement. As stated in the ELM (Richard E. Petty & John T. Cacioppo, 1986), consumers' involvement with the product influences how they process information. Highly involved individuals are more likely to process via the central route, paying attention to the arguments in a message (Richard E. Petty & John T. Cacioppo, 1986). In contrast, lower involvement triggers the peripheral route where persuasion is based on peripheral cues, such as the number of reviews (Park et al., 2007). The second objective of this study is to find out how the selected review cues differentially affect consumers' perceptions of usefulness and credibility depending on readers' level of involvement with the product. Argument strength and argument sidedness are commonly qualified as 'central' review elements (Filieri, Hofacker, et al., 2018), while the writing quality, rated review usefulness, summary review rating and number of reviews are usually considered as more peripheral cues (Filieri, Hofacker, et al., 2018; Park et al., 2007), and number of arguments can be considered either a central or a peripheral cue.

This study contributes to the literature on how online reviews are perceived and used by review readers to assess review usefulness and credibility by providing an integrative study of review cues through conjoint analysis, a method that takes real-life context into account, focusing on their relative importance. Very few studies have focused on the relative influence of multiple cues on the perceived credibility and helpfulness of the reviews. Brand managers can use the insights of our study to guide reviewers and to optimize the usefulness and credibility of online reviews on their e-commerce sites.

## Literature review

### *How do review cues influence review credibility and usefulness?*

We discuss seven review cues frequently encountered in practice and studied in academic research (e.g., Cheung et al., 2012; De Pelsmacker, Dens, et al., 2018; Zhang et al., 2014): argument strength, sidedness of the message, writing quality, number of arguments, number of reviews, rated review usefulness, and summary review rating.

#### *Argument strength*

The argument strength is the extent to which the message receiver perceives the argument as convincing or valid in supporting its position (Cheung et al., 2009). Schindler and Bickart (2005) explored the importance of the strength of review arguments in a qualitative study. They found that consumers will not readily believe the information in an online review if it does not contain sufficiently strong arguments about the product or service that they consider buying.

Thomas et al. (2019) found that argument quality, related to argument strength, is the primary factor affecting review credibility. Other research has proven that strong arguments serve as a significant predictor of the perceived usefulness of the review (Clare et al., 2018; Filieri, Hofacker, et al., 2018). For instance, Wang and Li (2019) found that information quality, also related to argument strength, is positively associated with the perceived usefulness of review websites. In summary, the strength or quality of arguments in a review seems to be a crucial element in influencing consumers' perceived review usefulness and credibility.

*Review sidedness*

The sidedness of a review refers to the presence or absence of (both) positive and negative information in a review. One-sided reviews are strictly positive or strictly negative, where two-sided reviews contain both positive and negative messages (Park et al., 2019). In line with what is typically documented in advertising research, two-sided messages are often deemed more credible because they provide a comprehensive overview on debated issues (Mayweg-Paus & Jucks, 2018). Park et al. (2019) found that two-sided reviews are only more credible than one-sided (positive) reviews for firm-sponsored reviews, and not for consumer-voluntary reviews. In contrast, one-sided messages can sometimes be perceived as more credible due to the confirmation bias in information processing (Metzger et al., 2020). Pentina et al. (2018) showed that (one-sided) positive reviews are perceived as more credible than two-sided reviews, while the difference between (one-sided) negative and two-sided reviews is not significant.

With respect to review helpfulness, prior research is also inconsistent. On the one hand, one-sided reviews offer readers a clear indication of what to do. Cao et al. (2011) found that reviews with extreme opinions (one-sided) receive more helpfulness votes than those with mixed or neutral opinions (two-sided). Pentina et al. (2018), too, showed that (one-sided) positive reviews are perceived as more helpful than two-sided reviews (although the difference between one-sided negative reviews and two-sided reviews was not significant). At the same time, Filieri, McLeay, et al. (2018) argued and found that two-sided reviews are more likely perceived as helpful because they better help readers to understand the strengths and weaknesses of a service and better evaluate whether it suits their needs. Considering the inconsistent findings, this study does not only assess the relative importance of review sidedness, but can also contribute to the debate on the direction of the effect of one-sided versus two-sided reviews.

*Writing quality*

The writing quality of an online review is related to factors such as spelling, structure, and grammar. Stylistic elements that may impair the clarity of a review, such as poor spelling and grammatical errors, cause the review to be perceived as less helpful (Schindler & Bickart, 2012; Wang et al., 2019) Poor grammar also makes readers question the competence of the author and dismiss the review's credibility as a result(Clare et al., 2018; Moran & Muzellec, 2017). A meta-analysis by Wang et al. (2019) showed that review readability (related to writing quality) is the most important factor in evaluating review helpfulness as it is directly related to the extent to which a review text is understood.

*Number of arguments*

The number of arguments refers to how many arguments are used in an online review. The more arguments a review has, the more comprehensive it is, making the message more complete and clear (Zhang et al., 2014). Previous studies have found a positive effect of the number of arguments in an online review on the perceived usefulness of the review (Chua & Banerjee, 2015). Reviews with many arguments are more helpful because they contain more details, which can help consumers with their purchase decision (Schindler & Bickart, 2012). According to Chua and Banerjee (2015), the number of arguments is important since reviews with substantial depth command a sense of adequacy and competence of the reviewer. In other words, a comprehensive number of arguments improves perceived source credibility, consequently enhancing the review's credibility.

*Number of reviews*

The number of reviews refers to how many online reviews are available for a particular product or service (Park & Kim, 2008). Zhang et al. (2014) state that the number of reviews is a helpful cue to assess the popularity of the product. Consumers are more likely to purchase products with many online reviews rather than with a few (Zhang et al., 2014). This is consistent with the consensus heuristic (Purnawirawan et al., 2014) which posits that people tend to consider that the majority's opinion is true. By seeing that a product is popular amongst others (because it has a high number of reviews), consumers may feel confident in trusting the choice of a big group of people, increasing the credibility and the perceived usefulness of the reviews (Purnawirawan et al., 2014). Thomas et al. (2019) point at the opposite direction, showing that review quantity has a negative impact on review credibility. Consumers might perceive a higher number of online reviews for a certain product or service as less credible if they suspect that companies have deceptively contributed to this multitude. In the present study, we will not only assess the relative importance of the number of reviews, but also the direction of its effect on perceived usefulness and credibility.

*Rated review usefulness*

The rated usefulness of an online review informs readers about how many (previous) users found the review useful (Kolomiiets et al., 2016). If many previous readers have indicated that they found a review useful, this will have a positive effect on the perceived usefulness of this review (Cheung et al., 2008). This influence may be explained by the bandwagon effect, which states that cues about others' behaviors guide our own decisions (Sundar et al., 2008). On the other hand, De Pelsmacker, Dens, et al. (2018) found that the rated review usefulness did not affect readers’ review impression (the extent to which the reader has a negative or positive impression from the review about the target object) and may be ignored when consumers have access to the review text. The rated review usefulness can also be relevant for readers to determine the perceived credibility of a review: the credibility of a review depends on how helpful it is perceived to be and vice-versa (Clare et al., 2018).

*Summary review rating*

In a review, consumers can sometimes summarize their overall appreciation of a product or a service in a summary score (e.g., from 1 to 5) or visual rating (e.g., from 1 star to 5 stars). Previous research has shown that summary review ratings have a strong influence on perceived review helpfulness and trustworthiness since this information is easy to process and allows an easy overall evaluation (Filieri, 2015). Also, the summary review star rating has proven to strongly influence e-tailer trustworthiness (Sebastianelli & Tamimi, 2018). The results of both Chevalier and Mayzlin (2006) and De Pelsmacker, Dens, et al. (2018) suggest that, in the presence of both review text and (star) ratings, review readers rely on the review text rather than on summary statistics such as ratings. That would imply that the effect of the summary review star rating on review helpfulness and credibility might be minimal. Considering these results, it is not clear what should be expected in terms of the relative importance of summary review star rating in the presence of other review cues such as argument strength.

In summary, previous research consistently suggests that cues such as the strength of the arguments, the number of arguments, and writing quality influence review credibility and usefulness positively. For cues such as summary review star rating and rated review usefulness, the results of previous studies are unclear regarding their effectiveness in the presence of other review elements, such as the review text. For still other cues, such as message sidedness and the number of reviews, the direction of their effect on perceived usefulness and credibility is unclear as prior studies document opposing effects. Besides, very few studies focused on the relative importance of several cues that are simultaneously present in a review, to determine perceived review credibility and helpfulness, and the methods used do not allow to study the relative importance of the cues in a format close to the real-life exposure to online reviews. No study so far has included the variety of the seven review cues that we included in the present study. Since the literature does not present a clear direction for the influence of all the cues on our outcome variables, we formulate the following research question:

*RQ1: What is the relative importance of argument strength, sidedness of the message, writing quality, number of arguments, number of reviews, rated review usefulness, and summary review rating, on perceived review credibility and usefulness?*

### *The moderating role of involvement*

Dual processing theories such as the Elaboration Likelihood Model (ELM) (Petty & Cacioppo, 1984; Richard E. Petty & John T. Cacioppo, 1986) explain how consumers process information in persuasive communication. According to the ELM, persuasion can occur through a central or a peripheral route (Petty & Cacioppo, 1984). When an individual processes a message via the central route, the information, and arguments present in the message are elaborated in-depth. Under these circumstances, the reader of the message will cognitively endeavor to process the available information and put more effort into evaluating the message, for instance, elaboration on the strength of the arguments. The peripheral route, on the other hand, implies less cognitive effort from the reader, or low elaboration. Individuals use simple signals or indicators, referred to as peripheral cues, to assess the message. For instance, in the case of online reviews, the average star rating of a product or service might serve as a peripheral cue (Baek et al., 2012).

The ELM has been applied to the study of online reviews to explain consumer cognitive processing of product reviews and evaluation of review messages, especially to understand the role of central and peripheral cues on consumers' decision-making processes (Baek et al., 2012;

Filieri, Hofacker, et al., 2018; Thomas et al., 2019). Central cues are related to the content of the message (Baek et al., 2012); in the present study, argument strength and sidedness of the message can be considered as central cues. Peripheral cues in an online review are non-content factors that do not require a lot of processing effort (Filieri, Hofacker, et al., 2018). Writing quality, the number of reviews, rated review usefulness, and summary review rating can be considered as peripheral cues. The role of the number of arguments is unclear. On the one hand, reviews with more arguments are more comprehensive, presenting the reader with meaningful extra content (Zhang et al., 2014) and longer reviews (i.e. with more arguments) provide more opportunities for consumers to elaborate on the message and its arguments and enhance counter-arguing (Kim et al., 2018). As a result they may be considered as central cues (Richard E. Petty & John T. Cacioppo, 1986). On the other hand, readers may use multiple arguments as a mere indication of the amount of available information, and consider them as a shortcut (i.e., a peripheral cue) (Richard E. Petty & John T. Cacioppo, 1986).

The use of the central or peripheral cues is determined by the consumer's motivation, ability, and opportunity to process the information. One of the determinants of elaboration motivation (and thus, the relative importance of central versus peripheral cues) is the review readers' degree of product category involvement. Readers who are less involved with the product category will more likely use the peripheral route. By using mental shortcuts (Richard E. Petty & John T. Cacioppo, 1986), the lowly involved individual will focus on easy to process non-content cues, such as star rating or the rated review usefulness. Indeed, Lee et al. (2008) found that low involvement readers tend to conform to the opinion expressed in the reviews regardless of the quality of the reviews, supporting the idea that they rely more on peripheral than on central cues. Similarly, Park et al. (2007) show that low-involvement readers are affected by the quantity (the "more-is-better" heuristic) rather than the quality of reviews.

In contrast, high involvement with a product will encourage readers to use the central route, in which a significant cognitive effort of the recipient is expected. Individuals with a high degree of product involvement are more likely to elaborately process and scrutinize the content of a review to evaluate the provided product information (Park & Lee, 2008). For example, De Pelsmacker, Dens, et al. (2018) found that the influence of (the valence of) the review text on evaluative responses is stronger for more highly involved people.

Considering the unclear role of some of these review characteristics in the presence of each other, we formulate the following research question:

*RQ2: What is the relative importance of the writing quality, number of reviews, rated review usefulness and summary review rating, number of arguments, argument strength and sidedness for highly involved individuals and lowly involved ones when evaluating review usefulness and credibility?*

## Method

### *Pre-test*

First, we conducted a pre-test to determine the product to be used in the main study. Because one of the purposes of the study is to test the moderating role of involvement, we wanted to select a product with a moderate level of and a substantial variation in involvement. We decided against the use of two products differing in involvement to avoid potential confounds due to the product itself. The product also had to be at least moderately appealing to consumers, to enhance the realism of the study (the idea being that people would never consult an online review for a product they have no intention of buying). In the pretest, respondents ($n = 16$) rated their product category involvement (3-item, De Keyzer et al., 2017$\alpha$=0.947) and purchase decision involvement (3-item, Dens & De Pelsmacker, 2010$\alpha$=0.85) on a 7-point semantic differential scale for 15 different products. The results indicated that a GPS is moderately involving ($M = 4.375$, $SD = 1.897$) and moderately appealing ($M = 3.958$, $SD = 1.804$) and had the largest variation in involvement. We constructed positive reviews for a fictitious GPS brand to avoid potential confounds due to prior brand experience. The reviews were based on actual reviews about existing GPSs.

### *Main study*

The relative importance of different cues for review credibility and usefulness was assessed through conjoint analysis. This research method is used to understand how consumers respond to stimuli varying in characteristics. The characteristics are attributes that have different ‘levels’ (Hair et al., 1992) (for instance the attribute ‘strength of the arguments’ with levels ‘low’ and ‘high’). In the current study, individuals are exposed to eight online reviews varying in attributes and levels (see hereafter) and are invited to rate the credibility and usefulness of each individual review. ‘Part-worth utilities’ for each level of each attribute (the extent to which each level contributes to credibility and usefulness) are the outcomes of the analysis. This, in turn, allows us to calculate the relative importance of each attribute for perceived review credibility and usefulness.

*Design*

The attributes selected for the conjoint analysis are the seven review cues discussed previously. Each attribute has two levels (see Table 2.1 for a detailed overview). The first attribute is the strength of the arguments (strong arguments about relevant functional features such as the processing speed and the price/quality relationship vs. weak arguments about less relevant features, such as the design, the availability of fun accessories, and the color of the product). The second attribute is sidedness (one-sided messages with only arguments in favor of the product vs. two-sided messages with arguments both for and against the product). Because the reviews were all positive, the two-sided review contained more positive than negative arguments. The third attribute is the number of positive review arguments. We used four positive arguments vs. two positive arguments to represent “more” or “fewer” arguments. In the two-sided conditions, we added two and one negative attributes, respectively. We did not want to use more than four positive arguments, because that could cause confounds due to review length. The fourth attribute is writing quality (good, a well-structured review incorrect language vs. poor, an unstructured review containing grammar and spelling errors). The fifth attribute is the number of available reviews (high = 274 other reviews available vs. low = two other reviews available). This number was displayed, but participants could not actually access the other reviews to avoid confounds. The sixth attribute is the rated review usefulness (high = 235 positive and seven negative usefulness ratings vs. low = seven positive and 235 negatives). The last attribute is the average product star rating (present, with four stars out of five vs. absent).

We used SPSS orthoplan to produce a balanced orthogonal design of eight cards (reviews) (Appendix 1). A balanced design means that each level of an attribute occurs an equal number of times over the different stimuli (De Meulenaer et al., 2015). In this study, we used a ‘full profile’ conjoint analysis where respondents score all eight cards, because of its perceived realism (Hair et al., 1992; Sebastianelli & Tamimi, 2018). Before seeing and scoring the reviews, all participants saw the same product description and specifications (such as the price or the memory capacity, see Appendix 1 for more detail) to enhance the tangibility of the product and to provide a standard context to all participants.

*Table 2.1. Items and alpha values for the measures adopted in the main study*

| **Argument strength** | **Sidedness** | **Number of arguments** | **Writing quality** | **Number of reviews** | **Rated usefulness** | **Star rating** |
|---|---|---|---|---|---|---|
| Strong | One-sided | More | Good | High | High | Present |
| *related to functional features such as speed and price/quality relation* | *only positive product reviews* | *4 arguments* | *well-structured review in correct language* | *274 other reviews available* | *235 positive, 7 negative* | *4 stars on a 5-star scale* |
| Weak | Two-sided | Fewer | Poor | Low | Low | Absent |
| *related to design, availability of accessories, and color* | *both positive and negative product reviews* | *2 arguments* | *unstructured review containing grammar and spelling errors* | *2 other reviews available* | *7 positive, 235 negative* | *no star rating displayed* |

*Participants and measures*

The study was conducted utilizing–an online survey in a convenience sample of Belgians recruited via social media. 287 people (47.7% female) completed the questionnaire. The average age of the respondents was 37 years (SD = 17.5) and 79.4% were educated beyond high school. The respondents first saw the product description and were then exposed to each of the eight cards in random order. Table 2.2 contains the details of the measures used in the questionnaire.

Respondents scored the perceived usefulness of each review (Purnawirawan, De Pelsmacker, et al., 2012 $\alpha$minimum = .932) on a three-item scale and review credibility (Soh et al., 2009) on a single item scale. After assessing all reviews, respondents indicated their involvement with the product (De Keyzer et al., 2017 $\alpha$ = .956) by means of a three-item seven-point scale and answered demographic questions (gender, age, and education). All constructs were measured on seven-point semantic differential scales. For further analysis, an average score across items was calculated for the two multi-item scales. Appendix 1 contains the stimuli and questionnaire used in this study.

*Table 2.2. Items and alpha values for the measures adopted in the main study*

| | **Cards** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** |
|---|---|---|---|---|---|---|---|---|---|
| Perceived Usefulness (Purnawirawan, De Pelsmacker, et al., 2012) | *1. I found this review useful*<br>*2. The reviews helped me to shape my attitude toward the GPS*<br>*3. The reviews helped me to make a decision regarding this GPS* | α = .963 | α = .932 | α = .967 | α = .961 | α = .960 | α = .965 | α = .948 | α = .966 |
| Review credibility (Soh et al., 2009) | *1. This review is not credible/ very credible* | N.A. | | | | | | | |
| Involvement with the product (De Keyzer et al., 2017) | *1. A GPS is unimportant – important to me*<br>*2. A GPS is meaningless – meaningful to me*<br>*3. A GPS does not matter to me – does matter to me* | α = .956 | | | | | | | |

## Results

We used IBM SPSS 25 to compute the relative importance of each attribute for each respondent, based on the estimated part-worth utilities for each attribute's level. The part-worth utilities and relative importance of the attributes are calculated for the total sample by averaging the individual scores. In the total sample, the correlation between the actual and predicted preferences is 1 and significant, indicating a good fit (Hair et al., 1992). Table 2.3 presents a summary of the utility estimates and the relative importance of the cues for review usefulness and credibility.

Answering RQ1, the results show that argument strength is the most important cue for both the perceived usefulness (35.6%) and credibility (23.6%) of a review (respectively). The presence or absence of a star rating (6.9% and 8%) and the number of reviews (7.9% and 9.5%) are the two least important cues for both review usefulness and credibility (respectively). For perceived usefulness, the other cues have the following importance: the number of arguments and writing quality (both 13.5%), message sidedness (11.6%), and the rated review usefulness (11.1%). For review credibility, the other cues have the following importance: writing quality (18%), message sidedness (15.1%), the rated review usefulness (13%), and the number of arguments (12.6%).

*Table 2.3. Utility estimates and importance values for perceived usefulness and review credibility*

| | | ***Perceived usefulness*** | | ***Review credibility*** | |
|---|---|---|---|---|---|
| | | *Utility Estimate* | *Importance* | *Utility Estimate* | *Importance* |
| **Arguments strength** | *Strong* | 1.127 | 35.550 | .617 | 23.569 |
| | *Weak* | -1.127 | | -.617 | |
| **Sidedness** | *One-sided* | .115 | 11.610 | .109 | 15.143 |
| | *Two-sided* | -.115 | | -.109 | |
| **Number of arguments** | *More* | .333 | 13.488 | .202 | 12.569 |
| | *Fewer* | -.333 | | -.202 | |
| **Writing quality** | *Good* | .219 | 13.463 | .375 | 17.962 |
| | *Poor* | -.219 | | -.375 | |
| **Number of reviews** | *High* | .056 | 7.892 | -.031 | 9.522 |
| | *Low* | -.056 | | .031 | |
| **Rated usefulness** | *High* | .237 | 11.081 | .305 | 13.029 |
| | *Low* | -.237 | | -.305 | |
| **Star rating** | *Present* | -.001 | 6.916 | .011 | 8.047 |
| | *Absent* | .001 | | -.011 | |
| (Constant) | | 3.851 | | 4.275 | |

Looking at the part-worth utilities, more arguments, stronger arguments, good writing quality and higher rated review usefulness all have positive effects on both review credibility and usefulness. One-sided messages are considered both more useful and more credible than two-sided messages. Fewer reviews (as opposed to more) and the presence of a (positive) star rating causes a review to be perceived as more credible while having more reviews and not presenting a star rating is better for perceived usefulness.

To analyze the influence of product involvement in determining the relative importance of each review cue, we divided the sample in a low- and a high-involvement subsample, using a median split. We excluded 45 participants that scored on the median (5 on a 7-point scale), resulting in 105 responses in the low-involvement group and 137 responses in the high-involvement one. The results are presented in Table 2.4.

*Table 2.4. Utility estimates and importance values for perceived usefulness and review credibility, for high and low involvement groups*

| | | ***Perceived usefulness*** | | | | ***Review Credibility*** | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | **High involvement** | | **Low involvement** | | **High involvement** | | **Low involvement** | |
| | | *Utility Estimate* | *Importance* | *Utility Estimate* | *Importance* | *Utility Estimate* | *Importance* | *Utility Estimate* | *Importance* |
| **Argument strength** | *Strong* | 1.156 | 35.735 | 1.125 | 35.547 | .653 | 24.773 | .597 | 22.311 |
| | *Weak* | -1.156 | | -1.125 | | -.653 | | -.597 | |
| **Sidedness** | *One-sided* | .075 | 10.950 | .141 | 12.239 | .101 | 15.719 | .103 | 14.105 |
| | *Two-sided* | -.075 | | -.141 | | -.101 | | -.103 | |
| **Number of arguments** | *More* | .378 | 14.787 | .292 | 12.200 | .236 | 13.437 | .160 | 11.449 |
| | *Fewer* | -.378 | | -.292 | | -.236 | | -.160 | |
| **Writing quality** | *Good* | .165 | 13.081 | .264 | 13.929 | .318 | 16.829 | .488 | 20.279 |
| | *Poor* | -.165 | | -.264 | | -.318 | | -.488 | |
| **Number of reviews** | *High* | .070 | 8.194 | .050 | 7.899 | -.021 | 9.143 | -.040 | 9.993 |
| | *Low* | -.070 | | -.050 | | .021 | | .040 | |
| **Rated usefulness** | *High* | .187 | 10.057 | .303 | 11.937 | .264 | 12.375 | .365 | 13.792 |
| | *Low* | -.187 | | -.303 | | -.264 | | -.365 | |
| **Star rating** | *Present* | -.025 | 7.195 | .008 | 6.249 | -.008 | 7.723 | .035 | 8.071 |
| | *Absent* | .025 | | -.008 | | .008 | | -.035 | |
| (Constant) | | 3.904 | | 3.755 | | 4.377 | | 4.102 | |

Answering RQ2, the relative importance of argument strength is higher for the high-involvement group than for the low involvement group for perceived credibility (25% > 22%). However, for perceived usefulness, there is no difference between the two groups (37.5% vs. 37.5%). Regarding the sidedness of the message, the results show that it is more important for the high-involvement group for perceived credibility (high-involvement = 15.7% > low-involvement = 14.1%), but not for perceived usefulness (high involvement = 10.9% < low involvement = 12.2%). Writing quality is more important for the low-involvement group than for the high-involvement group for review credibility (16.8% < 20.2%), but for perceived usefulness the difference is, again, negligible (13.1% and 13.9%). The importance of the number of reviews is not different for the high and low involvement individuals, since the difference is less than 1%, which is negligible, for both dependent variables. The rated review usefulness is more important for less involved individuals (11.9%, 13.8%) than for higher involved people (10.1%, 12.4%) when assessing both perceived usefulness and credibility (respectively). The difference between high and low involvement for the presence or absence of star rating is negligible for perceived usefulness (7.2% and 6.2%) and credibility (7.7% and 8.1%). Looking at how the number of arguments influences the perceived credibility and usefulness of a review, the results show that this cue is more important for the high-involvement group than for the low-involvement group when assessing perceived usefulness (14.7% > 12.2%) and credibility (13.4% > 11.4%).

# Discussion

We explored the relative importance of seven review cues in readers' assessment of the perceived usefulness and credibility of online reviews. Argument strength (a central cue) is the most important cue, and the number of reviews and the presence or absence of a summary review star rating are the least important cues for both review usefulness and credibility. Previous research also shows that argument strength is an important predictor of usefulness and credibility (e.g., Thomas et al., 2019; Wang & Li, 2019). Strong arguments contain diagnostic information that is useful for decision making (Filieri, Hofacker, et al., 2018). The second most important cue for perceived usefulness is the number of arguments. This finding supports the suggestion of  Schindler and Bickart (2012) that reviews with many arguments contain more details, which can help consumers with their purchase decision (Schindler & Bickart, 2012).

The second most important cue for perceived credibility (and third for perceived helpfulness) is the writing quality. This is consistent with previous findings that show that poor grammar makes readers question the competence of the author and dismiss the review's credibility as a result (Clare et al., 2018) and that readability (i.e., writing quality) is one of the most important variables determining review helpfulness (Singh et al., 2017). The low importance attributed to peripheral cues such as the number of reviews, rated usefulness, and summary review star rating is consistent with previous studies in that the effects of peripheral review cues on review impact are limited when central cues are present (Cheung et al., 2012; Chevalier & Mayzlin, 2006; De Pelsmacker, Dens, et al., 2018), especially for a single review. For instance, De

Pelsmacker, Dens, et al. (2018) found that summary review star rating does not affect review impression when people have a review text to rely on.

Previous research was inconclusive concerning the effect of message sidedness on review usefulness and credibility. In the current study, we found that consumers perceive one-sided review messages as both more helpful and more credible than two-sided messages. The unambiguous advice provided in one-sided reviews seems more useful to readers, which is consistent with the finding of Cao et al. (2011) that reviews with extreme opinions (be it positive or negative) receive more helpfulness votes than the ones with mixed opinions (two-sided). Our study also shows that, contrary to what was found by Cheung et al. (2012), one-sided messages affect credibility more positively than two-sided messages. This is in line with what was found by Pentina et al. (2018), as one-sided positive reviews are perceived as more credible than two-sided reviews. Since reviews are written by consumers who have no stake in the brand, the positive effect of two-sided messages in advertising on credibility does not occur (Schlosser, 2011). This is in line with Metzger et al. (2020), that state that one-sided messages can sometimes be perceived as more credible due to the confirmation bias in information processing.

The effect of the number of reviews differs between credibility and usefulness. Readers perceive the availability of more reviews as more useful. This finding is consistent with the idea that, by seeing that a high number of reviews are available, the consumers' confidence in them increases, as they see that many others are interested in that product (Purnawirawan, De Pelsmacker, et al., 2012). On the other hand, having fewer reviews causes a review to be perceived as more credible. In line with Thomas et al. (2019), our results suggest that a high number of reviews may be perceived by readers as unrealistic or fabricated, damaging the credibility of the reviews.

The central cue 'argument strength' is more important to determine perceived review credibility for highly involved than for lowly involved individuals. This finding is in line with previous research (Cheung et al., 2012; Filieri, Hofacker, et al., 2018). The peripheral cues 'writing quality' and 'rated review usefulness' are more important for low-involvement individuals than for highly involved ones. These cues serve as mental shortcuts (Richard E. Petty & John T. Cacioppo, 1986) for individuals that are not highly motivated to process the information in the reviews.

When evaluating review credibility, message sidedness is more important for highly involved individuals than for lowly involved ones. In this case, high-involvement individuals may be focusing on sidedness as a way to assess the completeness of information in the review (Cheung et al., 2012). On the other hand, and contrary to our expectations, message sidedness (a central cue) is more important for perceived usefulness of a review for lowly involved individuals than for highly involved ones. The reason for this may be that a clear-cut message is more helpful for low involvement individuals as they would not spend much cognitive effort to process dissenting opinions (Richard E. Petty & John T. Cacioppo, 1986), which is also consistent with the positive utility attributed to one-sided messages.

The number of arguments is also more important for higher involved than for less involved readers. This suggests that, in the context of the present study, the number of arguments is rather used as a central cue to assess usefulness. In agreement with the findings of Willemsen et al. (2011a) and Schindler and Bickart (2012), reviews with more arguments contain more information, which helps review readers with their decision about the product or service. Besides, the arguments used in this study are presented in short sentences, making it easier for the reader to process the information.

Finally, there is no difference between the relative importance of the number of reviews and summary review star rating in high and low involvement individuals when evaluating the perceived usefulness and credibility of the review. The differences in the relative importance of the number of arguments and rated review usefulness are also negligible between high and low involvement when assessing review credibility.

## Theoretical and Managerial implications

The results shed light on the relative importance of the most frequently studied online reviews cues, in each other’s presence. Therefore, the contribution of this paper to theory is threefold. First, we look into the relative importance of cues that are well studied in the context of online reviews offering a comprehensive analysis that not only allows to compare the relative importance of several cues, but also simulates a realistic context where consumers make (implicit or explicit) trade-offs between cues in a review. We confirm the importance of the review text over other cues, such as star rating (e.g., De Pelsmacker, Dens, et al., 2018; Thomas et al., 2019), and the limited importance of peripheral cues in the presence of central cues, at least when readers are only exposed to a single review (Cheung et al., 2012; Chevalier & Mayzlin, 2006; De Pelsmacker, Dens, et al., 2018). Second, we study the relative importance of cues whose role was not clear in previous literature and found that consumers perceive one-sided reviews as more useful and credible than two-sided reviews, which sheds new light onto the role of sidedness in online reviews. The fact that review volume contributes positively to usefulness, but negatively to credibility is also an important contribution. Consumers could perceive a higher number of online reviews as less credible because they suspect that they may be getting fake reviews (Y. Wu et al., 2020). Third, this study provides a test of the principles of the ELM to explain the effects of a set of characteristics of online reviews on persuasion. How review readers elaborate on certain cues will also depends on what other information is available in the review and competing for their attention. For instance, contrary to previous research that could not prove that peripheral cues (such as star rating and rated usefulness) were relatively more important for low involvement individuals (Kolomiiets et al., 2016), we find that the peripheral cues (for instance 'writing quality' and 'rated review usefulness') are more important for low involvement individuals than for highly involved ones. This confirms what could be expected based on the ELM (Richard E. Petty & John T. Cacioppo, 1986), namely that individuals who are not highly motivated take mental shortcuts to process the information. Another contribution to the use of the ELM to study online reviews is the role of the number of arguments in a review. Previous research showed contradictory findings for this review characteristic but, according to our results, the number of arguments appears to be processed centrally, being more important for highly involved individuals.

The current study provides insights for administrators of online review sites and marketers. As consumers often have many, sometimes contradicting, online reviews at their disposal (Gottschalk & Mafael, 2017), they need to simplify the processing of these reviews as they cannot consider all the available information. Perceived usefulness and credibility are important 'gatekeepers' to the further decision-making process. In general, considering the importance attributed to the text-related elements of a review, managers should request reviewers to write something, rather than merely provide a star rating, for example. Importantly, the reviews should contain strong arguments and should be impeccably written. Managers could incentivize strong arguments by rewarding reviews with a higher-rated usefulness or by suggesting important attributes or aspects that the review could mention. Writing quality could be ensured by providing automatic grammar and spelling controls. Reviewers should also be encouraged to write 'rich' reviews, with a sufficiently large number of arguments. Review platforms could for example provide people with a template or a set of criteria for reviewers to comment on or rate.

The rated review usefulness is relatively important and should, therefore, be highlighted, for example, by sorting reviews based on their helpfulness by default, or allowing users to do so. One-sided arguments create more favorable perceptions of credibility and usefulness than two-sided ones. By explicitly asking reviewers to write both positive and negative arguments, which some platforms (such as TripAdvisor) do, practitioners may be impairing the perceived credibility and usefulness of the review. A system in which reviewers are instructed to give their opinion, without specifically asking for positive and negative aspects would be preferable in this case. Importantly, each cue positively contributes to helpfulness and credibility, which means that reviewers should combine them to increase the helpfulness and credibility of their review. It is also possible that the cues would further reinforce each other, or could be reinforced by other cues not included here, a possibility which we could not explore within the current set-up. Ma et al. (2018), for example, showed that joining review texts and user-provided photos shaped the maximum performance, compared to text or photos alone. A few peripheral cues, such as the number of reviews and the presence of a star rating, are relatively unimportant and should thus not necessarily be included in case of single short reviews, as in the current study. Both these elements could become more useful, though, when people are exposed to a larger set of reviews and/or longer reviews.

Considering the differences between high and low involvement individuals, it may be interesting for practitioners to consider different website layouts depending on consumers' involvement. Personalization of web layout and content is increasingly feasible through artificial intelligence. Involvement could be deduced from, for example, previous searches for products in the same category, or likes or interests on social media. For more highly involved individuals, the central arguments of a review should be easily accessible and could be highlighted using bold font. For more lowly involved individuals, it would be more useful to see an overall assessment of the review, such as the rated review usefulness.

## Limitations and future research

Our study has some limitations that should be taken into account in future research. In conjoint analysis, the relative importance of attributes is determined by the selection of attributes and their levels. For example, for "star rating", we opted for the presence of a 4-star (out of 5) rating, versus no rating. A more extreme rating (5 out of 5 stars), or a comparison with a 1-star rating instead of "no rating" might lead to different results. Further research should, therefore, examine other levels of the cues to test the stability of our findings. For instance, in De Pelsmacker, Dens, et al. (2018), peripheral cues became irrelevant in the presence of a central cue. Further research should test the relative importance of cues (attributes) in the presence or absence of other cues.

Our sample is highly educated (71.3% were educated beyond high school), so the demographic characteristics of our participants may influence our results. It is possible that due to their high level of literacy, the individuals of our sample were more attentive to the writing quality of the reviews than a sample with other demographics. As such, future research should replicate our study in other demographic segments, such as lower educated individuals.

The present study considers how review cues differ in their relative importance depending on people's level of involvement with the product. Future studies should complement our findings by studying other product categories and other product types, for instance, search, credence and experience products, or utilitarian versus hedonic products, and compare the results across these types of products and product categories. For instance, central cues may be more important when reading reviews for search than for experience products since previous research found that consumers determine the credibility of a review for search products by the level of detail in the review (Jiménez & Mendoza, 2013). Besides, both sender and receiver characteristics, and the relationship between them may play a role in the perceived relative importance of review cues. In this study, we included involvement with the product as a receiver characteristic. Future research could also incorporate sender characteristics, such as whether the sender is a verified buyer or not. On social network sites, relational characteristics, such as homophily and tie strength, play an important role. For instance, it might be the case that the relative importance of the review cues, such as the role of writing quality on review credibility, is different when the reader has some connection with the reviewer. Future research should study the effects of these relational factors.

The stimuli presented to the participants were static, which means that respondents read eight static reviews. Reviews were therefore not displayed in their natural environment in which numerous reviews are accessible simultaneously, competing for the readers' attention. Other approaches may consider looking at the relative importance of these cues when multiple reviews are assessed simultaneously as it is possible that other elements, such as summary review star rating, will gain importance, as it provides a summary of the information available (Richard E. Petty & John T. Cacioppo, 1986).

Finally, considering that the orthogonal design adopted in this study does not account for interaction effects, further studies may look at how the different review attributes interact with each other. For example, while Ma et al. (2018) found that user-provided photos did not have the same influence as review texts, joining both elements shaped the maximum performance. In the current study as well, it is possible that some cues could be less important themselves, but could serve to reinforce the effect of other cues. Brand-related aspects should also be considered, as Wen et al. (2020) document a three-way interaction between review valence, brand familiarity, and price.

# 3. Valence and Attribute Repetition in Negative Sets of Online Reviews: (When) can positive reviews overcome negative ones?[7,8]

[7] This manuscript is accepted at the Journal of Electronic Commerce Research (authors: Lopes, A. I., Dens, N. and De Pelsmacker, P.); scheduled to appear in the February 2022 issue.
[8] An earlier version of this chapter was presented at 2019 International Conference on Research in Advertising, Krems an der Donau, Austria.

# Abstract

Review set valence (the degree of negativity or positivity of a set of online reviews) strongly determines review readers' responses. Previous research has mainly considered the mere number of positive and negative reviews to determine a review set's valence. This paper aims to study how increasing the number of important positive reviews influences readers' hotel staying intention, exploring the 'tipping point' at which important positive reviews compensate for the negative effect of a larger number of less important negative reviews. We further explore whether reader responses are more positive when all positive reviews address the same product attribute or different attributes. We present a 4 (review set valence) x 2 (attribute repetition vs. different attributes for the positive reviews) online experiment (N=408). The results show that a more positive review set leads to a higher staying intention only when the positive reviews discuss different attributes (and do not repeat the same attribute). The 'tipping point' at which positive reviews compensate negative ones is four positive reviews about different attributes in a set of 12. This study nuances the bandwagon effect, negativity bias and truth effect by showing that negative review sets can be positively evaluated.

# Introduction

eWOM (electronic Word-Of-Mouth) is any positive or negative statement made by customers about a product or company, made available to a multitude of people and institutions via the Internet (Ismagilova et al., 2017). The current study focuses on online reviews, product evaluations generated by users or experts based on their personal experience (Purnawirawan, De Pelsmacker, et al., 2012), as a specific type of eWOM. Online reviews strongly influence readers in their product or service-related purchase decisions (Baek et al., 2015; Chong et al., 2018). Consumers usually do not read just a single review, but instead interpret a compilation of different reviews that can be positive, negative, or both.

The valence of a set of reviews can be either positive, neutral, or negative, depending on the ratio of positive and negative individual reviews (Purnawirawan, De Pelsmacker, et al., 2012). The effect of review set valence on review readers can be explained by the bandwagon effect (Lee et al., 2018; Sundar et al., 2008), which states that people tend to make choices based on a perceived trend, without making judgments about the trend. Previous literature (e.g., Brunner et al., 2019) also shows that negative reviews are often more influential than positive reviews (negativity bias). Therefore, it would be logical to expect that products or services with mainly negative reviews (and thus a minority of positive reviews) would always be negatively evaluated. However, previous research showed nuances to the bandwagon effect and negativity bias (e.g., Hair & Bond, 2018; Wu, 2013), indicating that negatively valenced sets of reviews might not always lead to negative evaluations.

The importance of the arguments in a review could be one of the main factor driving these nuances (Filieri, Hofacker, et al., 2018). Review readers perceive reviews about important attributes as more diagnostic. Having positive reviews discuss important attributes (while the negative reviews pertain to less important attributes) may thus offset the negativity bias. As far as we are aware of, no previous research has yet addressed how predominantly negative sets of reviews are processed when the positive reviews in the set are about important attributes and the negative reviews are about less important attributes. In this study, we aim to investigate nuances to the bandwagon effect and the negativity bias. By manipulating the importance of positive and negative reviews, we explore the 'tipping point' where negative review sets can actually lead to positive booking intentions.

The effects of repeated exposure to advertising containing the same or different arguments are frequently studied in the advertising field (Chang, 2009). However, the effect of argument repetition in the context of online reviews is not clear. There are arguments in favor of repeating arguments to increase their believability (e.g., Dechêne et al., 2010) and others in favor of diversifying the arguments to increase information utility (e.g., Zhang et al., 2014). The effects of exposing online review readers to review sets discussing the same or different attributes on their behavioral intentions, have not been studied yet. Considering how common this phenomenon is in practice and how previous research fails to address it, we explore the effects of argument repetition on how online review information is processed.

Based on the previously mentioned research gaps, this study's first objective is to investigate how varying ratios of a majority of negative reviews about less important attributes and a minority of positive reviews about more important attributes influence review readers' intention to stay at a hotel. Consequently, our first research question is:

*How do varying ratios of a majority of negative reviews influence consumers' staying intention at a hotel when the positive reviews all address important attributes and the negative reviews address less important attributes? What is the 'tipping point' at which having positive reviews in a predominantly negative review set leads to a positive hotel booking intention?*

Moreover, we fill in the gap in previous literature by also looking at how repeating the same product attribute in the positive reviews vs. including positive reviews about different attributes moderates the effect of review set valence on consumers' responses. Therefore, the second research question guiding this study is:

*How do (multiple) positive reviews about the same attribute versus positive reviews about different attributes moderate the effect of the ratio of positive reviews on readers' intention to stay at a hotel?*

To answer these research questions, we first present a literature review on the effects of review set valence and attribute repetition. We then develop a 4 (ratio of positive reviews about important attributes to negative reviews about less important attributes) x 2 (attribute repetition vs. different attributes for the positive reviews) between-subjects full factorial design experiment.

## Literature Review and Hypotheses Development

Previous studies indicate that the valence of a set of reviews is a crucial determinant in the way consumers respond to eWOM (Floyd et al., 2014; Luo et al., 2021; Mafael et al., 2016; Zablocki et al., 2018). Review readers tend to follow the majority's opinion: they evaluate products positively after exposure to positively valenced review sets and negatively after exposure to negatively valenced sets (Brunner et al., 2019; Doh & Hwang, 2009; Xun & Guo, 2017). The effect of review set valence can be explained by the bandwagon effect (Lee et al., 2018; Sundar et al., 2008), a psychological phenomenon in which people tend to join what they perceive to be existing or expected majorities or dominant positions in society. In other words, when review readers perceived the majority of reviewers to be negative, they will "join them" and form a negative opinion of the reviewed object.

Moreover, positive and negative review information seems to carry different weights in judgment (Purnawirawan et al., 2015). In most situations, negative events are more salient and more influential than positive events (Rozin & Royzman, 2001). Negative online reviews are typically perceived as more usefulness than positive reviews (Jeong & Koo, 2015). Previous research on the negativity bias shows that negative reviews tend to be more influential on people's judgment of a product or service than positive ones (Lee et al., 2009). Notably, the relative weight of negative over positive reviews as predicted by the negativity bias will also depend on other review characteristics, such as the importance or quality of the arguments. Previous research has shown that argument importance determines how strongly online reviews influence review readers (Filieri, Hofacker, et al., 2018; Park et al., 2007; Thomas et al., 2019; Willemsen et al., 2011b). For instance, Filieri, Hofacker, et al. (2018) found that relevant reviews, i.e. reviews about important attributes, are perceived as diagnostic information, useful for consumers' decision-making. In the same vein, the argument strength seems crucial in influencing consumers' perceived review usefulness and credibility (Thomas et al., 2019).

We expect that having positive reviews about more important attributes in a mainly negative set of reviews could attenuate the bandwagon effect and the negativity bias. While the bandwagon effect and negativity bias are relatively well-established, a few studies already propose nuances. In a mainly positive review set, Hair and Bond (2018) found that review readers dismiss negative reviews as inconsequential when they discuss product attributes that are of low importance. Another study (Shoham et al., 2017) shows that including a negative irrelevant review in a positive review set does not harm product evaluations. In contrast, it can even improve them because consumers feel more confident that the information they have about the product is more complete (Shoham et al., 2017). Pentina et al. (2018) disconfirm the negativity bias by showing that positive reviews are perceived as more trustworthy, credible, and helpful than negative reviews. From an emotional value perspective, positive reviews influence consumers' decision making by enhancing the utility derived from positive feelings evoked by the review (Xia & Bechwati, 2008). Therefore, we expect that increasing the proportion of positive reviews in a negative set would affect review reader's intention to stay at the hotel by attenuating the negativity bias. We propose the following hypothesis to unveil at which ratio of positive reviews in a negative set there is a 'tipping point':

*H1: In a predominantly negative review set, having positive reviews about important attributes can lead to a positive intention (above the scale midpoint) to stay at a hotel.*

We also explore how including the same (repeated) or different attributes in a review set influences readers' intention to stay at a hotel. According to the truth effect, repeating arguments (i.e., reviews pertaining to the same attributes as the other reviews in the set) increases participants' subjective judgments of a statement's truth (Dechêne et al., 2010; Roggeveen & Johar, 2002). McCullough and Ostrom (1974) conducted an experiment with five similar advertisements using the same, but differently phrased, arguments and found a positive relationship between the number of repetitions and the attitude toward the product. Cacioppo and Petty (1989) state that moderate levels of repetition can increase persuasion when the arguments are strong (i.e., about important attributes). In the present study, the positive reviews are about important attributes: their strength is related to the importance that readers attribute to it (Cheung et al., 2009). These studies would predict a positive effect of attribute repetition on hotel staying intention. In contrast, a previous study in political communication showed that repeatedly presenting the same posters resulted in a negative attitude toward the presented political issue, mediated by a decrease in credibility judgments (Ernst et al., 2017). These findings suggest that there are limits to the truth effect.

Other studies point in a different direction regarding the effects of argument repetition. The repetition-variation hypothesis in advertising states that providing different arguments increases persuasion by increasing issue-relevant thoughts or by serving as a simple acceptance cue (Calder et al., 1974; Petty & Cacioppo, 1984). The persuasion literature also shows that messages with more arguments and reasons are more persuasive as they provide confidence in decision-making (Srivastava & Kalro, 2019). For example, having multiple speakers presenting multiple arguments enhances persuasion over having either multiple speakers or multiple arguments because of greater information utility (Harkins & Petty, 1981). In the context of online reviews, one could expect that the same effects take place. For a single review, increasing the number of arguments increases a review's perceived helpfulness (Baek et al., 2012). Similarly, Willemsen et al. (2011b) found that reviews are evaluated as more useful when they offer more arguments to back up their valenced statements. The more distinct the arguments presented to the reader, the more they affect consumers' purchase intention because the information about the reviewed product is more comprehensive (Lopes et al., 2020; Zhang et al., 2014).

Considering the limitations to the truth effect pointed out by Ernst et al. (2017) and the previous research on online reviews that finds a positive effect of presenting diverse arguments on review readers' intentions, we expect the following:

*H2: Having different attributes in positive reviews leads to a higher intention to stay at a hotel than repeating the same attribute.*

Wu (2013) indicates that, in the context of eWOM, the negativity bias can be attenuated or even reversed because the quality of a review plays a determinant role when consumers assess the usefulness of eWOM. As argued in the development of H2, increasing the number of arguments in an online review makes the review more complete and clear (Lopes et al., 2020; Zhang et al., 2014), contributing to its quality. This argumentation can be transposed to a context with multiple reviews: increasing the number of arguments across a set of reviews will impact the perceived quality of the set and its effect on decision making. This effect on decision making originates on the increased amount of available information that consumers can use to make their judgment, as expected based on the accessibility-diagnosticity theory (Herr et al., 1991). Information diagnosticity refers to the ability of the information in online reviews to enable readers to learn and evaluate the quality and performance of services (information diagnosticity) before purchasing them (Filieri, Hofacker, et al., 2018). The greater the information diagnosticity of reviews, the higher will be the influence on purchase intentions (Filieri, 2015; Herr et al., 1991). Information relevancy is one of the most important predictors of perceived information diagnosticity (Filieri, Hofacker, et al., 2018). The relevance of adding positive reviews that simply repeat information already provided by other reviewers is smaller than when the reviews add new arguments. Adding positive reviews about diverse attributes can help reduce uncertainty about more attributes. Therefore, we expect that the benefits of increasing the number of positive reviews (given the bandwagon effect) are greater when the reviews discuss different attributes, compared to a single attribute. This reasoning is also in line with advertising studies showing that wear-out effects occur with greater repetition (Schmidt & Eisend, 2015). Wear-out occurs because of redundancy or boredom (Berlyne, 1970), which result in negative thoughts (Cacioppo & Petty, 1979) that outweigh the positive ones.

In line with these arguments, an increasing number of positive reviews about different important attributes would benefit the intention to stay at the reviewed hotel more than when the positive reviews are all about the same attribute. Therefore, we propose:

*H3: The positive effect of adding positive reviews to a negative review set on the intention to stay at a hotel is reinforced by having different attributes rather than repeating the same attribute.*

## Empirical study

We developed a 4 (ratio of positive reviews about important attributes to negative reviews about less important attributes: 5 positive/7 negative; 4 positive/8 negative; 3 positive/9 negative; 2 positive/10 negative) x 2 (attribute repetition vs. different attributes for the positive reviews) between-subjects full factorial design experiment, creating eight experimental conditions (see Table 3.1). The main study encompasses an experiment in which each participant reads 12 reviews, as in the studies developed by Hair and Bond (2018). Another reason to choose using 12 reviews is that Purnawirawan (2013) found that people read at least five to ten reviews per search session, so by presenting 12 reviews, we aim to provide the reader with sufficient information for decision making.

*Table 3.1. Overview of the 4x2 between-subjects design*

| **Conditions** | **Review set ratio** | **Attribute repetition** |
|---|---|---|
| 1 | 2 Positive reviews/ | Positive reviews: all different attributes |
| 2 | 10 Negative reviews | Positive reviews: repetition of the same attribute |
| 3 | 3 Positive reviews/ | Positive reviews: all different attributes |
| 4 | 9 Negative reviews | Positive reviews: repetition of the same attribute |
| 5 | 4 Positive reviews/ | Positive reviews: all different attributes |
| 6 | 8 Negative reviews | Positive reviews: repetition of the same attribute |
| 7 | 5 Positive reviews/ | Positive reviews: all different attributes |
| 8 | 7 Negative reviews | Positive reviews: repetition of the same attribute |

*All the positive reviews relate to important attributes*
*All the negative reviews refer to different unimportant attributes*

### *Pre-tests*

In our main study, we test sets of 12 reviews, the overall valence of which is negative, with a varying number of positive reviews (either 2, 3, 4, or 5). The positive reviews include important attributes, while the negative reviews discuss less important attributes. An all-inclusive resort was chosen as the experiment setting since most participants could easily relate to this context, and reviews about holiday resorts are extensively available and consulted by travelers (Yang et al., 2018). Two pre-tests were carried out. First, to select relatively important and relatively less important resort attributes in the decision to stay at a hotel. Second, to test the perceived valence of the reviews for use in the main study. In both pre-tests, we provided a scenario in which respondents planned to spend their holidays in an all-inclusive resort and were asked to evaluate hotel reviews. In the first pre-test, 30 respondents (46.7% female; Mean of age = 33, Standard Deviation = 7.5) recruited through Prolific (online recruitment platform) rated the perceived importance of 50 attributes (e.g., “The size of the hotel lobby”) on a 5-point scale. The list of 50 attributes was based on the attributes used in the study of Purnawirawan, Dens, et al. (2012), complemented with other attributes that online reviews for actual all-inclusive hotels on TripAdvisor frequently mention. We selected the 15 least important attributes (with average scores between 2.77 and 3.47; e.g., “The variety of gym appliances”) and the 10 most important attributes (average scores between 4.33 and 4.67; e.g., “The cleanliness of the room”). All the less important attributes were significantly less important in the decision to stay at a hotel than the more important attributes (t-tests, all $p < 0.002$).

In the second pre-test, we formulated 43 reviews using these attributes. We recruited 28 respondents through Prolific (53.6% female; Mean of age = 34, Standard Deviation = 10.8) to rate the reviews' perceived valence on 5-point scales. Because the wording of a review could influence the perceived importance of an attribute, we developed various reviews per attribute and tested the importance again as well. We ultimately selected 10 negative reviews about less important attributes and 5 positive reviews on important attributes (see Table 3.2). The selected positive reviews were significantly more positive and significantly more important than the negative reviews (t-tests, all $p < 0.001$). Between 71.4% and 92.9% of participants rated the importance of the negative (less important) reviews as 3 or less while 96.4% to 100% scored the positive (more important) reviews 4 or 5 (on a 5-point scale), showing that respondents consistently evaluated the less important and the more important reviews as intended.

All the negative reviews are about different less important attributes, and all the positive reviews discuss important attributes, as established in the pre-tests. In the conditions in which the same positive attribute was repeated, the positive reviews all discussed the freshness of the food, in different words, to enhance the realism of the review set. The stimuli contained only the review text presented in Table 3.2 to avoid possible confounds.

The reviews were distributed across conditions so that the average importance and valence of the positive reviews and the average importance and valence of the negative reviews, respectively, were equal across conditions. The purpose of this distribution is to guarantee that our results for the different conditions are not due to a specific review that might, for instance, be about a more important attribute than the reviews included in the other sets (Table 3.3).

We also set out to test differences between conditions in which the positive reviews all discuss the same attribute versus different attributes. When the positive reviews discuss the same attribute, we selected the attribute *'The food was always freshly made, amazing'* (Mean Importance = 4.68; Mean Valence = 4.82). The importance and valence of this attribute were closest to (and not significantly different from) the average importance and valence of the selected positive reviews about different attributes (see Table 3.2). We then developed 4 more slight variations of this review to enhance the realism of the set.

*Table 3.2. Average values of importance and valence for all conditions*

| | **10N/2P** | **9N/3P** | **8N/4P** | **7N/5P** |
|---|---|---|---|---|
| **Negative reviews:** | | | | |
| It was a shame that the minibar didn't offer much choice. | X | X | X | X |
| I was disappointed that the hotel does not offer any motorized watersports. | X | X | X | X |
| The sunbeds were very uncomfortable. I couldn't lie on them for a long time. | X | X | X | X |
| The hotel lobby was small, I felt crammed when we were checking in. | X | X | X | X |
| The gym offers little variety in equipment. More choices would have been better. | X | X | X | X |
| The best spots by the pool were always taken. | X | X | X | X |
| There was no one at the hotel of our age. | X | X | X | X |
| It was too bad that the hotel didn't have a wellness area. | X | X | X | |
| I wanted to rent a boat and had to find a rental company in the town… It would have been easier to book it through the hotel. | X | X | | |
| The hotel garden was very small, I really felt like I could use some more green. | X | | | |
| **Positive reviews (different attributes):** | | | | |
| The room was incredibly clean… It was really amazing to have such a tidy place to stay. | X | X | X | X |
| It was great that we could access the Wi-Fi in the room, really nice. | X | X | X | X |
| The bed was very comfortable. It was great to rest. | | X | X | X |
| The food at the hotel restaurant was delicious, I loved it. | | | X | X |
| The food was always freshly made, amazing. | | | | X |
| **Mean importance (across all 12 reviews)** | **2.79** | **2.97** | **3.16** | **3.34** |
| **Mean valence (across all 12 reviews)** | **2.65** | **2.85** | **3.05** | **3.27** |
| **Positive reviews (same attributes):** | | | | |
| The food was always freshly made, amazing. | X | X | X | X |
| It was great that the buffet always had fresh food available. | X | X | X | X |
| I really enjoyed the food at the hotel, always fresh! | | X | X | X |
| I loved the freshness of the food! | | | X | X |
| The best fresh food. Definitely a plus. | | | | X |

**The cells marked with an X mean that the corresponding review is included in the set.*

*Table 3.3. Average importance and valence of reviews per condition (Pre-test)*

| | **Importance negative reviews** | | **Importance positive reviews** | | **Valence negative reviews** | | **Valence positive reviews** | |
|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Condition 1 | 2.40 | .18 | 4.77 | .08 | 2.24 | .19 | 4.71 | .25 |
| Condition 2 | 2.40 | .18 | 4.68 | .55 | 2.24 | .19 | 4.82 | .39 |
| Condition 3 | 2.39 | .19 | 4.71 | .11 | 2.24 | .20 | 4.68 | .19 |
| Condition 4 | 2.39 | .19 | 4.68 | .55 | 2.24 | .20 | 4.82 | .39 |
| Condition 5 | 2.40 | .21 | 4.68 | .11 | 2.23 | .11 | 4.70 | .16 |
| Condition 6 | 2.40 | .21 | 4.68 | .55 | 2.23 | .11 | 4.82 | .39 |
| Condition 7 | 2.39 | .21 | 4.68 | .10 | 2.23 | .12 | 4.72 | .15 |
| Condition 8 | 2.39 | .21 | 4.68 | .55 | 2.23 | .12 | 4.82 | .39 |

### *Procedure and sample of the main study*

We recruited 463 participants from the United States of more than 18 years old through Prolific (the same platform as the pre-tests). As in the pre-tests, the questionnaire first presented a scenario with a description of an all-inclusive resort. Respondents then indicated their experience with all-inclusive resorts. Fifty-five respondents were excluded, fifty due to a lack of previous knowledge of or experience with all-inclusive resorts, and five because of failing two or three of the three attention checks in the questionnaire. The final sample (N = 408) consisted of 50.2% female, ranging from 18 to 66 years old (Mean of age = 35, Standard Deviation = 10.7). 46.6% had a Bachelor's degree, 34.6% had completed high school, 18.6% had a Master's degree or higher, and 0.2% attended primary school. The mean age was not significantly different ($p$ = .605) between the three samples (35 for the main experiment, 33 in the first pre-test, and 34 in the second pre-test). The proportion of men and women was also comparable ($p$ = .867: 50.2% women for the main experiment, 46.7% women in the first pre-test and 53.6% women in the second pre-test), as was the level of education ($p$ = .360).

Participants were randomly assigned to one of the eight conditions. The order of the reviews in each set was randomized to avoid confounding effects of the display order (Kolomiiets et al., 2016; Nan et al., 2017; Purnawirawan, Dens, et al., 2012). Respondents rated their intention to stay at the presented resort after reading the reviews in the condition to which they were assigned. Intention to stay was measured using the seven-point scale developed by Netemeyer et al. (2005) ($\alpha$ = .962), anchored by ‘strongly disagree’ and ‘strongly agree’ (3 items, e.g. ‘*It* *is very likely that I will stay at this resort'*). Appendix 2 contains the stimuli and questionnaire used in this study.

## Results

We conducted a series of one-sample t-tests to test the first hypothesis, which states that even in a predominantly negative review set, having enough positive reviews about important attributes leads to a positive (above the scale midpoint) intention to stay at a hotel. We, therefore, tested if the stay intention at each of the four ratios differed significantly from the scale midpoint. The results show that review sets with a valence ratio of 10 Negative/ 2 Positive (Mean Difference = -.657, $p < .001$) and a valence ratio of 9 Negative/ 3 Positive (Mean Difference = -.603, $p < .001$) score significantly below the scale midpoint. The stay intention for review sets with a valence ratio of 8 Negative/ 4 Positive (Mean Difference = -.157, $p = .115$), and a valence ratio of 7 Negative/ 5 Positive (Mean Difference = .062, $p = .564$) does not differ significantly from the scale midpoint. These results do not support our first hypothesis: although the intention to stay at a hotel is above the scale midpoint for a review set with 7 Negative and 5 Positive reviews, the intention scores are not significantly different from the scale midpoint.

To test how the ratio of positive and negative reviews and attribute repetition affect the intention to stay at the hotel (H2 and H3), we conducted an ANOVA with a Scheffé post hoc test to compare groups. The analysis revealed a significant positive main effect of the ratio of positive to negative reviews [$F(3, 400) = 12.96$, $p < .001$, partial $\eta^2 = .09$], showing that having relatively more positive reviews increases the intention to stay at the hotel. The results for attribute repetition show a significant negative main effect [$F(1, 400) = 10.78$, $p = .001$, partial $\eta^2 = .03$] meaning that, compared to a review set repeating the same positive attribute, having different positive attributes increases the intention to stay at the hotel. This confirms H2.

The overall interaction effect between ratio and repetition (Figure 3.1) is not significant [$F(3, 400) = 1.61$, $p = .186$, partial $\eta^2 = .01$]. Importantly, however, the post hoc test results show that when the same positive attribute is repeated (left-hand side of Figure 3.1), there is no significant difference in staying intention between the conditions with different ratios. In contrast, when the positive reviews discuss different positive attributes (right-hand side of Figure 3.1), there is a significant difference in staying intention between the set with 4 positive reviews and the one with 3 positive reviews (Mean Difference = .742, Standard Deviation = .194, $p$ = .044). The differences between 5 and 3 positive reviews (Mean Difference = .829, Standard Deviation = .191, $p$ = .010) and 5 and 2 positive reviews (Mean Difference = .801, Standard Deviation = .192, $p$ = .017) are also significant. A set with 4 positive reviews is not significantly different from a set with 5 or 2 positive ones. There is also no significant difference between having 2 or 3 positive reviews in the set of 12.

The results for the conditions with positive reviews about different attributes further show that the intention to stay at the hotel (Mean = 3.163, Standard Deviation = .869) exceeds the scale midpoint with 4 or more positive reviews in the set. The conditions with only 2 or 3 positive reviews score significantly below the scale midpoint ($p < 0.01$). The tipping point where the positive reviews overcome the negativity of the set is thus at 4 positive reviews (out of 12), as long as the reviews discuss different attributes. When the same positive attribute is repeated, the intention to stay at the hotel never exceeds the scale midpoint. We thus find support for H3, which predicted that the positive effect of adding positive reviews to a negative set is reinforced by having different attributes rather than repeating the same attribute.

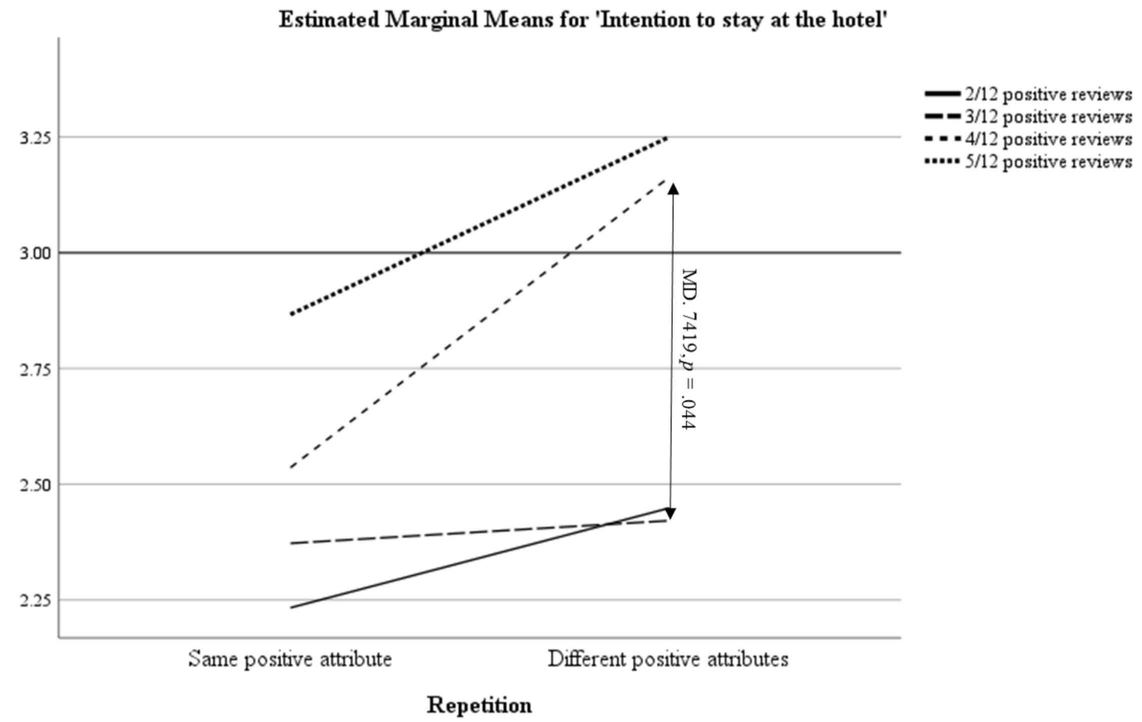



*Figure 3.1. Mean differences for the different conditions*

## Discussion

The current study explores the effect of the degree of positivity in predominantly negative review sets on behavioral intention and the moderating role of attribute repetition on this effect. It also explores how many positive reviews about important attributes are needed to compensate for a larger number of negative reviews about less important attributes. The results show that an increasing number of positive reviews enhances the intention to stay at the hotel. However, this intention only becomes positive when the positive reviews pertain to different attributes. When a review set only presents a single positive reason to stay at a hotel, it does not compensate for the multiple reasons to avoid the hotel. Repeating that argument cannot significantly increase people's intention to stay at the hotel. These results imply that, in the context of online reviews, repetition may not necessarily increase truth perceptions, as would be expected according to the truth effect (Dechêne et al., 2010). This finding is consistent with the findings of Ernst et al. (2017) that repeatedly presenting a message leads to a decrease in credibility judgments, which in turn leads to negative attitudes. Our findings do lend support to the repetition-variation hypothesis. They are consistent with the literature suggesting that more arguments enhance persuasion (Calder et al., 1974; Petty & Cacioppo, 1984; Willemsen et al., 2011b) because they make the message more complete and clearer to the reader. According to Hair and Bond (2018), the prominence of negative over positive attribute information depends on attribute importance: when positive reviews are about important attributes, even a relatively smaller number of positive reviews in the set is enough to compensate for the effect of more reviews with negative information.

A noteworthy finding in our study is that there is a 'tipping point' at which positive reviews can compensate for a review set's overall negativity, but only when the positive reviews are about different attributes. This tipping point occurs when we move from 3 to 4 positive reviews in a set of 12 (and thus from 9 to 8 negative reviews). These results point at a nuance of both the bandwagon effect and the negativity bias (Carstensen & DeLiema, 2018; Rozin & Royzman, 2001; Wu, 2013). People do not necessarily follow the "majority" opinion and that negative reviews do not always carry more weight than positive reviews. Rather, the effect depends on the importance of the attributes and the inclusion of different arguments, which can be related to information relevancy and completeness. Previous research has already found that a single positive review can have a positive effect on consumers' attitude (Purnawirawan et al., 2015; Tata et al., 2020), review credibility, trustworthiness and helpfulness (Pentina et al., 2018), hotels' revenue (Phillips et al., 2017) and purchase intention (Tata et al., 2020). This effect can be explained by the fact that review readers will tend to dismiss the negative reviews since they are about less important attributes (Hair & Bond, 2018). Besides, as stated by the emotional value perspective (Xia & Bechwati, 2008), positive reviews enhance positive feelings evoked by the review, influencing the consumers' decision making. In this study, the majority of reviews were negative in all conditions and there was still a positive intention to stay at the hotel when the positive reviews were only 4/12. These results might be explained by the role of involvement. Previous research by De Pelsmacker, Dens, et al. (2018) shows that the influence of review text valence on evaluative responses is stronger for more highly involved people than for lowly involved individuals. Given the task at hand in the current study (decide on an all-inclusive resort where they would spend their holidays), it is expected that the participants in this study were relatively highly involved with the task. Therefore, in light of De Pelsmacker, Dens, et al. (2018), a negative set with 4 positive reviews out of 12 will generate positive evaluations from the review readers since they were highly involved in the task of reading and assessing the reviews.

Our results show that the intention to stay does not increase steadily from condition to condition by adding a single positive review to the set. Other factors besides the increasing positive ratio of reviews might influence perceptions. For instance, including more reviews can contribute to increasing the sense of information completeness and lead to positive evaluations (Rucker et al., 2008; Shoham et al., 2017). When consumers have access to both positive and negative information to make their assessment, they are more likely to conclude that their attitudes are based on more complete information (Shoham et al., 2017). They will feel more confident in their hotel choice since it allows them to assess more accurately whether the hotel's weaknesses are acceptable and the strengths are good enough (Purnawirawan et al., 2015). Our research reinforces previous findings (Hair & Bond, 2018; Pentina et al., 2018; Wu, 2013) pointing at the volatility of the negativity bias in the context of online reviews by showing that negative online reviews about less important attributes are outweighed by positive reviews focusing on diverse important attributes. This encourages a new theorization on eWOM, exploring other characteristics in the online reviews besides valence.

## Implications

This study provides insights into the combined role of review valence ratio, attribute importance and attribute variation in review sets. Our findings contribute to further theory development, as it challenges and tests the bandwagon effect, negativity bias (two well-established psychological theories that are frequently used to explain behaviour in the context of online reviews) and the truth effect. First, consumers do not always follow blindly the majority opinion, as proposed in the bandwagon effect (Sundar et al., 2008), as in our study review readers show a positive hotel staying intention even when most reviews are negative. It seems that the bandwagon effect might only take place when review readers tend to peripheral cues and not when they access the review text. Second, we show that the negativity bias (Rozin & Royzman, 2001; Wu, 2013) does not completely explain how review readers assess positive and negative information. Using a negative set of reviews with a majority of negative reviews about less important attributes and a minority of positive reviews about important attributes, we established that other elements than the negativity or positivity of the reviews in a set are taken into consideration by review readers. This shows that consumers attend to the importance of statements in their decision making when buying a product or selecting a service. Third, the results on the effect of providing different arguments in a set of reviews indicate that information richness and completeness are also important determinants of consumer's intentions. We thus find that the truth effect (Dechêne et al., 2010) does not apply in the context of online reviews, as review readers prefer more diversified information than the same argument in all positive reviews. By showing nuances of these three well-established psychological mechanisms, we contribute to a better understanding of these effects and shedding light on how future studies should consider them when theorizing on the adoption of online review information.

In terms of practical implications, our findings show that businesses can be positively evaluated even when most reviews are negative. Receiving little to no negative reviews about your product or service is often considered an ideal scenario. Our research shows that brand managers should not necessarily fear negative reviews. Previous research shows that the inclusion of an irrelevant negative review in a positive review set improves product evaluations (Shoham et al., 2017), which already suggests that a small amount of negativity is not necessarily detrimental and can even benefit a business. Our study now shows that this can be the case even when there are more negative than positive reviews, as long as the positive reviews are about more important attributes than the negative ones. As such, practitioners should adopt strategies to incentivize positive eWOM about important product attributes (Wang et al., 2018). Moreover, online review managers should encourage diversity in online reviews, for instance, by asking reviewers to comment on aspects neglected in previous reviews. This could be automated in review platforms using artificial intelligence by generating a list of attributes that are not being mentioned commonly or recently and presenting the reviewer with this list as a suggestion of aspects to comment upon. Besides, when selecting testimonials from online reviews to be shown on the website, practitioners should try to diversify the arguments picked to be displayed. For instance, in the context of resorts, instead of displaying several reviews mentioning the cleanliness of the room, practitioners should select diversified reviews that mention the quality of the food or the pool amenities. It is not sufficient to merely increase the number of positive reviews, the reviews should also preferably highlight different strengths. Companies should understand consumers' critical decision criteria and strive for excellence in more than one of these. This will increase the chance that customers will mention different important attributes in their positive reviews, which can compensate for negative reviews about less important attributes. The findings imply that businesses should not merely focus on review valence but also on the importance and variety of their arguments.

## Limitations and future research

The current study has some limitations that offer opportunities for further research. The first limitation is that the use of a scenario means that the importance of the attributes must be read in light of that specific scenario. Other studies manipulating attribute importance could opt for other scenarios and contexts (e.g., for other services or products), which would also contribute to our findings' generalizability. Moreover, future studies could add other variables such as characteristics of the relationship between the writer and the reader of the review (e.g., homophily or tie strength), personality traits of the respondents, or other review cues such as star or usefulness ratings. Considering the previous findings that multiple sources presenting multiple arguments enhances persuasion over having either alone (Harkins & Petty, 1981), more research should be devoted to understanding the effect of source credibility on how review readers interpret negative information. Previous studies in political communication show that repeatedly presenting the same posters resulted in a negative attitude toward the presented political issue, mediated by a decrease in credibility judgments (Ernst et al., 2017). This mechanism of information adoption where credibility moderates the effects on intentions can also apply to online reviews and should, therefore, be further studied. Future research could also focus on how the tipping point we found (4 out of 12 versus 3 out of 12 positive reviews) evolves in larger or smaller review sets. Further research could also study other valence ratios and expand our findings to positively valenced review sets, as well as look at the effect of attribute repetition in the negative reviews. Exploring varying ratios of positive and negative reviews in predominantly positive review sets to test the 'positivity effect' (Shoham et al., 2017) would allow to find nuances on when a positive review set might harm intentions. These studies would enable to further refine how valence affects consumers' attitudes and behaviors toward products and brands. Other moderators should also be studied to improve our understanding of the nuances to the bandwagon effect and negativity bias. For instance, Wu (2013) studies how review quality (i.e., readability) and quantity (i.e., length) can attenuate or even reverse the negativity bias in the context of eWOM. Therefore, other studies can focus on these or other moderators when studying how review sets influence intentions.

Finally, the participants in this study only saw the review text. Further research should investigate if the nuances that we found for the bandwagon effect still hold when other review characteristics are present. When people have a large number of reviews, it may become impossible to read them all (Park & Lee, 2008). In such a case, they could rely on the majority's opinion by considering aggregated information. For instance, 60% of the users rated the hotel negatively (regardless of the actual content of these reviews).

# 4. Managerial response strategies to eWOM: a framework and research agenda for webcare[9]

[9] Manuscript under review as Lopes, A.I., Dens, N., De Pelsmacker, P., Malthouse, E.C. Managerial response strategies to eWOM: a framework and research agenda for webcare

## Abstract

Managers increasingly address client feedback online, a practice known as webcare. Based on previous research on webcare, this review provides a framework that aims to identify potential generalizations, discuss possible explanations for inconsistencies that require further investigation, and identify the under-researched areas concerning the managerial responses to online reviews. This framework answers several practical and theoretical questions on eWOM (electronic Word-of-Mouth). Should practitioners respond to eWOM or not? If they do respond, what kind of eWOM should they respond to and what strategies should they use: who should respond, when, on what platforms, in what style? How should they specifically respond to negative reviews? Future research should disentangle the many contradictory effects (e.g., when to use defensive webcare) and cover under-researched topics (e.g., webcare strategies for Positive WOM specifically or the underlying mechanisms explaining the effects of different webcare strategies).

# Introduction

A fundamental change in the field of marketing occurred when consumers became empowered by digital media to easily communicate with and about firms, for instance, through electronic word-of-mouth (eWOM), creating the need for brands to manage these communications (Deighton & Kornfeld, 2009). Many studies exist on the antecedents and effects of eWOM in general and online reviews in particular (e.g., De Keyzer et al., 2017; de Matos & Rossi, 2008). Several authors attempt to consolidate the knowledge on the effects of eWOM on consumer responses and sales through literature reviews (e.g., King et al., 2014) and meta-analyses (e.g., Babić Rosario et al., 2016; Purnawirawan et al., 2015). The last 10 years document dozens of papers on a track of eWOM research that studies the effects of webcare. Webcare is defined as the act of engaging in online communication to address client feedback (Edwards & de Kool, 2015). While much research on webcare focuses on mitigating the effects of NWOM (negative eWOM) (Dens et al., 2015; Van Noort & Willemsen, 2012), webcare is, in fact, an integrative organizational tool combining customer care, public relations and marketing that can increase consumer engagement (Edwards & de Kool, 2015; Schamari & Schaefers, 2015; Van Noort et al., 2015). Considering this definition of webcare, we will study strategies applicable to eWOM and online reviews in particular. We will use the terms eWOM and online reviews interchangeably as online reviews are a form of eWOM frequently studied in previous literature.

Previous research shows that providing webcare positively affects consumers' attitudes, intentions, brand evaluations and, consequently, business outcomes (e.g., Colliander et al., 2015; Sheng et al., 2019; Wang & Chaudhry, 2018; Xie et al., 2016). As such, webcare is an essential tool to mitigate the negative influence of NWOM and boost PWOM (positive eWOM). Although webcare is originally directed at the person who provides the eWOM (e.g., a reviewer), it also affects other consumers ("bystanders") that read the eWOM and managerial response (Kim et al., 2016; Wang & Chaudhry, 2018).

Previous studies on webcare tackle different aspects of responding to eWOM. Many studies focus on the effects of responding versus not responding to online reviews (e.g., Proserpio & Zervas, 2017; Van Noort & Willemsen, 2012; Wang & Chaudhry, 2018). As mentioned, overall, responding to eWOM seems to elicit positive effects. Importantly, the effects of webcare will depend on who responds to the review, the response timing, what is said (the tone, length, …), among other things. It will also matter whether the original review was positive or negative. A response strategy, or webcare strategy, refers to the type of answer employed by businesses to reply to online reviews (e.g., apologize for mistakes). Previous research is often inconclusive about what are adequate webcare strategies and the extent to which they deliver results.

Previous literature reviews on webcare (Stevens et al., 2018; Van Noort et al., 2015) focus on a limited number of webcare characteristics in response to NWOM. Van Noort et al. (2015) focus on timeliness, content (e.g., apologizing and taking corrective action), and stylistic elements (e.g., use of conversational human voice, message personalization) of the response. Stevens et al. (2018) focus their literature review on three principles that managers should base their webcare on; timeliness, transparency and trust. The current review provides an integrative framework for webcare, identifying multiple components of webcare strategies in response to both NWOM and PWOM and structuring the processing variables. This approach allows us to (1) identify potential generalizations from the findings of previous research, (2) discuss possible explanations for inconsistencies that need to be further explored, and (3) identify the under-researched areas with respect to managerial responses to online reviews. Such framework would provide guidance to managers seeking to manage and respond to eWOM. With the increasing volume of eWOM, it is crucial for organizations to know whether and how to invest their efforts in webcare to achieve positive business results.

After explaining the different steps for this systematic literature review, we develop the review along the lines of the categories shown in Figure 4.1. For each category, we will provide an overview of prior findings. In the concluding section, we establish which findings are robust across studies, focus on the issues on which there is a discrepancy among studies and highlight under-researched strategies.

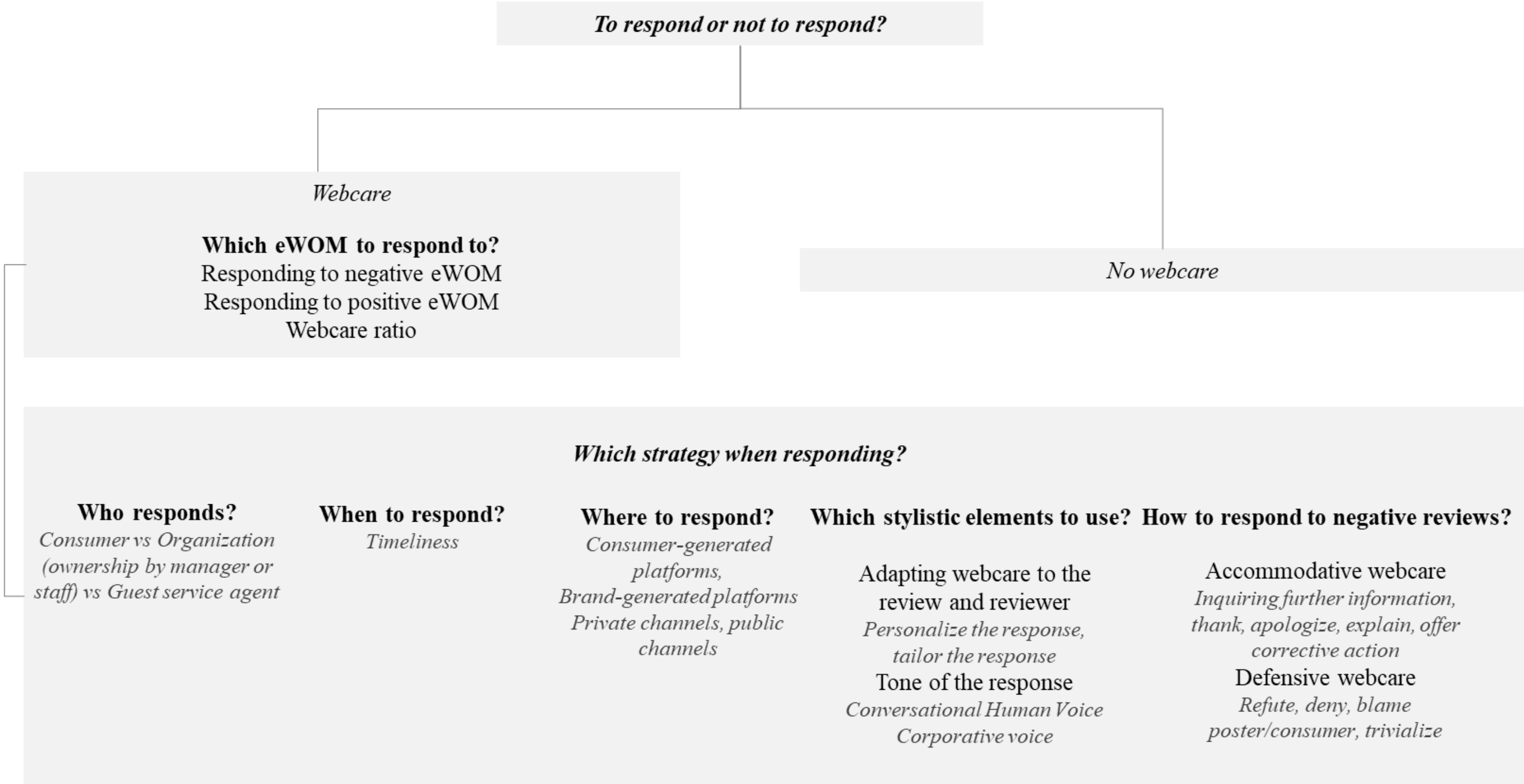



*Figure 4.1. Conceptual framework of webcare*

## Literature search

To conduct this literature review, we consulted papers that adopt different methodological approaches from different research fields within marketing. The literature search started by looking for articles published under the keywords of *managerial responses to reviews*, *webcare*, *service failure* and *service recovery*, *complaint handling* and *complaint recovery*, *online communities* and *online firestorms*, *response strategies to online reviews*, *service intervention*, *reputation management* and *customer care* in *Google Scholar* and *Web of Knowledge,* from 2000 until 2020. After identifying a study, we examined its references to find further studies. The initial sample is composed of 97 articles. From this sample, we selected all papers in English that reported empirical results and were related to managerial responses to eWOM or online reviews. A final list of 71 articles was retained. An overview of the studies included in this literature review can be found in appendix 3. Many of the papers included in this literature review concern the hospitality industry; the reason for this is that online reviews, and webcare by extension, play a decisive role in business success within this industry (Sheng et al., 2019; Xie et al., 2014). Although most of the papers (around 70%) are centered exclusively on webcare in response to NWOM, we will cover previous studies regardless of eWOM valence.

## Which eWOM to respond to?

After receiving an online review, the first issue to deal with is deciding whether or not to provide webcare (i.e., to respond). Considering that this is the first issue organizations have to deal with, the effect of providing webcare or not has been widely studied. The studies cited in this section do not make an explicit distinction between webcare towards PWOM or NWOM. They look at the effects of responding to eWOM in general. The valence of this eWOM could often not be determined because, for instance, the authors used aggregate data across multiple reviews, making it is impossible to see the effects of individual replies.

According to previous studies, providing webcare positively affects review helpfulness (Kwok & Xie, 2016), perceived review credibility (Kniesel et al., 2016), consumer sentiment (Homburg et al., 2015), future review ratings (Proserpio & Zervas, 2017; Sheng et al., 2019; Wang & Chaudhry, 2018; Xie et al., 2016), future review volume (Chen et al., 2019; Proserpio

& Zervas, 2017; Sheng, 2019; Xie et al., 2016), and subsequent review length (Proserpio & Zervas, 2017). For instance, Proserpio and Zervas (2017) find that once hotels start providing webcare, the volume of subsequent negative reviews decreases, although the average length of negative reviews increases. The authors suggest that this might indicate that dissatisfied consumers become less likely to leave short refutable reviews when they expect hotels to respond. Ma et al. (2015) find that webcare incentivizes future voicing (e.g., posting of reviews), which is also related with review volume. Webcare also influences reviewers by increasing their motivation to post subsequent reviews (Chevalier et al., 2018).

While there is much literature to suggest a positive effect of webcare on consumer perceptions, the effect on business outcomes is less clear. For instance, Anderson and Han (2016) find that not responding to online reviews (compared to responding) harms hotel revenue. In contrast, Xie et al. (2016) did not find any significant effect of webcare on hotel performance, while Xie et al. (2014) even find a negative effect of responding to online reviews on business performance (RevPAR). Other research shows that a response is not required when only a minority of reviewers are dissatisfied (Dens et al., 2015). Lui et al. (2018) find that responses have a more substantial positive impact when they address extreme reviews.

### *Responding to negative eWOM*

Most prior research focuses on the effects of answering NWOM. Most research on this topic frames the importance of responding to negative reviews on the social or restorative justice theory (e.g., Liu, Jayawardhena, Dibb, et al., 2019). According to this theory, by offering webcare, brands are restoring justice in a situation in which customers feel like they were treated less than equitably.

There is a strong consensus that there is a positive effect of responding (compared to not responding) to negative reviews. From the perspective of the reviewer, it increases complaint satisfaction (Einwiller & Steilen, 2015), customer satisfaction (Gu & Ye, 2014), reviewer motivation to post (or eWOM continuance) (Chevalier et al., 2018; Liu, Jayawardhena, Dibb, et al., 2019). Webcare in response to NWOM also improves reviewers' relationship with the brand (Ma et al., 2015), their attitude towards the organization (Anderson & Han, 2016; Liu, Jayawardhena, Dibb, et al., 2019), and behavioral intentions (i.e., recommendation intentions) (Kim et al. 2016). Interestingly, webcare to negative reviews harms customer satisfaction for people who see managerial responses to previous reviews without receiving a response to their own negative review (Gu & Ye, 2014).

From the perspective of review readers or bystanders, webcare to negative reviews positively influences future ratings (Wang & Chaudhry, 2018), brand evaluations (Anderson & Han, 2016; Van Noort & Willemsen, 2012; Weitzl, 2019), brand reputation (Rose & Blodgett, 2016), attitude towards the company (Esmark Jones et al., 2018), and purchase intentions (Casado-Díaz et al., 2020). It also boosts trust and diminishes concerns (Sparks et al., 2016) and benefits business performance (Kim et al., 2015; Xie et al., 2017). Only Bhandari and Rodgers (2018) find a negative effect of webcare to NWOM on review readers' purchase intentions.

Finally, previous research shows that, it might not be worth it to respond to NWOM in certain circumstances, when no webcare strategy seems to mitigate adverse reactions. For example, in the case of vindictive complaints, complaints by 'revengeful loyalists' (committed, revengeful customers mainly driven by webcare-independent motives) or when there are multiple failures (Weitzl et al., 2018; Weitzl & Einwiller, 2020). We discuss specific strategies to respond to negative reviews later in this paper.

### *Responding to Positive eWOM*

Considering that most eWOM posted online is positive (Chevalier & Mayzlin, 2006; Resnick & Zeckhauser, 2002), the literature on the effects of responding to PWOM is quite scarce and sometimes contradictory. For instance, Schamari and Schaefers (2015) find a positive effect of webcare directed at positive reviews on observers' brand engagement intentions. According to these authors, other consumers who can see the interaction could perceive webcare directed at positive reviews as a "reward" for the comments (Schamari & Schaefers, 2015). Other authors find a negative effect of webcare to positive reviews on future review ratings (Anderson & Han, 2016; Wang & Chaudhry, 2018) and sales revenue (Li et al., 2018). J. Wu et al. (2020) find that the positive effect of responding to positive reviews might actually depend on the content and style of the response. According to these authors, an active-constructive response (validating the good experience being shared and showing enthusiastic support for future events) increases consumer repurchase intention, while a passive-constructive (giving understated, minimal support) does not. Using a friendly communication style (vs. official style) reinforces an active-constructive response's positive effect (J. Wu et al., 2020). This is, to our knowledge, the only paper focusing on specific response strategies to positive online reviews.

### *Webcare ratio*

The next question is: what is the optimal ratio (fraction of reviews that gets a response) of responses? In a qualitative study, Park and Allen (2013) investigate how hotel managers' perspectives about online reviews link to how often they engage in webcare. The authors find that hotels that respond to more reviews consider them to be an honest gauge of consumer sentiment, and promote regular meetings and consultations with the internal staff to discuss their content (Park & Allen, 2013). On the other hand, the 'non-responders' believe that reviews represent only extraordinarily positive or negative views and typically rely on external corporate managers to handle social media (Park & Allen, 2013). These findings indicate that the amount of webcare signals the level of commitment that brands put into providing webcare.

According to Homburg et al. (2015), responding to online reviews benefits consumer sentiment, but shows diminishing returns with an increased response rate (the percentage of reviews that receive a response). In the same vein, Anderson and Han (2016) find that the effect of webcare in TripAdvisor positively influences hotel revenue, but only up to a response rate of about 40 percent. Higher response rates are detrimental, and hotel revenue declines (Anderson & Han, 2016). These authors find that when managers respond to more than 85 percent of reviews, revenues are lower than if they do not respond at all (Anderson & Han, 2016). Xie et al. (2016) show that the response rate reinforces the review rating effect on hotel sales. Other authors do not mention diminishing returns in their findings that increasing the response volume increases review volume (Sheng, 2019) and a firm's competitive performance (Lui et al., 2018).

# Which strategy when responding?

## *Who responds?*

Platforms such as Booking.com allow only hotels to reply to their respective reviews. Other platforms, such as TripAdvisor or Google, allow other consumers to participate in the dialogue by commenting on others' reviews. Previous research studies the effects of reading a response by the business itself (webcare) versus by other consumers. For bystanders, a managerial response is perceived as less trustworthy than a response by another consumer, hurts the attitude towards the company (Esmark Jones et al., 2018), and leads to lower purchase intentions (Brunner et al., 2019; Esmark Jones et al., 2018; Ullrich & Brunner, 2015). Interestingly, this might not be true for all brands. Brunner et al. (2019) show that the source effect (brand vs customer) on readers' purchase intentions is moderated by brand strength: if a strong brand responds to a negative review, the purchase intentions of the bystanders are similar to those generated by a customer's response. In contrast to most other studies, Weitzl and Hutzinger (2017) find that managerial responses (vs consumer responses) have a more positive effect on brand responses (brand attitude and trust, purchase intention, PWOM intention). Although not controlled by the firm, consumer responses to reviews should not be ignored by organizations, as their effects can be very beneficial. For instance, previous research shows that having other consumers replying to reviews enhances the effects of webcare for negative reviews (Jiang et al., 2019; Schaefers & Schamari, 2016).

Tathagata and Amar (2018) explore how 'webcare ownership' (the person/team within the organization that responds) reflects the organization's acceptance of responsibility. The authors find that webcare towards negative reviews is more credible (for reviewers) with high ownership – provided by an individual with personal details (such as name and designation) – than low ownership –provided by a team representing the company (Tathagata & Amar, 2018). By seeing a name, consumers have the opportunity to attribute their blame to this identifiable person and therefore tend to forgive the firm even after a failure (Tathagata & Amar, 2018).

Apart from the degree of personal detail about the person responding, this person's role inside the organization may also matter. A secondary data analysis from Xie et al. (2017) finds that webcare provided by hotel executives lowers future financial performance compared to webcare provided by the staff. In contrast, an experiment by Kniesel et al. (2016) shows no significant difference in bystander's attitude towards the brand between answers from managers and staff members.

### *When to respond?*

Deciding on whether to provide webcare opens the door to other questions regarding 'when' to provide it. Previous findings on the effects of timeliness on both reviewers and bystanders seem to be consistent across studies. Giving a timely response to a complaint or negative review, compared to a late response, positively influences reviewers' perception of justice (Gelbrich & Roschk, 2011). However, it does not seem to significantly affect complaint satisfaction (Einwiller & Steilen, 2015; Min et al., 2015). Timely responses also benefit future review volume (Sheng, 2019; Sheng et al., 2019) and financial performance (Xie et al., 2017). When most reviews are negative, a timely response increases readers' trust, decreases their concern (Sparks et al., 2016), and leads to higher levels of forgiveness (Ghosh, 2017).

### *Where to respond?*

There is also the matter of 'where' to perform webcare. Van Noort and Willemsen (2012) find that the platform (consumer vs. brand-owned blog site) where brands provide webcare makes a difference. Both reactive (posted in reply to a customer's request to respond to their complaint) and proactive webcare (not preceded by any direct or indirect requests to respond) in response to NWOM benefit readers' evaluations of a company compared to no response. The effect of responding is less prevalent on a consumer-generated than on a brand-owned blog site (Van Noort & Willemsen, 2012; Van Noort et al., 2015).

Kemp et al. (2020) find that one of the main drivers for managers to engage in proactive webcare is the socially prescribed perfectionism: managers feel that their business is under scrutiny and that they need to be perfect, therefore all brand-related messages posted online need to be addressed. However, on social networking sites (typically "reserved" for consumer-to-consumer interactions), proactive webcare can be seen as unsolicited and lead to substantial feelings of privacy infringement (Demmers et al., 2018). Grégoire et al. (2015) suggest that firms should publicly contact complainants and invite them to engage in a private conversation. Channel changes can be active (the person providing the webcare actively transfers the complaint to another channel) or passive (the reviewer is requested to contact the company through another channel). In many cases, a private response (not made available for review readers) is appropriate for dealing with consumer feedback (Zhang et al., 2019). Private responses could help avoid the virality of negative information (how much a message spreads online) or online firestorms (NWOM that receives substantial support from other customers in a short period) (Zhang et al., 2019). Herhausen et al. (2019) argues differenty, finding that suggesting a channel change further fuels the storm at an evolved stage of an online firestorm. According to these authors, changing the channel initially blocks and disengages from elaborate online discussions, but at a later stage, once the NWOM has gathered support within an online brand community, these disengagement approaches are not advisable.

## Which stylistic elements to use?

A lot of research has been conducted on the "ideal" style of webcare, especially in response to negative reviews.

### *Adapting webcare to the review and reviewer*

Regarding the way to respond, previous research looked into message tailoring and personalization. While message tailoring refers to adapting the response to individual reviews, message personalization means that a response includes personal information of the reviewer or the respondent (e.g., the reviewer's name or the name of the person undersigning the webcare). As message personalization can be an element of the tone of voice of a reply (Van Noort & Willemsen, 2012), which we address in the next section, we only focus on tailoring here. Previous research mostly shows that specific (or tailored) management responses to negative reviews lead to more positive outcomes than generic responses (Raju, 2019; Wei et al., 2013). Liu et al. (2015) also find that a targeted (i.e., tailored) response strategy significantly improves online hotel ratings. Lappeman et al. (2018) show that, when faced with the option of replying to a cluster of complaints or each individual complaint, individual replies engender a more positive brand reputation. Min et al. (2015) find that paraphrasing a complaint (a form of tailoring) in response to a negative review will cause potential guests to evaluate the response more favorably than a response that does not paraphrase the complaint. The only study we found documenting an adverse effect of webcare tailoring is by Xie et al. (2017), who show that message tailoring in response to negative reviews harms financial performance.

### *Tone of the response*

The style or tone of voice of webcare also matters. Sparks et al. (2016) find that, compared to using a more professional tone of voice, using a more conversational human tone increases bystanders' trust and makes them less concerned about the problem expressed in the review. A conversational tone can be achieved by personalizing responses, which transmits a more humane treatment and fosters feelings of trust (Stevens et al., 2018). In the same fashion, Crijns et al. (2017) find that personalizing a response to negative reviews positively affects brand reputation through higher perceptions of conversation human voice and sequentially lowers consumer skepticism. However, this effect is not significant with a personalized response to positive reviews. This might explain why Kniesel et al. (2016) did not find that using a more humane tone (vs corporate tone) leads to a more positive brand attitude from review readers, as they study this effect across both negative and positive reviews. In contrast, Sheng et al. (2019) find that showing little sentiment in webcare increases subsequent ratings.

Other studies find that an empathetic voice increases the perceived usefulness of the review (Liu & Ji, 2019) and complaint satisfaction (Javornik et al., 2020; Min et al., 2015). Herhausen et al. (2019) find that showing empathy diminishes virality in the first stages of an online firestorm, but has the opposite effect in later stages. Einwiller and Steilen (2015) report that the use of empathy is uncommon in webcare.

### *Length of the response*

Previous research shows that webcare length is positively related to review volume (Sheng, 2019), the perceived helpfulness of a review (Liu & Ji, 2019), and financial performance (Xie et al., 2017). The effect of webcare length might depend on the valence of the review. Chen et al. (2019) find that managers should provide detailed (longer) responses to negative reviews, but brief ones to positive reviews. According to the authors, detailed webcare to positive reviews may unduly emphasize the few negative points mentioned in the (overall positive) reviews, which may undermine the positive influence of webcare (Chen et al., 2019).

## How to respond to negative eWOM

Previous studies mainly classified responses to NWOM as either accommodative – complaisant and comprising corrective action, compensation and/or mortification – or defensive – denial and evasion of responsibility (Einwiller & Steilen, 2015).

### *Accommodative webcare*

Accommodative webcare can be provided in many ways (or a combination of them): showing understanding, inquiring further information, expressing gratitude, offering an apology, providing explanations, and taking corrective actions (e.g., offering compensation). Lee and Cranage (2014) find that, in a negative set of reviews, accommodative responses are more effective than no response at preventing negative bystander attitudes. According to Einwiller and Steilen (2015), inquiring further information is the most common strategy in practice, but does not lead to satisfaction with how the complaint was handled. Expressing gratitude for the review, the second most common webcare strategy, in contrast, does lead to complaint satisfaction.

Apologizing is one of the most commonly used and studied response strategies to NWOM (Nghiêm-Phú, 2018; Zhang & Vásquez, 2014). Despite the widespread use of apology, there is little evidence that merely apologizing is sufficient to deliver positive outcomes for brands. Kim et al. (2016) found that bystanders that see webcare towards NWOM containing an apology have less negative behavioral intentions than those who do not. On the other hand, Dens et al. (2015) test different accommodative and defensive responses to negative reviews and find that offering only an apology does not raise readers' attitudes or patronage intentions significantly compared to no response, even when most reviews are positive. Einwiller and Steilen (2015) find that apologizing does not significantly correlate with reviewers' complaint satisfaction. Herhausen et al. (2019) find that, overall, apologizing is useful to help avoid the virality of NWOM, but only when the crisis has just started and not when the negative situation is already spread out.

Managers can also attempt to provide explanations for the service or product failure. According to Einwiller and Steilen (2015), complainants who receive an explanation are not more satisfied than those who do not. Notably, the quality of the explanation matters: webcare containing strong explanations can produce higher consumer forgiveness compared to less plausible explanations (Ghosh, 2017; Tathagata & Amar, 2018).

Perhaps better than merely explaining is taking some form of corrective action, assuring reviewers and readers that the failure will not occur in the future. The use of this strategy seems to foster complaint satisfaction (Einwiller & Steilen, 2015), and benefits readers' purchase intention and brand perceptions (Treviño & Castaño, 2013) as well as brand reputation (Rose & Blodgett, 2016). According to Sparks and Bradley (2017), one of the most effective corrective actions is offering a compensation (e.g., a discount on a future purchase). Previous research shows that offering compensation leads to a higher perception of justice, which positively affects customers' future behavioral intentions (Ha & Jang, 2009). However, this strategy might not be appropriate in all circumstances. Herhausen et al. (2019) find that offering compensation only mitigates the virality of NWOM when used in evolved stages of online firestorms. Valentini et al. (2020) find that monetary compensation (vs. a voucher or a free product/service) is the only tool that can attenuate negative emotions from complainers, although it does not boost positive emotions.

Previous research has studied how combining different accommodative response strategies might lead to more positive customer reactions. Rose and Blodgett (2016) find that an apology with the assurance of future satisfaction and an apology with corrective action notification equally boost company reputation. Dens et al. (2015) find in an experiment that the optimal combination of response strategies depends on the review set balance. When most reviews are negative, more effort from the organization is required to create positive attitudes and encourage behavioral intentions with review readers. In this case, the most effective response includes both an apology, explanation and compensation (Dens et al., 2015). Sreejesh et al. (2019) find that this same combination is needed to boost attitude and patronage intentions from review readers. These articles by Dens et al. (2015), Rose and Blodgett (2016), and Sreejesh et al. (2019) are the only ones to our knowledge that study the combination of different webcare strategies.

### *Defensive webcare*

Defensive webcare entails refuting what is written in the review, accusing the reviewer or a third party, or trivializing a complaint. Previous studies do not show an obvious negative effect of giving a defensive response compared to no response. Lee and Song (2010) find that a defensive response decreases the problem attribution to the company and positively affects company evaluation. Based on interviews with prospective hotel customers, Treviño and

Castaño (2013) find that hotels performing any type of webcare, even defensive, are perceived as giving more importance to customer service and guests than hotels that do not respond to negative reviews. Lee and Cranage (2014) find that when there is low consensus in a negative set of reviews (meaning that some reviews are positive), defensive responses are more effective than no response at preventing negative bystander attitude. In a high consensus situation, however, not responding is more effective.

Exploring specific defensive strategies, Weitzl and Hutzinger (2017) study the effect of accusing the customer, denying fault, accusing a third party, or trivializing the complaint on several brand outcomes. They find that the only strategy with a significant adverse effect on failure attribution is vouching (i.e., countering negative comments with favorable statements) (Weitzl & Hutzinger, 2017). Weitzl and Hutzinger (2017) show that credible, defensive responses might strengthen bystander-brand relationships. Dens et al. (2015) find that refuting negative reviews, a specific defensive reaction, is the worst response (compared to strategies such as apologizing, explaining, and offering compensation) when at least half of the reviews are negative. Similarly, Weitzl (2019) find that defensive responses stimulate future negative WOM. When reviews are overall positive, refutation is an adequate strategy to boost attitudes and patronage intentions, but not for PWOM intentions. Honisch and Manchón (2019) find that a humorous strategy (e.g. satire) is the least recommendable strategy, worse than refuting what is written in the review.

Scholz and Smith (2019) present a different perspective on defensive approaches, stating that ‘flyting’ (a ritualized exchange of insults between two or more interlocutors) can help brands bolster their ideological positioning. Their results signal that, in certain circumstances, defensive responses might be a valid strategy as long as it is consistent with the brand positioning.

### *Comparing accommodative to defensive webcare*

The majority of papers seem to agree that accommodative webcare is, in most circumstances, the preferred strategy. Bach and Kim (2012) explore how accommodative and defensive webcare links with business performance, showing that low-performing businesses tend to have a defensive approach. Studies find that, compared to defensive webcare, accommodative webcare exerts a more positive effect on the company evaluation (Lee & Song, 2010), reputation (Honisch & Manchón, 2019), PWOM intention (Xia, 2013), satisfaction (both from the reviewer and the bystander perspective) (Einwiller & Steilen, 2015; Xia, 2013) and booking intentions of bystanders (Casado-Díaz et al., 2020). Chang et al. (2015) also show that accommodative responses (vs. defensive responses) contribute to lower attribution of internal *locus* (if the failure is internal or external), enhancing organizational reputation and reducing negative WOM (Chang et al., 2015). When there are few failures, an accommodative response leads to the smallest attribution of *locus* (Weitzl et al., 2018). In contrast, Xia (2013) finds that the difference between providing defensive or accommodative webcare is not significant for readers' purchase intentions. The kind of reviews to which webcare is applied might help explain these contrasting results. For instance, Li et al. (2018) find that defensive webcare increases sales revenue when applied to 'ordinary reviews' (reflecting dislike, mismatched preferences, unrealistic expectations or occasionally unreasonableness on the part of the customer), while sales revenue decreases with accommodative answers to such reviews. In contrast, accommodative webcare increases sales revenue when the reviews mention product failures, while defensive responses to such reviews decrease revenue (Li et al., 2018). Another studied strategy is the provision of two-sided webcare: accepting some accusations and denying others, two-sided webcare presents a mix of arguments and counterarguments to the posted reviews (Tathagata & Amar, 2018). According to Tathagata and Amar (2018), two-sided webcare leads to higher forgiveness than its one-sided (accommodative, accepting all accusations) counterpart.

## Managerial implications

As we will discuss in detail in the next section, not many consistencies can be drawn from the studied papers. Despite this, we provide guidance for practitioners based on what previous research tends to find leads to positive outcomes.

To respond or not to respond? Although there is some disagreement (e.g., Xie et al. (2014) find a negative effect of responding on financial performance), most studies suggest a positive effect of responding to eWOM. Therefore, we suggest businesses that, when they have the necessary means, providing webcare is, by default, the best way to manage eWOM.

Which eWOM to respond to? Despite some contradictory findings, responding to NWOM has an overall positive effect. Therefore, we recommend managers to respond consistently to negative reviews since previous research has inclusively proven that it positively affects financial performance (Xie et al., 2017). There are circumstances, however, when managers can consider leaving some eWOM unanswered, especially if businesses have scarce resources: vindictive complaints, complaints by ‘revengeful loyalists’ (committed, revengeful customers mainly driven by webcare-independent motives) or when there are already multiple failures (Weitzl & Einwiller, 2020).

The few studies on webcare for PWOM present different findings for different variables. However, considering that the study from Li et al. (2018) shows a negative effect of responding to PWOM on sales revenue, we advise managers to leave positive reviews unanswered, especially in the case of scarce resources to reply in an personalized manner (J. Wu et al., 2020).

In terms of webcare ratio, previous research shows that there is little added value in answering all reviews (Anderson & Han, 2016; Homburg et al., 2015). Therefore, and considering our previous advice that there is no need to always reply to positive reviews, we advise managers to concentrate their webcare efforts in replying to NWOM. This advice might need to be revisited once there is more evidence regarding the effects of responding to PWOM.

Which strategy when responding? Who responds? Previous research shows many benefits of having other consumers replying to eWOM (e.g., Jiang et al., 2019). While this is not strictly within a firm’s control, businesses could think of setting up ambassadorship programs or other ways to encourage satisfied customers to join the conversation. Research shows that webcare (especially towards negative reviews) should have a high level of ownership (i.e., signed with the name of the person responding) (Tathagata & Amar, 2018). Also, we suggest that managers can delegate this task to their staff as previous research finds a negative effect on future financial performance when managers reply (Xie et al., 2017) and no significant difference in the attitude towards the brand between answers from managers and staff members (Kniesel et al., 2016).

When to respond? Regarding the response’s timeliness, the literature is consistent, as webcare given within a short time frame leads to the most favorable outcomes (e.g., Xie et al., 2017).

Where to respond? The literature related to the platforms on which to provide webcare does not present consistent results. Therefore, we advise managers to reply to eWOM regardless of the platform, but to bear in mind that responding to reviews can sometimes lead to privacy infringement feelings (Demmers et al., 2018). We tend to follow the advice from Grégoire et al. (2015) that firms should publicly contact complainants and invite them to engage in a private conversation, especially when the interaction is in an early stage (Herhausen et al., 2019). Providing webcare to NWOM communicates to bystanders that the firm cares about customer satisfaction (providing webcare with a request to change the channel should be better than seeing no webcare at all), while reviewers receive attention to their complaints. Nevertheless, when the NWOM is not necessarily problematic for future customers (e.g., the issue reported was not crucial and was already fixed), a public answer should be adequate.

How to write the response? The majority of previous research emphasizes a positive effect of tailoring webcare to NWOM (e.g., Lappeman et al., 2018). Therefore, we recommend the use of tailored webcare, even considering the findings from Xie et al. (2017) that message tailoring in response to NWOM negatively influences financial performance, as the authors also operationalized tailoring in their study as a repetition of the topics in the review. In terms of tone of the response, the overall recommendation based on previous studies (e.g., Sparks et al., 2016) is that webcare towards NWOM should use a conversational human tone, be personalized and show empathy. We would also recommend rather lengthy webcare to NWOM, providing details.

How to respond to negative reviews? Accommodative webcare (e.g., apologizing) is the strategy leading to the most positive outcomes (compared to defensive webcare) when managing NWOM (e.g., Bach & Kim, 2012) and should, therefore, be the preferred strategy to deal with dissatisfied customers. However, the use of defensive strategies can be preferable in some circumstances. Defensive webcare increases sales when applied to reviews that reflect mismatched preferences, unrealistic expectations or occasionally unreasonableness, while sales decrease when accommodative answers are applied to these reviews. On the other hand, accommodative webcare increased sales revenue when it was answering reviews mentioning product failures, while defensive responses to these reviews decreased revenue (Li et al., 2018). Therefore, the review characteristics should be considered before opting between an accommodative and a defensive strategy. A combination of both (presenting a mix of arguments and counterarguments to the posted reviews) can also be considered (Tathagata & Amar, 2018). When opting for an accommodative webcare, a combination of strategies (for instance apologizing, explaining and, when possible, offering compensation) leads to the most positive outcomes (Dens et al., 2015).

# Discussion and research agenda

Having analyzed the literature on webcare published in the last 20 years, in this section our purpose is to look into the consolidated knowledge on the different webcare strategies to find consistent results and to point out inconsistencies or gaps that indicate paths for further research.

## *Consistent findings*

Some consistent empirical findings emerge from this literature review. Webcare to negative reviews, regardless of the strategy used, generally brings more positive outcomes than not responding, and this response should be timely and personalized. Moreover, when a greater number of consumers are dissatisfied, accommodative answers are usually better than defensive answers. These generalizations can be useful for practitioners in implementing policies to manage online reviews, as described above.

## *Inconsistent findings and under-researched topics*

Many possible research questions emerge from the inconsistencies found in previous literature in terms of specific webcare strategies. Additionally, several important topics are under-researched. Table 4.1 shows a summary of the topics in our proposed research agenda.

### *Inconsistent findings*

The first inconsistency relates to the effects of responding versus not responding. While some studies find a negative effect of not responding on future review volume (regardless of review valence) (Anderson & Han, 2016), others did not find a significant difference (Xie et al., 2016), or even a negative effect of responding (Xie et al., 2014). The reason for these differences need to be further investigated. Proserpio and Zervas (2017) find that webcare diminishes the future volume of negative reviews, while other authors (Chevalier et al., 2018; Ma et al., 2015) find that it increases future volume. These differences require further investigation, as the dependent variable in prior studies differed: is it the case that by responding to eWOM, managers are lessening the volume of future negative reviews while increasing the volume of future positive reviews? Previous research into the effects of webcare content suggests that it is not responding vs not responding influencing review volume, but rather what is written in the responses.

*Table 4.1. Topics for further research*

| Topic | Possible Research questions |
|---|---|
| | *Disentangle inconsistent findings* |
| Effects on review volume | Does webcare indeed diminish the volume of future negative reviews and increase the volume of positive reviews? |
| Webcare ratio | What is the "optimal" ratio of reviews for which to provide webcare? |
| Effects on business performance | Is webcare helping consumer attitudes and intentions but hurting business? How? |
| Channels for webcare | In what channel(s) is it beneficial to engage in webcare? How does the platform influence the effectiveness of webcare? |
| Tailoring | What is the best way to tailor webcare for PWOM and NWOM? |
| Accommodative vs. defensive | Are accommodative responses always better than defensive ones? When can defensive responses be appropriate? |
| | *Under-researched topics* |
| Responding to positive reviews | Does it make sense to respond to positive reviews? If yes, what are the appropriate strategies to respond to positive reviews? |
| Combining strategies | What are the effects of combining different webcare strategies in the same response? |
| Cultural differences | How is webcare understood across cultures and languages? |
| Professionalize webcare | Does hiring external professionals to provide webcare bring positive outcomes for the business? |
| Timeliness | What can be considered a timely response?<br>Does responding later bring negative consequences (and therefore is it not worth responding anymore), or will it just have a smaller positive effect? When is it 'too late' to answer? |
| Unravel mechanism | What are the mechanisms that underlie the effects of the different webcare strategies? |
| Reviewer or bystander | How does webcare affect reviewers versus bystanders?<br>Should managers focus their webcare on reviewers or bystanders? |
| Methodological diversity | How can qualitative approaches to study webcare bring different insights on how it is provided and perceived?<br>Can the findings for the hospitality industry be applied to other services and product categories? |

There are conflicting findings on the "optimal" amount of webcare. While Anderson and Han (2016) and Homburg et al. (2015) find that responding to too many reviews might be detrimental for brands, other authors present the positive effect of webcare as being linear (e.g., Sheng, 2019). Also, according to Gu and Ye (2014), once managers embrace webcare, they should do it consistently for all reviews to avoid problems with future guests who do not receive webcare. As such, future research should clarify the effect of the ratio of webcare (fraction of reviews that gets a response) on attitudes, intentions, and business performance, to inform managers on the amount of effort they should put into webcare.

There are also inconsistencies in terms of the effects of having different people responding to and signing webcare. While Xie et al. (2017) find that webcare by hotel executives lowers future financial performance (revenue, average daily rate, and occupancy), Kniesel et al. (2016) did not find a significant difference between the answers from managers or from staff members. As such, further research should investigate this discrepancy, by testing different boundary conditions that might explain it (we go further in our reflection on boundary conditions next, in Figure 4.2). For instance, with negative reviews, the severity of the failure might be an essential moderator to this effect, as complaints about more severe issues might require webcare from a member higher in the hierarchy.

Research regarding the channels in which webcare should take place also leaves room for further studies. For instance, the findings from Zhang et al. (2019) indicate that consumers do not have a preferred channel for webcare. In many cases, responding in private channels (email, messaging systems in social media or even through chat boots) is acceptable to consumers. Private responses could help avoid virality of negative information or online firestorms (Herhausen et al., 2019). More research should aim to understand how the platform or channel influences the effectiveness of webcare in different contexts (e.g. severe problems, early-stage or later on,…).

In terms of style, tailored webcare seems to lead to the most positive outcomes (e.g., Lappeman et al., 2018; Liu et al., 2015), but this is not consistent for all studies (Xie et al., 2017). The reason for this should be further investigated through well-controlled experiments, for instance. Also, more research is needed to determine the best way to tailor webcare to both PWOM and NWOM to maximize results for reviewers and bystanders.

In terms of content, it seems clear that, in most circumstances, offering accommodative webcare is superior to defensive webcare. However, by apologizing, which is accommodative, firms are assuming guilt in the reviewers' accusations (Lee & Song, 2010; Weiner, 2000, 2010). Besides, defensive responses have also proved to be better in some cases than not responding (Xia, 2013). Are there contexts in which defensive webcare responses actually bring positive outcomes to companies? Johnen and Schnittka (2019) showed that defensive responses could be superior in hedonic contexts, but inferior in utilitarian ones (Johnen & Schnittka, 2019). Findings from previous research (Dens et al., 2015; Lee & Cranage, 2014) show that the degree of consensus in a review set might also determine if accommodative or defensive strategies are preferable. As pointed out by Ro and Wong (2012), there are occasions when customers knowingly and incorrectly report service failures or make illegitimate complaints. In these cases, companies should be able to refute these complaints to overcome dishonest complaints. Scholz and Smith (2019) present an interesting perspective that ‘flyting’ allows brands to boost their ideological brand positioning. The circumstances under which defensive responses are appropriate should be further investigated. For instance, one could study what reviewers and review readers understand as good quality webcare in different circumstances, using tools like the webcare quality scale by Tathagata and Mandal (2020). Finally, a possible reason for the inconsistencies in the effects of the use of accommodative and defensive webcare strategies might be related to how these variables are operationalized in the different studies. For instance, accommodative responses include only apologizing, apologizing + offer compensation + explanation, or other combinations. Looking at previous research on the combination of different strategies, we know that these different approaches to an accommodative response might yield different outcomes.

*Under-researched topics*

The biggest gap in literature seems to be the lack of research dedicated to webcare strategies for positive reviews. Looking at the section dedicated to ‘What to write in the response’, it is clear that most content strategies are focused on NWOM. The only study that focuses on specific strategies to respond to positive reviews is by J. Wu et al. (2020) (active-constructive response vs. passive-constructive response). Because (potential) customers attribute such importance to negative reviews, researchers seem to consider studying webcare to NWOM as the obvious choice. However, as mentioned, positive comments account for the majority of reviews (Chevalier & Mayzlin, 2006; Resnick & Zeckhauser, 2002), so appropriate webcare strategies for positive feedback should not be neglected. As such, future research should address the lack of empirical studies on its effects. The first step should be to conduct exploratory research to uncover the strategies used to respond to positive reviews, since these categories have not yet been discussed in the literature. As such, either qualitative approaches or unsupervised machine learning techniques for topic modelling would be interesting exploratory strategies. After this first stage, the effects of these strategies on attitudes, intentions, behavior and business performance should be studied.

When considering webcare in practice, different webcare strategies are often combined by integrating different accommodative strategies or including accommodative and defensive elements in the same response. However, apart from the articles by Dens et al. (2015), Rose and Blodgett (2016) and Sreejesh et al. (2019), the combination of different webcare strategies have received litte attention in academic studies. As such, more research should be conducted to better understand the effect of combining webcare strategies. Also, more research is needed to find the best webcare strategy to deal with different types of reviews, not only in terms of valence, but also in terms of what is expressed by the consumers. For instance, further research should consider exploring the type of issue reported in the review (Zhang et al., 2019) before deciding the type of webcare to be provided. Other methodologies , such as secondary data analysis, could also be applied to explore what strategies are commonly offered together.

An under-researched aspect of webcare is how webcare is provided and understood across cultures and languages. Previous research shows that values (fundamental beliefs held by the managers) and culture (local culture and beliefs) are dimensions to be considered when looking at how businesses provide webcare (Mate et al., 2019). For instance, Cenni and Goethals (2020) find that webcare provided to negative reviews by Dutch and English businesses is similar (mostly accommodative), while Italian businesses tend to be more defensive. Further research should explore how businesses in other cultures provide webcare and how consumers from different parts of the world perceive these different strategies.

Besides, researchers should dwell on the effects of having outsourced guest service agents providing webcare (e.g., Revinate, HotelSpeaker, ReviewPro). Although the service provided by these companies is widespread within the hospitality industry, research is lacking to understand if professionalizing these services brings positive outcomes for the business. This is a task that might require combining quantitative and qualitative methods (e.g., interviews with hotel managers). Text mining only, for instance, may not fully capture if webcare is outsourced or not.

In terms of timeliness, the literature consistently reports that timely responses are the best for several outcomes. However, what can be considered a timely response? In practice, responding within 24 hours seems to be the threshold for timely webcare. However, does responding later bring negative consequences (and is it therefore no longer advised to still respond after that), or will it just have a smaller positive effect? When is it ‘too late’ to answer?

*Unravel the mechanism underlying the effects*

Another aspect that stands out, especially when analyzing Appendix 3, is the diverse range of mediators, moderators, and dependent variables studied. It seems that many variables can help explain how webcare affects readers and reviewers. However, this diversity leaves a gap in understanding which variables are the most important to focus on, for researchers and practitioners. According to De Pelsmacker et al. (2021), marketing communications' objectives (or variables) can be conceptualized as process (the extent to which customers should have processed a specific communication stimulus) versus effectiveness (the effect of the whole campaign on the brand or the organization) objectives. Transposing this thought to webcare, some of the studied variables, such as perceived fairness, can be considered process variables. On the other hand, some effectiveness variables are brand related while others refer to the impact of webcare on business, such as hotel bookings or product sales, being considered commercial outcomes. Figure 4.2 portrays the variables included in prior webcare studies in light of De Pelsmacker et al. (2021)'s conceptualization.

As we can see in Figure 4.2, many process variables can affect commercial outcomes directly or indirectly through brand-related outcomes. However, the mechanism through which webcare processing occurs is not clear: which strategies lead to which process variables? Which process variables lead to which brand-related outcomes? Which brand outcomes lead to which business outcomes? What the moderators of these mechanisms? It is relevant that further research unravels the mechanism for webcare to clarify the relationships between these variables, shedding light on the findings on the effects of different strategies to manage eWOM. For instance, it is undeniable that one of the most critical variables that webcare can affect is business performance. However, although webcare seems to positively affect consumer sentiment (Homburg et al., 2015) and future review ratings (Proserpio & Zervas, 2017; Wang & Chaudhry, 2018), among others, it hurts business performance (Xie et al., 2014). Is it possible that webcare is helping consumer attitudes and intentions but hurting business? Previous research shows that measuring intentions does not entirely reflect actual buying behavior (Morwitz et al., 2007). One of the intentions closely related to financial performance is purchase intention. Bhandari and Rodgers (2018); Esmark Jones et al. (2018) find a positive direct effect of webcare on purchase intentions while Bhandari and Rodgers (2018) find the opposite in their experiment. These different effects might be explained by the fact that Bhandari and Rodgers (2018) included brand trust as a mediator in their design and find that it mediates a positive effect on the dependent variable. Xie, So and Wang (2017) find that, when webcare is directed at negative reviews, it has a positive effect on financial performance.

Therefore, while there is a negative effect of webcare on (immediate) purchase intentions, the effect on (long-term) business performance is positive as trust is a dimension of brand equity. Further research should aim to disentangle these inconsistencies.

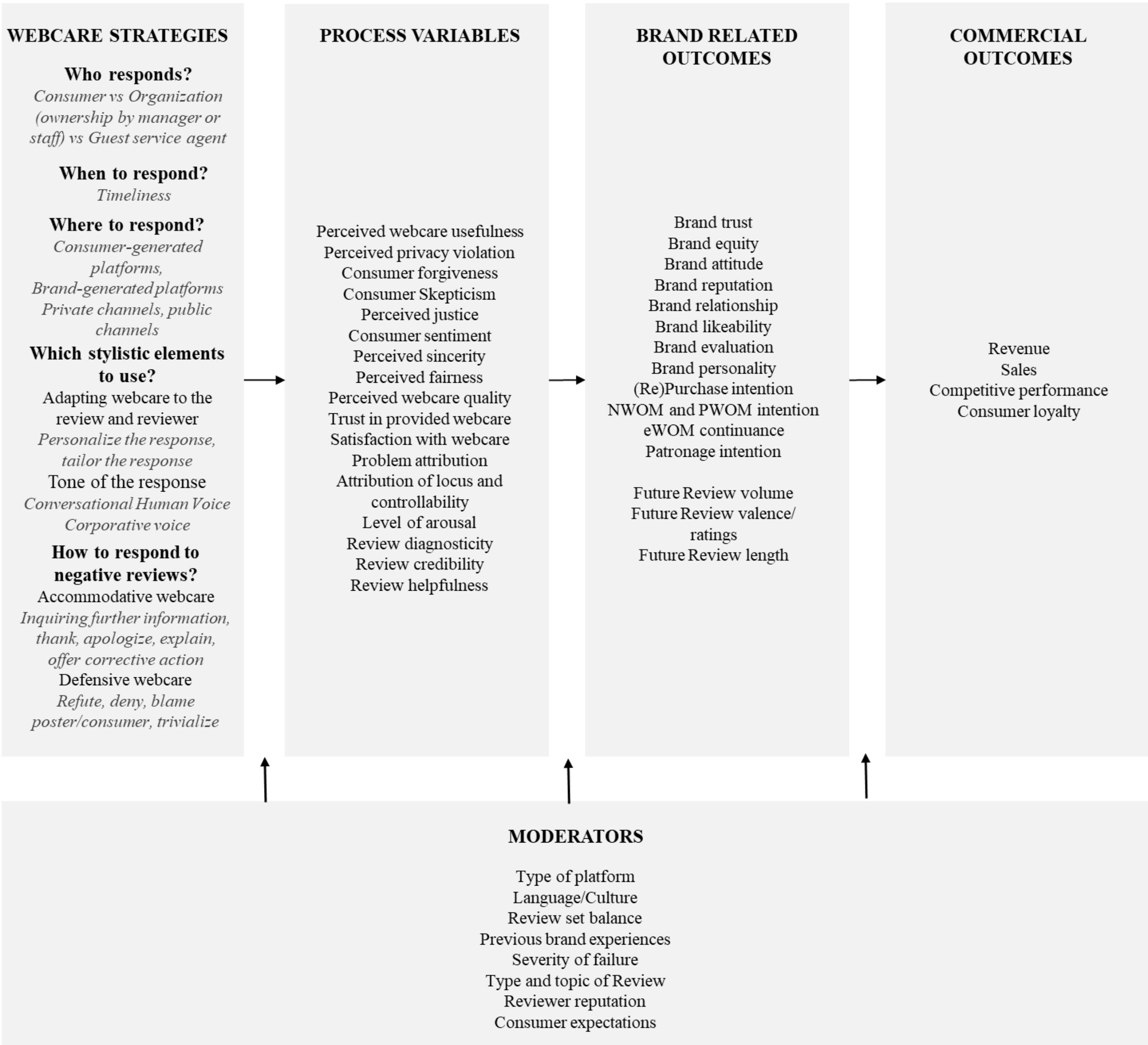



*Figure 4.2. Webcare related process and effectiveness variables*

Besides taking into account the mechanisms leading to positive and negative effects on commercial outcomes, researchers should clarify if the dependent variables used in their research are from the perspective of the reviewer, or from the perspective of review readers or bystanders. This will allow drawing more accurate conclusions on the effects of the webcare strategies.

In terms of methodology, we notice that most of the published studies opt for a quantitative approach, mostly experiments and secondary data analysis techniques. Although these methods bring interesting contributions and allow for empirical generalizations (Blattberg et al., 1995), more qualitative research should be applied to understand webcare. For instance, although the webcare provided by a certain firm can be analyzed through secondary data analysis techniques, the actual firm policies to respond to reviews might not entirely be revealed using these methods. For example, further studies should follow the steps of Homburg and Fürst (2007), whose paper on complaint handling explores if and how firms are taking actions based on the feedback they receive. Is the growing amount of feedback brands receive affecting the quality of the service/ products? Are there actual internal changes to the organization based on negative feedback? Or even: is it the webcare that drives the future ratings and review volume, or is it that the companies providing webcare are generally more attentive to customer feedback and therefore more likely to improve their service? Qualitative studies might unravel how these webcare strategies emerge in practice, by speaking with managers and companies that offer this service, to understand which guidelines (if any) they use to provide webcare. This comment on the methods is also valid the other way around. Qualitative studies, such as the one by Zhang and Vásquez (2014), should be followed by well-controlled experiments, in order to test the effect of the uncovered strategies.

Another noticeable aspect on the studies under analysis is that many of the 70 articles are studies on the hospitality industry. Previous research has shown that online reviews are especially important in the hospitality industry, as reviews about hotels and restaurants are extensively available and consulted by travelers (Yang et al., 2018). However, these strategies might also be applied to other services and product categories (e.g., search vs. experience) to see in which way the results hold or change depending on the product/service being reviewed.

# Conclusion

Knowing how to deal with the increasing volume of eWOM is a task that has kept academics and practitioners busy over the last decades. In this literature review, we find little consensus on the best strategies to deal with online reviews. Therefore, we suggest that academics keep developing studies on webcare to solve the inconsistencies and under-researched areas. This will help practitioners to implement strategies that will most strongly benefit their business. These studies are of the utmost importance, for instance, to develop automated responses like chatbots (Dao & Theotokis, 2021; Li et al., 2021; Liebrecht & van Hooijdonk, 2019). While companies with smaller review volume might tackle online reviews manually, other might need to apply advanced artificial intelligence algorithms to face the volume of reviews that they receive. In either case, knowing the best way to respond to each review is crucial to overcome the nefarious effects of NWOM and boost the positive effects of PWOM.

# 5. Is Webcare good for business? A machine learning approach to the effect of managerial response strategies to online reviews on hotel bookings[10]

[10] Manuscript under review as Lopes, A.I., Malthouse, E.C., De Pelsmacker, P., Dens, N. Is webcare good for business? A machine learning approach to the effect of managerial response strategies to online reviews on hotel bookings

# Abstract

Engaging in webcare – i.e., responding to online reviews – can positively affect consumer attitudes, intentions, and behavior. However, previous research is often scarce or inconsistent regarding the effects of specific webcare strategies on business performance. Therefore, we test whether and how several webcare strategies affect hotel bookings. After testing various machine learning classifiers, BERT yields the best performance for classifying webcare variables. The strategies that have a positive effect on bookings are directing reviewers to a private channel, being defensive, offering compensation and having managers sign the response. Webcare strategies to be avoided are apologies, merely asking for more information, inviting customers for another visit, and adding informal non-verbal cues. Strategies that do not appear to affect future bookings are expressing gratitude, personalizing, and having staff members (rather than managers) sign webcare. These findings help hotel managers to optimize their webcare strategy for better business results and develop automated webcare.

# Introduction

Electronic word-of-mouth (eWOM) has rapidly gained importance in consumer decision-making (Rosario et al., 2020). Apart from its effect on decision-making, eWOM is a major source of information that allows companies to understand consumer preferences or even predict financial performance or sales (Babić Rosario et al., 2016; Chevalier & Mayzlin, 2006; Ye et al., 2009). Online reviews are a form of eWOM in which users or experts evaluate products or services based on their own experience by giving specific suggestions, for instance, about restaurants, hotels, or attractions (Sotiriadis & Van Zyl, 2013). Previous research has shown that online reviews are critical in the hospitality industry, as reviews about hotels and restaurants are extensively available and consulted by travelers (Guo et al., 2017; Phillips et al., 2017; Tsao et al., 2015; Yang et al., 2018).

As the volume of eWOM increases constantly, it is crucial for organizations to know how to manage online reviews to achieve positive business results (Schamari & Schaefers, 2015; Williams & Buttle, 2011). Webcare - the act of engaging in online communication to address client feedback, including eWOM (Edwards & de Kool, 2015) - is an essential tool to mitigate the negative influence of NWOM (negative eWOM) and boost PWOM (positive eWOM) (Chevalier et al., 2018; Sheng et al., 2019). Although webcare is typically directed at the reviewer, it also has an effect on “bystanders” (potential customers) who read the review and the response (Wang & Chaudhry, 2018). Prior research highlights the potential of webcare to improve both reviewers’ and bystanders’ attitudinal and behavioral responses (Schamari & Schaefers, 2015; Van Noort & Willemsen, 2012).

Most commonly, studies on webcare focus on the effects of responding versus not responding to online reviews (Proserpio & Zervas, 2017; Van Noort & Willemsen, 2012; Wang & Chaudhry, 2018). Most studies indicate that responding to online reviews improves subsequent ratings and reviews (Wang & Chaudhry, 2018; Xie et al., 2016), reviewers’ and bystanders’ attitudes and intentions (Bhandari & Rodgers, 2018; Brunner et al., 2019; Tathagata & Amar, 2018), and consumer sentiment (Homburg et al., 2015; Ma et al., 2015). However, some studies come to a different conclusion. Bhandari and Rodgers (2018) find a negative effect of providing webcare on purchase intention, while Xie et al. (2014) find that webcare has a negative effect on hotel performance (RevPAR). Xie et al. (2016) did not find any significant effect of responding on hotel sales.

These discrepant findings from previous literature constitute a gap that needs further research, as they seem to indicate that there are aspects other than merely responding that play a role in terms of the effects of webcare. Previous studies point out that the effect of responding depends on the actual strategy used in the response (Dens et al., 2015; Van Noort & Willemsen, 2012; Weitzl & Hutzinger, 2017; Xia, 2013). A webcare strategy refers to the type of reply from a business to online reviews. Considering how vital eWOM is for business, the current study investigates the effects of webcare on hotel bookings as a measure of business performance. We study the effects of providing webcare versus not doing so, and we explore the effect of several specific webcare strategies: who should respond (i.e., sign the webcare message), how soon to respond, where to respond (i.e., private versus public channel), which stylistic elements to use (i.e., message tailoring and tone of voice), and which strategies to employ when responding to negative reviews in particular.

Despite being a common practice for businesses with a strong online presence, organizations often struggle to know which webcare strategies to employ to yield the best results (Van Noort et al., 2015), which demonstrates the need for this study. Moreover, many previous studies, particularly those looking into the effect of specific webcare strategies, measure attitudes or purchase intentions and not actual (buying) behavior. Previous research has shown that measuring attitudes and (purchase) intentions does not necessarily reflect or predict actual buying behavior (Morwitz et al., 2007). Considering the contradictory findings and the lack of knowledge on how webcare strategies actually affect business (hotel bookings in this case), this study builds upon past research on the effect of these strategies on attitudes and intentions by investigating how webcare in general and specific webcare strategies in particular influence (future) hotel bookings. We thus aim to answer the following research questions:

*How does webcare affect hotel bookings? How do specific webcare strategies affect hotel bookings?*

To answer these questions, we test the effect of webcare and several webcare strategies on the actual bookings received by seven hotels in the city of Antwerp (Belgium) through Booking.com for a period of four years (2016-2019). Most previous webcare studies are based on experimental designs with limited external validity and/or explore the effect of webcare on (future) hotel ratings or booking intention and/or are limited to one hotel only. The current study, therefore, contributes to the existing knowledge on webcare in several ways. First, it uses a machine learning approach to explore the effect of specific webcare strategies on the actual bookings of seven hotels while controlling for other factors that can affect business in the hospitality industry, such as bookings' seasonality. Second, this is, to our knowledge, the first comprehensive study that simultaneously looks into the effects of several common webcare strategies whose effects on business performance are under-researched or for which prior research presents contradictory findings. Besides looking into how webcare affects business performance, a third contribution is to develop automated machine learning tools and methods for coding webcare responses. Several machine learning text classifiers are thoroughly tested for the different variables in the study and their performance evaluated. This study, therefore, helps to develop automated webcare, for instance, chatbots, which are fundamental tools in the future of webcare management (Li et al., 2021; Liebrecht & van Hooijdonk, 2019; Navío-Marco et al., 2018). We will conclude with insights for further research, hotel (webcare) managers, and review platform managers.

## Literature Review

Managerial responses to online reviews are not only read by those who have written the review but also by other (prospective) customers ("bystanders") who read the reviews and responses (Wang & Chaudhry, 2018; Weitzl & Hutzinger, 2017) and use this information to decide if they will buy the reviewed product or service. This effect of webcare on bystanders can be explained by the social learning theory, which postulates that individuals predominantly learn from observing others' behaviors (direct reinforcement) and/or the consequences of those behaviors (vicarious reinforcement) (Bandura & McClelland, 1977). Applying this theory to the context of online reviews and webcare, bystanders learn from online reviews and managerial responses to them. This explains why bystanders as potential customers are motivated to read webcare (Schamari & Schaefers, 2015), to help them build their attitude towards the brand (Weitzl & Hutzinger, 2017), develop trust perceptions (Ku et al., 2021), and make purchase decisions (Kim et al., 2016). Previous research finds that online reviews have a higher influence on potential customers' booking intentions for less familiar hotels than for familiar ones(Vermeulen & Seegers, 2009). Therefore, knowing how to provide webcare is particularly important for businesses when targeting bystanders, who are often potential new customers. Based on these insights, we aim to understand how different webcare strategies will influence actual behavior from bystanders by focusing on the bookings that hotels receive.

In order to fill this gap, we focus on frequently used webcare strategies, starting with the effect of replying or not to reviews. When a response is provided, we study how the way in which webcare is signed, the timeliness of the response, changing to another channel, tailoring the message, using a conversational tone (personalize the response, ask for more information and use non-verbal cues), asking for more information, showing gratitude, apologizing, offering compensation, and giving defensive answers influence future hotel bookings. The choice of these webcare strategies emerges from previous literature that focused on identifying the most commonly used and studied webcare strategies (e.g., Einwiller & Steilen, 2015; Sparks & Bradley, 2017; Zhang & Vásquez, 2014). For instance, previous research found that expressing gratitude is the most commonly used webcare strategy, followed by apologizing and inviting for another visit (Zhang & Vásquez, 2014). However, from these strategies, only apologizing is often studied. Therefore, this research looks not only at the effects of widely studied strategies (e.g., apologizing), but also strategies often neglected by previous research (e.g., inviting for another visit).

### *Effects of webcare strategies on future hotel bookings*

Some previous studies find that responding (vs. not responding) to online reviews harms purchase intention (Bhandari & Rodgers, 2018) and hotel performance (RevPAR) (Xie et al., 2014). Despite finding a direct negative effect of responding to negative reviews on purchase intentions, Bhandari and Rodgers (2018) also find an indirect positive effect (through an increase in brand trust) on the same variable. According to the authors, brand feedback helps to reinforce the brand’s implied promise to deliver a product of value and thereby lowers the amount of risk cues after reading a negative eWOM message (Bhandari & Rodgers, 2018). Most previous research finds a positive effect of webcare on subsequent rating (Proserpio & Zervas, 2017; Sheng et al., 2019; Wang & Chaudhry, 2018; Xie et al., 2016) and subsequent review volume (Chen et al., 2019; Proserpio & Zervas, 2017; Sheng, 2019; Xie et al., 2016), attitudes (Bhandari & Rodgers, 2018), consumer sentiment (Ma et al., 2015) and hotel revenue (Anderson & Han, 2016). Previous studies on complaint handling show that perceived justice (with how the complaint was handled and with the services and service provider as a whole) plays a role in restoring customer satisfaction (Gelbrich & Roschk, 2011). Providing webcare is, therefore, a form of restoring justice for the reviewer while also impacting the bystanders who see the review and response, as expected according to the social learning theory (Bandura & McClelland, 1977). While attitudes and sentiment are not the same as actual behavior, they are often considered as proxies or antecedents. We, therefore, expect the effect of providing webcare to be positive and propose the following hypothesis:

*H1: Providing webcare (vs. not providing webcare) positively influences hotel bookings.*

In practice, once managers decide to engage in webcare, they are faced with many choices. The first is **who should respond**? Tathagata and Amar (2018) explore how ‘webcare ownership’ (the person/team within the organization who responds) reflects the organization’s acceptance of responsibility. The authors find that webcare towards negative reviews leads to higher forgiveness (from bystanders) when it contains high ownership – webcare provided by an individual with personal details (such as name and designation) (e.g., owner, manager) – than webcare provided by a team or department (Tathagata & Amar, 2018). By seeing a name, consumers have the opportunity to attribute their blame to this identifiable person and therefore tend to forgive the firm even after a failure (Tathagata & Amar, 2018).

When signed by an individual, Xie et al. (2017) find that webcare provided by hotel executives lowers future financial performance compared to webcare provided by the staff because staff’s operational insights allow them to address consumer comments relevantly and helpfully. In contrast, Kniesel et al. (2016) finds no significant difference in bystanders’ attitudes between managers’ and staff members’ answers. We formulate the following research question:

*RQ1: What is the effect on hotel bookings of the presence of a manager’s name and function (vs. staff members or hotel name) in response to online reviews?*

An important factor for the effectiveness of webcare is timeliness: **how soon to respond?** According to a content analysis conducted by Mate et al. (2019), the majority of the customer reviews are responded to within four days of being received, which is what the authors consider to be a timely response. Previous findings seem consistent: giving a timely response, compared to a late response, positively influences reviewers’ perception of justice (Gelbrich & Roschk, 2011), increases readers’ trust, decreases their concern (Sparks et al., 2016), and leads to higher levels of readers’ forgiveness (Ghosh, 2017). Besides, a timely response increases future review volume (Sheng, 2019; Sheng et al., 2019), improves the valence of future reviews (Sheng et al., 2021), and benefits financial performance (Xie et al., 2017). Considering the consistent positive result of timeliness, including business performance, we propose the following hypothesis:

*H2: Providing timely (versus late) webcare positively influences future hotel bookings.*

Regarding the content of the response, managers are faced with the question of **which stylistic elements to use?** *Tailoring webcare*, i.e. adapting webcare to what is mentioned in the review, is one of the most commonly studied strategies (e.g., Crijns et al., 2017; Li et al., 2018). Quite some research documents more positive outcomes for tailored webcare than generic responses (e.g., Lappeman et al., 2018; Raju, 2019; Wei et al., 2013). Moreover, Min et al. (2015) find that paraphrasing a complaint (a form of tailoring) in response to a negative review causes potential guests to evaluate the response more favorably than not paraphrasing the complaint and therefore giving a more generic answer. The authors argue that rather than offer a polite but empty generic response, hotels should indicate that they have read the review thoroughly by rephrasing the review in their response (Min et al., 2015). This is in line with active listening, frequently referred to in service recovery literature, that advocates that businesses should show consumers that they are paying attention to the complaint (Drollinger et al., 2006). Tailoring the responses by referring to what is mentioned in the review and by providing explanations should, therefore, lead to positive outcomes. *Providing explanations*, another frequently used response strategy (Cenni & Goethals, 2020), can also be considered a form of tailoring since they typically refer to what is mentioned in the review. As commonly found for tailoring, previous research finds that webcare directed at negative reviews containing strong explanations can produce high consumer forgiveness (Ghosh, 2017; Tathagata & Amar, 2018). It is important to note that the only study on how tailoring affects business performance, shows a negative effect of simply repeating the topics mentioned in the review (i.e., paraphrasing as a form of tailoring) on hotel financial performance versus not repeating the topics mentioned in the review (Xie et al., 2017). This finding reinforces the idea that providing explanations is an important part of tailored webcare, as simply repeating what is mentioned in the review might lead to negative outcomes. Therefore, given the positive effects found for tailoring, we posit:

*H3: Compared to a generic response, providing tailored webcare positively influences hotel bookings.*

Another frequently studied stylistic element is the *tone of voice* of webcare. Unlike a professional tone, a conversational human tone is achieved by personalizing responses (i.e., including the reviewer's name), inviting guests to visit again, or using non-verbal cues (e.g., abbreviations, emoticons, words in upper case) (Liebrecht et al., 2021). Compared to using a professional tone, using a conversational human tone increases bystanders' trust (Stevens et al., 2018; Van Noort & Willemsen, 2012; Zhang & Vásquez, 2014) and makes them less concerned about the problem expressed in the review (Crijns et al., 2017; Sparks et al., 2016). Javornik et al. (2020) show that using a conversational human voice (or tone) leads to more positive observer perceptions of complaint handling, as opposed to when a corporate voice is employed. Only Kniesel et al. (2016) did not find that using a human tone (versus a corporate tone) leads to a more positive brand attitude from bystanders. Considering that most previous research finds positive effects of using a conversational tone, we posit:

*H4: Compared to a professional tone, using elements suggesting a conversational human tone of voice (personalizing, inviting guests for another visit and using non-verbal cues) positively influences hotel bookings.*

### *Effects of specific webcare strategies to negative reviews*

Finally, while the previous strategies apply to both positive and negative reviews, a substantial part of the webcare literature pertains to **how to respond to negative reviews in particular**? Previous studies mainly classify responses to negative reviews as either accommodative – complaisant and comprising corrective action, compensation and/or mortification – or defensive – denial and evasion of responsibility (Einwiller & Steilen, 2015). Considering the panoply of webcare strategies in these categories, especially accommodative ones, we will look into them separately, drawing from previous research on the effects of changing to a private channel, inquiring further information, expressing gratitude, apologizing and offering compensation.

*In which channel to respond?* Previous research suggests that firms should publicly contact reviewers (especially dissatisfied ones) and invite them to engage in a private conversation, changing the channel through which the conversation occurs (Grégoire et al., 2015; Zhang et al., 2019). Initial public replies show bystanders that the organization cares about the dissatisfied reviewer and is prepared to solve the issues mentioned in the review. Steering the conversation away from the public eye could help avoid the virality of negative information or online firestorms (NWOM that receives substantial support from other customers in a short period) (Zhang et al., 2019), especially at an early stage (Herhausen et al., 2019). Therefore, we predict:

*H5: Inviting reviewers to follow up on a review in a private channel positively influences hotel bookings compared to not issuing such a request.*

Previous research finds that *inquiring further information* is the most common strategy in practice, but does not lead to satisfaction with how the complaint was handled (Einwiller & Steilen, 2015). Expressing *gratitude* is the second most common strategy (Cenni & Goethals, 2020; Sparks & Bradley, 2017) and does lead to satisfaction with complaint handling (Einwiller & Steilen, 2015). Despite their frequent use, the effects of these strategies on bystanders and business performance are too scarce to formulate a directional hypothesis. Therefore, we posit:

*RQ2: What is the effect of asking for more information and expressing gratitude on hotel bookings?*

A frequently used and studied accommodative strategy is *apologizing* (e.g., Mate et al., 2019; van Hooijdonk & Liebrecht, 2021; Zhang & Vásquez, 2014). Typically, apologizing is seen as representing a caring attitude and showing compassion for the negative event experienced by the customer (Mate et al., 2019). Previous research finds that bystanders who see webcare towards NWOM containing an apology (versus no apology) have less negative behavioral intentions than those who do not (Kim et al., 2016). However, most research that studies the effects of apologizing (versus not apologizing) does not find significant effects or even finds a negative influence on consumers attitudes and intentions (Lee & Song, 2010). Dens et al. (2015) find that apologizing in itself does not significantly raise readers' attitudes or patronage intentions compared to no response, even when most reviews are positive. In the same vein, van Hooijdonk and Liebrecht (2021) find that the presence (versus absence) of an apology did not enhance brand reputation. This could be because by apologizing, firms are assuming guilt in a reviewer's accusations (Lee & Song, 2010). Therefore, we propose the following research question:

*RQ3: What is the effect using apologies (versus not using apologies) on hotel bookings?*

Sparks and Bradley (2017) note that one of the most effective responses after a negative review is to offer *compensation* (e.g., discount on a future purchase). Compensation benefits readers' purchase intention and brand perceptions (Treviño & Castaño, 2013) as well as brand reputation (Rose & Blodgett, 2016). It leads to a higher perception of justice, which positively affects customers' future behavioral intentions (Gelbrich & Roschk, 2011; Ha & Jang, 2009). It has been suggested as the optimal response for less severe failures (Liu, Jayawardhena, Dibb, et al., 2019). Offering compensation also mitigates the virality (how much a message spreads online) of NWOM when used in evolved stages of online firestorms (Herhausen et al., 2019). These effects are consistent with social exchange theory (e.g., Lamb et al., 2020) which posits that people make decisions based on a cost-benefit evaluation. In this sense, if bystanders see that unsatisfied customers are offered compensation, the "cost" of choosing a hotel that might not satisfy them diminishes, since they expect that a similar compensation would be offered to them in such case. Considering the positive results associated with compensation in prior research, we posit:

*H6: Compared to webcare that does not offer compensation, providing webcare that does offer compensation positively influences future hotel bookings.*

Previous research finds that *defensive* response strategies has a stronger impact on the consumers' perception that the company was at fault, than ‘‘no action" strategies (Dens et al., 2015; Lee & Song, 2010). However, other previous studies do not show an obvious negative effect of giving a defensive response compared to no response. Looking into defensive strategies, Weitzl and Hutzinger (2017) find that the only webcare strategy with a significant adverse effect on failure attribution is vouching (i.e., countering negative comments with favorable statements), while trivializing what is said in the review or doubting the reviewer do not affect failure attribution. Credible, defensive responses might even strengthen bystander-brand relationships (Weitzl & Hutzinger, 2019). Also, 'flyting' (a ritualized exchange of insults between two or more interlocutors) could help brands bolster their ideological positioning by engaging with opposing sides in verbal contests (Scholz & Smith, 2019). Lee and Cranage (2014) find that when there is little consensus in a negative set of reviews (meaning that some reviews are positive), defensive responses are more effective than no response to prevent a negative bystander attitude. Treviño and Castaño (2013) find that hotels performing any type of webcare, even defensive, are perceived as giving more importance to customer service and guests than hotels that do not respond to negative reviews. Therefore, given that the majority of previous research finds a positive effect of using defensiveness, we posit:

*H7: Using defensive webcare positively influences hotel bookings.*

Figure 5.1 shows the expected effects of the previously mentioned webcare strategies on the number of bookings that hotels receive.

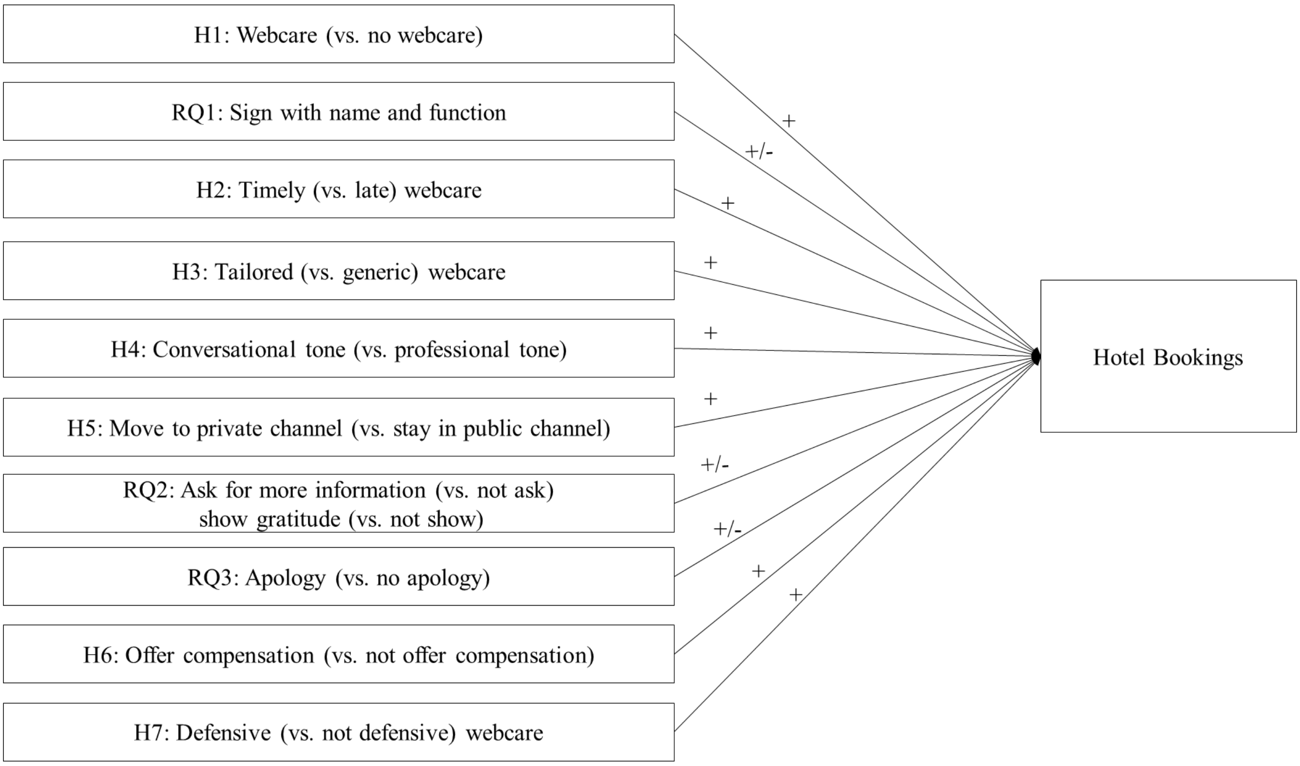



*Figure 5.1. Expected effects for the webcare variables on future hotel bookings*

## Methodology

We used a machine learning approach to study how different webcare strategies affect hotel bookings by combining online reviews and their managerial responses on Booking.com with booking data through Booking.com from seven Belgian hotels for almost four years (February 2016 until December 2019). On Booking.com, the managerial responses are displayed under each review, making this information noticeable to bystanders as prospective customers. The online reviews and managerial responses were supplied by a company that provides hotels with integrated tools and processes to manage guest satisfaction. The hotels provided us with the number of bookings they received each week through Booking.com. In sum, for each hotel, we know what reviews were posted (and when), the managerial responses (if any) provided, and the number of weekly bookings received through Booking.com. The dataset contains 18,320 reviews posted on Booking.com during the mentioned period, 5,564 of which received managerial responses. The 5,564 replies are first categorized for the purposes of this study. Considering the goal of this study – how different webcare strategies affect the number of bookings that hotels receive – we followed several methodological approaches to create the variables to be included in the model. Below, we describe them in detail.

### *Description and operationalization of webcare variables*

Several steps were involved in creating variables to measure the use of different webcare strategies. The first step was to use the Google Translate API to translate all the replies to English since many were in other languages. The second step was to develop a comprehensive document defining all webcare variables discussed in the literature review, and establish how they should be measured in the most accurate and parsimonious way. Table 5.1 summarizes these variables, including a short definition and the chosen method for each one (more details on the method are provided next). Appendix 4 contains all the details for the operationalization of each variable.

At this point, the variable timeliness was dropped from the study because, when hotels replied, they almost always did so on the same day that the review was posted.

The next step was to measure variables that could not be automatically coded (i.e., there is no dictionary). Due to the large volume of responses and our goal of developing automated approaches, we coded a training sample with human coders and then applied leading machine classifiers to determine which one worked best for classifying webcare responses. To create the training sample, we developed a codebook for our human coders with operational definitions (section 1 of Appendix 4) initially drawn from previous work and subsequently adjusted after reliability tests with two coders to establish pre-coding reliability. One researcher coded a sample of 810 managerial responses, around 15% of the total available cases (5,564). The second coder coded 15% of the instances from the coded sample, which is an acceptable subsample to calculate intercoder reliability (Riffe et al., 2019). After three training sessions, acceptable (>.67, Krippendorff's alpha [KA]) intercoder reliability levels were achieved for all variables (Krippendorff, 2018) (Table 5.2). Therefore, the sample of 810 coded managerial responses was used to evaluate several classifiers' performance in coding the remaining sample.

*Table 5.1. Description of the webcare variables*

| | **Variable** | **Definition** | **Operationalization** |
|---|---|---|---|
| | Webcare | The review received a response | Automatic (the response field is not empty) |
| Signature | Signed by department | Response signed with the department's name | BERT[11] |
| | Signed with hotel | Response signed with the name of the hotel | BERT |
| | Signed by manager | Response signed with by the manager | BERT |
| | Signed with name | Response signed with the name of the person that wrote the response | BERT |
| | Signed by staff | Response signed by the staff | BERT |
| Timeliness | Timeliness | Response given no later than the next day (following the review) | Automatic |
| Tailoring | Paraphrasing | Response mentions the topics mentioned in the review | BERT |
| | Explaining | Response explains details of a specific issue mentioned by the review (positive or negative) | BERT |

*(table continues on the next page)*

[11] More details on how BERT was the selected classifier comes later in this chapter.

*(continuation of table 5.1.)*

| | | | |
|---|---|---|---|
| Tone of voice | Personalization | Response mentions name (or alias) of the reviewer | Automatic (match between the name of the reviewer and the name mentioned in the response) |
| | Invite for visit | Hotel explicitly invites reviewer to return to the hotel | BERT |
| | Non-verbal elements | Response uses abbreviations, emoticons, words in upper case, onomatopoeia, sound stretching, or punctuation to convey message | BERT |
| Accommodative | Changing to private channel | Response directs reviewers to another channel | Dictionary approach (Herhausen et al., 2019) |
| | Ask for more information | Response requests more information from the reviewer | Automatic (use of question marks) |
| | Showing gratitude | Response uses gratitude words | Dictionary approach (developed by the authors) |
| | Apologizing | Response uses apology words | Dictionary approach (Herhausen et al., 2019) |
| | Compensating | Response offers compensation (reimbursement, voucher, discount, etc.) | Dictionary approach (Herhausen et al., 2019) |
| Defensive | Refutation | Response disagrees with/ argues against the review | BERT |

*Table 5.2. Intercoder reliability*

| **Variable** | **KA** |
|---|---|
| Signed with department | NA[12] |
| Signed with hotel | .91 |
| Signed by manager | .95 |
| Signed with name | 1 |
| Signed by staff | .8 |
| Tailoring | .78 |
| Explanation | .78 |
| Invitation for visit | .78 |
| Non Verbal cues | NA[13] |
| Defensiveness | .72 |

After creating the training sample, the next step was to evaluate the performance of leading machine classifiers, including support vector machines (SVM), boosted trees, random forests (RF), naïve Bayes (NB), and the bidirectional encoder representations from transformers (BERT) (Devlin et al., 2018). Details of our methodology are provided in Appendix 5. With 810 training cases, we used 10-fold cross-validation to compute out-of-sample accuracy (percentage of correctly classified cases) and AUC (area under a ROC curve) measures. We concluded that BERT outperforms the other methods and provide performance metrics for this classifier in Table 5.3. Accuracy and AUC were highly correlated and tell a consistent story. Most of the dimensions could be classified with near-perfect accuracy, but the machine classifiers had more difficulty classifying the more subjective dimensions of explanations, invitations, and defensiveness. In the worst cases (defensiveness and invitations), BERT still correctly classified at least 84% of the cases with AUC values greater than 90%.

Having trained the BERT text classifier, we applied it to the whole sample of 5,564 managerial responses. The next step was to aggregate the data by week to match it with the booking information provided by the hotels. Using weekly data avoids within-week seasonality issues. The data was aggregated by summing all the instances in a given week where a certain strategy was used for each hotel. In total, we have 192 weeks of data for the seven hotels, totalizing 1,344 observations. Below we explain how these variables were used in a panel-data model.

[12] Not enough cases to compute reliability. Since this variable has high face validity, we proceeded with the coding of the full sample regardless.
[13] Not enough cases to compute reliability. Since this variable has high face validity, we proceeded with the coding of the full sample regardless.

*Table 5.3. Cross-validated Accuracy and AUC scores achieved by BERT*

| Variable | Accuracy | AUC |
|---|---|---|
| Signed with department | .98 | 1 |
| Signed with hotel | .99 | .99 |
| Signed by manager | .96 | .99 |
| Signed with name | .98 | .99 |
| Signed by staff | .98 | .99 |
| Tailoring | .95 | .98 |
| Explanation | .88 | .93 |
| Invitation for visit | .85 | .90 |
| Non Verbal cues | .99 | .99 |
| Defensiveness[14] | .84 | .93 |

### *Description of the model*

This section describes the model we used to study how the webcare variables are associated with bookings during the next week. The first step was to study relationships between predictor variables to understand where multicollinearity could be problematic. Five webcare variables measure who signed the response. Some had very low occurrences (e.g., signed by the department). Principal component analysis (see detailed results in Appendix 6) shows that these variables load on three different factors. Subsequently, they were combined into three new variables, combining signed by staff or department (*sigDepStaf*), combining signed by manager or with name (*sigNameMgr*), and keeping signed by hotel (*sigHotel*) as a separate variable.

A correlation matrix (see Appendix 6) showed high correlations between some other variables ($r$ > .81), namely between tailoring and defensiveness, invite for a visit, explanation, and personalization, as well as between defensiveness and explanation, explanation and invite for a visit, and gratitude and invite for a visit. Therefore, the next step was to factor analyze the variables for potential grouping. Results (see Appendix 6) show that the tailoring (*tailor*), defensiveness (*defensive*), invite for a visit (*invitevisit*), explanation (*explain*), personalization (*personal*), and *gratitude* loaded on one factor, consistent with the results from the correlation

[14] For the variable defensiveness, we used a balanced training set (same number of cases for each label).

matrix. However, these variables present distinct constructs that do not conceptually reflect one single variable. Therefore, for the time being, they were included in the model separately. The remaining variables, non-verbal (*nonverbal*), compensation (*compensate*), channel change (*chachange*), *apology,* and ask more information (*information*) did not load on the main factor and had more modest correlations with each other.

We created a panel Poisson regression model in which the number of weekly bookings was the dependent variable (*nextbook*). We controlled for seasonality (*time*), bookings the week before (*book),* and idiosyncratic hotel characteristics by including six 0-1 dummy variables (*hotel*). This panel design is robust to many threats to internal validity. The time variable should account for outside events that could affect bookings, such as a festival or holiday. Hotel dummies account for systematic differences between the hotels. Including lagged bookings (*book*) as a predictor further controls for these threats.

Previous research shows that review volume (number of reviews) and review valence (the positivity of a review) positively influence business performance (e.g. Floyd et al., 2014). Therefore, we initially included *valence* (the average review rating from all the hotel ratings in the previous 24 months) and *volume* (the number of reviews), but found them to be highly correlated with the *hotel* variable. *Hotel* was kept in the model over *valence* and *volume* because, besides these variables, it also controls for important differences between the hotels, such as size, promotions, advertising campaigns, and even specific platform-related aspects, such as the order in which potential saw the hotels. The model also included a dummy variable for the *webcare* treatment that takes the value 0 when there were no hotel responses and 1 when the hotel responded to at least one review during a given week (*webcare*). A log transformation was performed on the variables *nextbook*, *book*, *defensive*, *invitevisit*, *nonverbal*, *apology*, *compensate*, *chachange*, *gratitude*, *information, personal*, *sigDepStaf* and *sigNameMgr* since these were right-skewed with very few observations in the tail. Logging the variables reduces the influence of outliers.

The first model we estimated included all variables. Appendix 6 reports parameter estimates and variance inflation factors (VIF). This initial model had multicollinearity issues, with VIF > 10 for the variables *tailor*, *explain*, *personal*, *invitevisit,* and *sigNameMgr.* Since *tailor* and *explain* are highly correlated with the other variables (VIFs of 24 and 11, respectively), we dropped them from the analysis. The signature variable appeared to be correlated with *sighotel,* so it was also excluded, as *sigNameMgr* is more informative for this study. Table 5.4 presents a summary of the variables included in the final model.

*Table 5.4. Variables included in the final regression model*

| **Variables in the model** | **Definition** | **Transformations** |
|---|---|---|
| *nextbook* (DV) | number of bookings that the hotel received in the following week (time $t$+1) | - |
| *hotel* | dummy variable for the hotels in study | - |
| *book* | number of bookings that the hotel received in a given week | log transformation |
| *s(time)* | seasonality | (spline) |
| *webcare* | dummy variable for the existence of responses in a given week | - |
| *sigDepStaf* | number of responses signed by the staff in a given week | log transformation |
| *sigNameMgr* | number of responses signed by the manager in a given week | log transformation |
| *chachange* | number of responses changing to a private channel in a given week | log transformation |
| *personal* | number of responses using the name of the reviewer in a given week | log transformation |
| *invitevisit* | number of responses inviting guests for another visit in a given week | log transformation |
| *nonverbal* | number of responses using non-verbal cues in a given week | log transformation |
| *information* | number of responses asking for more information in a given week | log transformation |
| *gratitude* | number of responses showing gratitude for the review in a given week | log transformation |
| *apology* | number of responses apologizing to the reviewer in a given week | log transformation |
| *compensate* | number of responses offering compensation in a given week | log transformation |
| *defensive* | number of responses using defensive arguments in a given week | log transformation |

The final model estimated the following Poisson regression using a GLM:

**nextbook ~ hotel + book + s(time) + webcare + sigDepStaf + sigNameMgr + chachange + personal + invitevisit + nonverb + info + gratitude + apology + compensate + defensive,**

where *nextbook* was the number of bookings that the hotel received in the following week (time $t$+1); the independent variables are all at time $t$; s was a univariate smoothing spline representing possible nonlinear effects of time. We used the default log link function for a Poisson model, and therefore the model predicted the log-mean number of bookings. We fitted the above model in order to test how several webcare strategies influence future bookings. Compared to the previous model that contained all the webcare variables (presented in Appendix 6), this regression model presented appropriate VIF values (less than 10).

## Results

A look into the frequency of occurrence of each strategy shows that inviting for another visit is the most commonly used strategy occurring 3,686 times in the 5,564 responses, followed by personalizing the response with the name of the reviewer ($n$ = 2,678) and by showing gratitude ($n$ = 2093). We find a similar number of instances where hotels were defensive ($n$ = 1,477) and apologetic ($n$ = 1,102). The less commonly used strategies were offering compensation ($n$ = 510), asking for more information ($n$ = 418) and the use of non-verbal cues ($n$ = 169). Regarding how many times responses were signed, we find 2,045 signatures from the department or the staff, and 5,715 signatures that include a name or come from the manager.

Table 5.5 presents the results for the final model. In order to test for the robustness of the model, we also fit a lasso regression with the same variables. Comparing the estimates of the GLM and lasso models shows that the effect sizes are robust across models, with similar variables retained in both models. This indicates the robustness of the model since in linear models there is no penalization for the model's choice of weights, which could lead to overfitting. Lasso regression, on the other hand, penalizes the model for the sum of absolute values of the weights (Hastie et al., 2009). This means that in a lasso regression variables are often dropped, and the penalty is selected with cross-validation. Considering that roughly the same results are achieved with the two models, our model can be considered robust.

*Table 5.5. Results for regression model (GLM)*

| | **Coefficients GLM** | | | | **VIF** | **Lasso** |
|---|---|---|---|---|---|---|
| | **Estimate** | **Std. Error** | **z value** | **Pr(>\|z\|)** | **GLM** | **Estimate** |
| (Intercept) | 2.13 | 0.06 | 37.17 | < 2e-16 *** | | 2.02 |
| hotel_id=2 | -0.08 | 0.01 | -7.13 | 9.86e-13 *** | 43.03 | -0.06 |
| hotel_id=3 | -1.18 | 0.03 | -37.70 | < 2e-16 *** | | -1.09 |
| hotel_id=4 | -0.84 | 0.02 | -36.38 | < 2e-16 *** | | -0.77 |
| hotel_id=5 | -0.67 | 0.02 | -34.51 | < 2e-16 *** | | -0.61 |
| hotel_id=6 | -0.01 | 0.01 | -0.63 | .53 | | 0.002 |
| hotel_id=7 | -0.58 | 0.02 | -27.50 | < 2e-16 *** | | -0.54 |
| Book | 0.61 | 0.008 | 78.70 | < 2e-16 *** | 4.82 | 0.63 |
| Time | Spline, figure available on request | | | | | |
| Webcare | 0.055 | 0.01 | 5.16 | 2.45e-07 *** | 2.57 | 0.04 |
| sigDepStaf | 0.002 | 0.004 | 0.54 | .59 | 2.45 | 0.0001 |
| sigNameMgr | 0.02 | 0.005 | 4.03 | 5.56e-05 *** | 5.89 | 0.02 |
| Chachange | 0.04 | 0.009 | 3.96 | 7.37e-05 *** | 1.36 | 0.03 |
| Personal | -0.002 | 0.009 | -0.29 | .77 | 9.28 | . |
| Invitevisit | -0.02 | 0.007 | -3.29 | .001 ** | 7.94 | -0.01 |
| Nonverbal | -0.05 | 0.01 | -3.67 | .0002 *** | 1.10 | -0.05 |
| Information | -0.04 | 0.01 | -4.32 | 1.53e-05 *** | 4.18 | -0.02 |
| Gratitude | -0.009 | 0.006 | -1.50 | .13 | 4.03 | -0.005 |
| Apology | -0.05 | 0.007 | -6.29 | 3.42e-10 *** | 3.34 | -0.04 |
| Compensate | 0.02 | 0.008 | 2.27 | .02 * | 2.11 | 0.01 |
| Defensive | 0.02 | 0.007 | 3.14 | .002 ** | 4.21 | 0.01 |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

Null deviance: 120,407 on 1,143 degrees of freedom

Residual deviance: 9,988 on 1,119 degrees of freedom

The strongest effect ($z$=+79) is for the control variable *book*: the more bookings a hotel has in one week, the more it tends to have during the following week. Not surprisingly, this variable explains a lot of the variation in bookings. This and the other controls (time and hotel dummies) account for differences across hotels and time periods, strengthening our findings' internal validity.

Responding to at least one of the online reviews a hotel receives in a given week has a positive effect ($b$=0.055, $z$=5.2, $p$<.0001) on the number of bookings received the next week, confirming H1. Because we are predicting the log expected bookings, we can say that responding to reviews is associated with an increase of $e^{0.055}$ = 1.056 (5.7%) in bookings.

Answering RQ1, managers signing with their name or function (*SigNameMgr*) positively influences future bookings ($b$=0.02, $z$=4.0, $p$<.0001), while having staff members or departments signing (*sigDepStaff*) is not significantly associated with future bookings ($p$>.05). For each unit increase in the log count of manager signatures, the number of bookings is expected to increase by 2% ($e^{0.02}$ = 1.020).

We obtain different effects for the several components of conversational tone. *Non-verbal* cues ($b$=-0.05, $z$=-3.7, $p$ <.0001) and an invitation for another visit (*invitevisit*) ($b$=-0.02, $z$=-3.3, $p$=.001) have significant negative associations with bookings, representing a decrease of 4.9% ($e^{-0.05}$ = 0.9512) and 2% ($e^{-0.02}$ = 0.9802) respectively, in the number of bookings that a hotel receives the following week. The use of personalization (*personal*) did not significantly affect future bookings ($p$>.05). Therefore, we cannot confirm H4 stating that a conversational tone positively influences hotel bookings.

Asking to change to a private channel positively influences future bookings (*chachange*, $b$=0.04, $z$=4.0, $p$ <.0001), as expected in H5, causing an increase of 4.1% ($e^{0.04}$ = 1.040) in the bookings received in the following week.

Regarding RQ2, we conclude that the effect of asking for more information (*information)* ($b$=-0.04, $z$=-4.3, $p$<.0001) is negatively associated with the number of bookings that hotels receive the following week. This means that asking for more information lowers hotel bookings by 3.9% ($e^{-0.04}$ = 0.9607). Expressing *gratitude* does not significantly affect future hotel bookings ($p$>.05). Answering RQ3, we find that the use of *apology* ($b$=-0.05, $z$=-6.3, $P$<.0001) has a negative effect on future bookings. Apologizing lowers future hotel bookings by 4.9% ($e^{-0.05}$ =

0.9512). As predicted in H6, offering *compensation* increases future bookings ($b$=0.02, $z$=2.3, $p$=.04) by 2% ($e^{0.02}$ = 1.020). Finally, as predicted in H7, being *defensive* positively influences future bookings ($b$=0.02, $z$=3.1, $p$=.01), increasing the number of bookings the following week by 2% ($e^{0.02}$ = 1.020). Considering that we had to drop some of the variables from the analysis, we could not test H2 (timeliness) and H3 (tailoring).

## Discussion

This study investigates effects of several commonly found and frequently studied webcare strategies on future hotel bookings. We find that responding improves the number of future hotel bookings, in line with most previous studies on webcare (e.g., Chen et al., 2019). This result contradicts some previous research that finds negative associations of responding and purchase intention or business performance (Bhandari & Rodgers, 2018; Xie et al., 2014). Our finding of a positive effect of responding might be explained by the fact that our model includes control variables for possible influences of external factors (i.e., seasonality and hotel effects) and several webcare strategies. This addition is relevant considering that previous research points out that the effect of providing or not providing webcare might depend on the actual content of the response (Van Noort et al., 2015).

Looking into specific webcare strategies, we find that the one with the most positive effect on bookings is changing to a private conversation channel by inviting the reviewer to further discuss the situation via email or phone. Through this strategy, hotels show bystanders that they care about dissatisfied reviewers, but can avoid firestorms (e.g., Herhausen et al., 2019), particularly on platforms that allow a continuous dialogue between reviewers and organizations (e.g., Twitter). The positive association between changing to a private channel and future hotel bookings shows that bystanders looking to book a hotel search for indications that potential negative events will be addressed. Inviting the reviewer to engage in a private conversation signals that hotels are interested in discussing the situation in person. When managers want to give a public response, the best strategy is to offer compensation, as also found previously (e.g., Rose & Blodgett, 2016). Offering compensation is seen as a way to restore justice (Liu, Jayawardhena, Dibb, et al., 2019) and, therefore, positively influences bystanders considering booking a particular hotel. In line with the social exchange theory (Lamb et al., 2020), the cost of choosing a hotel that might not satisfy them diminishes, since they expect to be offered similar compensation in such case.

Being defensive also yields positive business results, as expected from previous research that finds that not being complacent can be good for firms (Johnen & Schnittka, 2019; Scholz & Smith, 2019). This result seems to defy the established idea that accommodative webcare, exclusively, is the only type that can result in positive outcomes (e.g., Casado-Díaz et al., 2020). Together with other studies that defensive strategies can be the best ones to adopt (e.g., Xia, 2013), our findings point at a new perspective on webcare on positive effects of the use of defensiveness. This is consistent with previous findings that indicate that, by being accommodative, firms accept the blame for what went wrong in the review, with adverse downstream effects (Johnen & Schnittka, 2019). While review writers naturally tend to make internal attributions and blame the service firms involved (Cowley, 2005), review readers act as third parties. They also take other cues into account when making their judgments (He & Bond, 2015), such as their impression based on the hotel pictures, reviewer characteristics, other situational factors, and also webcare. Besides, as Ro and Wong (2012) point out, there are occasions when customers knowingly and incorrectly report service failures or make illegitimate complaints. In these cases, companies should be able to refute these reviews to overcome dishonest complaints. Findings from previous research (Dens et al., 2015; Lee & Cranage, 2014) show that the degree of consensus in a review set might also determine if accommodative or defensive strategies are preferable. If a hotel receives many negative reviews, readers are more likely to infer that the hotel does indeed possess flaws, and defensive responses will not seem credible. Most of the reviews posted online are actually positive (Chevalier & Mayzlin, 2006; Resnick & Zeckhauser, 2002). Bystanders are thus confronted with predominantly consistent positive review sets. Being defensive to one or a few negative reviews can, in such case, lead to positive results.

Another strategy that yields positive results is having the manager sign the response with their title or name. This strategy results in more bookings, confirming earlier findings (Tathagata & Amar, 2018). Such responses signal that the manager (as an individual) takes responsibility for the situation, as a result of which it is easier to forgive the company for the failure (Tathagata & Amar, 2018). Besides, when managers sign, this may indicate that the hotel takes the reviews seriously by having someone with higher responsibility (than staff members) providing webcare, making the actual guests feel valued both when they provide positive or negative feedback, and bystanders feel that any concerns they may have will also be addressed in the same manner.

Our findings show that certain webcare strategies should be avoided, since they hurt future bookings. Apologizing signals review readers that the firm assumes guilt for reviewers' accusations (Lee & Song, 2010; Weiner, 2010). This negative effect of apologizing, together with the positive effect found for defensiveness, can be explained by attribution theory (e.g., Chang et al., 2015; Weitzl et al., 2018): review readers might interpret a hotel that does not counter-argue a negative review as taking the blame for the complaints. Unlike compensation, which increases the perceptions of justice (Liu, Jayawardhena, Dibb, et al., 2019) and minimizes the cost of choosing a bad hotel (Lamb et al., 2020), offering an apology to dissatisfied reviewers does not reassure bystanders. This finding, again, contradicts the idea that accommodative strategies, such as apologizing, are always linked with more positive outcomes. Asking for more information in a platform such as Booking.com, which does not allow reviewers to reply back to the webcare message, also reduces future hotel bookings. It is preferable to invite the reviewer to a private channel, since by publicly asking for more information, the hotel is undermining the feeling of closure that bystanders might have by seeing a negative complaint being solved.

Inviting customers for another visit and using non-verbal cues, two elements that determine the tone of voice used in the webcare message, exert adverse effects on future hotel bookings. These findings, together with the non-significant effect found for personalizing the response with the name of the reviewer, seem to point to a possible negative influence of using a conversational tone over a professional tone. Inviting the reviewer for another visit might be interpreted as a strategy to sell, undermining the fact that hotels respond to reviews because they care about their costumers and making it seem more like a sales attempt. Moreover, inviting guests to visit again is such a frequently used webcare move (Zhang & Vásquez, 2014) that bystanders might see it as meaningless, leading to a negative association with its use. The negative effect of non-verbal cues might be explained by previous research that finds that non-verbal cues used in marketer-consumer interactions, such as smiles, convey friendliness and increase the perceived warmth but decrease perceived competence (Wang et al., 2017).

Finally, besides personalizing the response with the name of the reviewer, expressing gratitude and having staff members signing the webcare also do not have a significant effect on future bookings. Similarly to what happens with inviting for another visit, thanking the reviewer for their comment does not foster positive outcomes considering how commonly it is used (Zhang & Vásquez, 2014). Having staff members signing the messages does not lead to more bookings, as it happens when managers sign, showing that ownership of top executives is indeed important in how bystanders have their booking decisions influenced by webcare (Tathagata & Amar, 2018).

Regarding the comparison of the different classifiers to create the webcare variables, we find that BERT outperforms the other tested machine learning algorithms, consistent with previous studies using text classifiers for similar tasks (e.g.,González-Carvajal & Garrido-Merchán, 2020). The results from SVM, boosted trees, random forests, and naïve Bayes are never substantially discrepant, contrary to what is found in previous literature (Hartmann et al., 2019).

## Theoretical and methodological implications

There are several theoretical and methodological contributions from this study. First, it adds to previous research in advancing the theoretical framework for webcare by simultaneously assessing the effects of several commonly used webcare strategies (managerial responses to online reviews) on actual business outcomes (bookings). It disentangles contradictory findings (for instance, for the effects of having people with different roles in the organization replying to the reviews) and focuses on under-researched strategies (for example, the effect of asking for more information).

Second, the findings from this study help demystify the idea that accommodative webcare, exclusively, fosters positive brand and business outcomes. As previously mentioned, defensive webcare can, in many circumstances, be the preferred approach as it contributes to lower the attribution of guilt for what is mentioned in the review to the hotel (Johnen & Schnittka, 2019).

Third, this paper has methodological contributions by testing several machine learning classifiers to identify webcare strategies. Often, researchers in the field of tourism in general, and webcare studies in particular, have to resort to machine learning or big data approaches to fulfill the goals of their studies (Line et al., 2020). However, the panoply of classifiers and technical options to deal with a large volume of data can seem overwhelming without proper guidance and benchmarks. Our study shows that BERT performs better than the other classifiers to label this type of online data. Besides, by documenting the steps taken to create each variable, we are laying the foundations for future research that faces similar challenges.

## Managerial implications

In terms of managerial contributions, our findings help hotel managers optimize their webcare strategy for better business results. Our findings show that the webcare strategies used by hotels matter above and beyond the other components that have shown to be related to hotel business performance (specific hotel characteristics, previous volume of bookings, seasonality, etc.). Providing webcare increases futures bookings, so managers should make it a priority. More specifically, hotels should embrace the strategies that show a positive effect on future bookings: directing reviewers to a private channel, being defensive, offering compensation, and having managers signing the response. Webcare strategies that should be avoided are apologies, asking for more information, inviting customers for another visit, and non-verbal cues, since our findings show that they lower future hotel bookings. Strategies that do not appear to have any effect on future bookings are expressing gratitude, personalizing, and having staff members signing webcare; therefore, webcare efforts should not be directed at employing those.

Our findings seem to indicate that in the event of a negative online review, there are two possible approaches that will have a positive impact on business results. When the complaint is legitimate (or might seem legitimate to potential customers), hotels should offer compensation and/or direct the complainer to a private channel. Considering that compensations can be costly for businesses (Liu, Jayawardhena, Dibb, et al., 2019), managers that cannot offer them consistently should consider to first direct the reviewer to another channel and then offer compensation privately. This avoids that future guests are disappointed if they do not get a similar offer, which they would expect, in light of the social exchange theory (e.g., Lamb et al., 2020). However, when there is ground for the premises mentioned in the reviews to be refuted, managers should defend their perspective and counter-argue what is mentioned in the review. These insights can also help develop automated webcare, for instance, chatbots (Li et al., 2021; Liebrecht & van Hooijdonk, 2019). Since there is a crescent volume of reviews online, it becomes challenging for businesses to find the resources to manage all online feedback. Besides, our contribution to the development of tools for automated webcare is enriched by showing how to identify and define webcare strategies. Testing different machine learning classifiers to label the strategies also contributes to the automation of webcare by finding which approach is the most adequate to analyze raw review and webcare data.

Our findings are also useful for hotel review platforms. Based on which strategies have shown to lead to positive or negative results, review platforms can develop systems similar to those employed for how the reviews are displayed (i.e., organized according to the rated review helpfulness). For instance, it can both help businesses and bystanders looking for information if the platforms prioritize reviews with responses that lead to higher bookings. This would be rewarding for hotels that reply adequately.

## Limitations and further research

A relevant area for further research is to explore which webcare strategies work best for which types of reviews, i.e., look into how review characteristics moderate the effects of different webcare strategies. For instance, Allard et al. (2020) find that consumers are empathetic towards firms that receive unfair reviews, showing that review readers are also affected by the arguments used in the reviews. In fact, besides the well-established effect of review valence and volume (e.g., De Pelsmacker, Dens, et al., 2018; Kim et al., 2018; Maslowska, Malthouse, & Bernritter, 2017), previous research also shows that the importance of the arguments determines how online reviews influence review readers (e.g., Filieri, Hofacker, et al., 2018; Thomas et al., 2019).

Future studies should also look into the effects of combining multiple strategies on business performance. Previous research points out that using several strategies in one response might yield the most positive results on consumers (Dens et al., 2015; van Hooijdonk & Liebrecht, 2021) and, consequently, affect future business. Considering that, in practice, most managerial responses use more than one strategy, it is interesting for future research to look into different combinations of strategies and how they influence future business. Besides, more research is needed to understand how tone of voice influences business performance. While conversational versus professional are the more commonly studied tones, other formats can be important for the tone used on webcare. For instance, previous research finds that consumers appreciate interactions containing humor or comedy attempts from brands (Warren et al., 2018). Specially regarding responding to positive reviews, humor can possibly be a webcare strategy leading to positive outcomes.

A limitation of our study is that we were not able to assess the effect of two frequently used and commonly studied strategies. Therefore, other studies should focus on the effects of timeliness and tailoring on business performance since these strategies could not be tested in this study. Finally, our study should be replicate in the context of other tourism industries to test the generalizability of our findings.

# 6. Conclusions, implications and further research

This concluding chapter first describes a summary of the main findings reported in the previous chapters of the dissertation. Next, it discusses the contributions and implications of the present work to both theory and practice. Finally, it discusses its limitations and elaborates on suggestions for future research.

## Main findings

The goal of this dissertation was twofold: to study (1) the influence of different eWOM characteristics on consumers' responses; and (2) the effects of webcare on consumer's responses and business performance. Next, we will present the main findings for each of these two parts.

Considering the first part, we aimed to **find out which online review cues are important to determine the credibility and helpfulness of a review and how the composition and content of a set of reviews affect consumer responses.** Chapter 2 answers the following question:

*What is the relative importance of argument strength, sidedness of the message, writing quality, number of arguments, number of reviews, rated review usefulness, and summary review rating, for perceived review credibility and usefulness?*

The results show that argument strength is the most important cue that review readers use to assess the perceived usefulness and credibility of a review, in line with findings from previous research (Thomas et al., 2019; Wang & Li, 2019). The presence or absence of a star rating and the number of reviews are the two least important cues for both review usefulness and credibility. This reinforces previous findings that summary review star rating does not affect review impression when review text is provided (e.g., De Pelsmacker, Dens, et al., 2018). In general, more arguments, stronger arguments, good writing quality and higher rated review usefulness all positively affect both review credibility and usefulness, as previously found (e.g., Kim et al., 2018; Schindler & Bickart, 2012). One-sided messages are both more useful and more credible than two-sided messages, which is in line with the confirmation bias in information processing (Metzger et al., 2020). Fewer reviews (as opposed to more) and the presence of a (positive) star rating cause a review to be perceived as more credible, while having more reviews and not presenting a star rating is better for perceived usefulness. In the same chapter, we also answer the following question:

*What is the relative importance of the writing quality, number of reviews, rated review usefulness and summary review rating, number of arguments, argument strength and sidedness for highly involved individuals and lowly involved ones when evaluating review usefulness and credibility?*

When assessing review usefulness, the most important cues for lowly involved individuals are argument strength and writing quality, while for highly involved individuals, the most important cues are argument strength and the number of arguments. In terms of review credibility, the most important cues for both lowly and highly involved individuals are argument strength and writing quality.

Chapter 3 builds upon the previous chapter by studying cues related to the review arguments, which we found to be the most important review element. This chapter goes deeper into the role of arguments' content and number by looking into how varying reviews' argument importance, number, and repetition influence consumers' intentions. With this objective in mind, the chapter presents an experimental study aiming to find out the following:

*How do varying ratios of a majority of negative reviews about less important attributes and a minority of positive reviews about more important attributes influence consumers staying intention at a hotel, and where is the 'tipping point' at which a number of positive reviews in a predominantly negative review set leads to a positive hotel booking intention?*

Additionally, we look into the effects of message repetition by studying the following:

*How does having (multiple) positive reviews about the same attribute versus positive reviews about different attributes moderate the ratio of positive reviews on readers' intention to stay at a hotel?*

The results show that a more positive review set leads to a higher staying intention only when the positive reviews discuss different attributes (and do not repeat the same attribute). When a review set only presents a single positive reason to stay at a hotel, it does not compensate for the multiple reasons to avoid the hotel. Contrary to what is expected on the basis of the truth effect (Dechêne et al., 2010), repetition may not necessarily increase truth perceptions in the context of online reviews. These findings support the repetition-variation hypothesis, where more arguments enhance persuasion (Calder et al., 1974; Petty & Cacioppo, 1984; Willemsen et al., 2011b). The 'tipping point' at which positive reviews compensate negative ones is four positive reviews about different attributes in a set of 12. Our results point at a nuance to the bandwagon effect and the negativity bias (Carstensen & DeLiema, 2018; Rozin & Royzman, 2001; Wu, 2013). Contrary to what is expected in light of the bandwagon effect, people do not necessarily follow the ‘majority’ opinion, and negative reviews do not always carry more weight than positive reviews, as posited by the negativity bias.

Having studied how several review characteristics influence consumer responses in the first part of this thesis, the second part of the manuscript relates to the effects of webcare on consumers and, by consequence, on business performance. As explained in chapters 2 and 3, understanding eWOM is very important for academics and practitioners since it influences many outcomes that are determinants of business success. Considering its importance, it is crucial to also look into how should eWOM be managed. Therefore, the next two chapters of the dissertation aim to **build up a framework for webcare and explore the effects of different webcare strategies on business performance**. Chapter 4 reports the findings from a literature review that answers the question:

*What are the webcare strategies that yield positive (and negative) results for businesses?*

Most studies suggest a positive effect of responding to eWOM, especially to negative WOM (NWOM). However, replying to positive WOM might not be advisable. In fact, there is little added value in answering all reviews (Anderson & Han, 2016; Homburg et al., 2015). Also, webcare (especially towards negative reviews) should have a high level of ownership (i.e., signed with the name of the person responding) (Tathagata & Amar, 2018), and there is no advantage in having managers reply to reviews (e.g., Xie et al., 2017). The literature is consistent finding that webcare given within a short time frame leads to the most favorable outcomes (e.g., Xie et al., 2017). Firms should publicly contact complainants and invite them to engage in a private conversation when the online firestorm (NWOM that receives substantial support from other customers in a short period) is in an early stage (Grégoire et al., 2015; Herhausen et al., 2019). Most research emphasizes a positive effect of tailoring webcare (i.e., adapting the response to individual reviews) to NWOM (e.g., Lappeman et al., 2018). In terms of the tone of the response, studies (e.g., Sparks et al., 2016) show that webcare towards NWOM should use a conversational human tone (as opposed to a professional one), be personalized and show empathy. When responding to NWOM, accommodative webcare leads to a more positive outcome than defensive webcare, especially when several accommodative strategies are combined (Dens et al., 2015). However, the use of defensive strategies (e.g., refuting) can be preferable in some circumstances (e.g., Li et al., 2018).

To propose an agenda for future research on webcare, in chapter 4, we also answer the question:

*Which webcare strategies show inconsistent outcomes and are under-researched requiring further studies?*

Areas that need further research are the effects of providing webcare on future review volume and the 'optimal' ratio of reviews for which to provide webcare. Besides, future research should also look into exploring the effects of webcare on business performance, rather than merely on consumer responses, since the effects of webcare on business performance are crucial but under-researched. There are also inconsistent findings in terms of the benefits of engaging in webcare in different channels or platforms. More research is also needed to get insights into the best way to tailor webcare for PWOM and NWOM, as previous findings are also contradictory. Another inconsistency is the superiority of accommodative strategies over defensive ones. Previous research indicates that, although accommodative responses can also lead to positive outcomes, there are circumstances in which being defensive appears to be preferable (e.g., Li et al., 2018) and can even help cement the brand's personality (Johnen & Schnittka, 2019). However, more research is needed to make clear when defensiveness and accommodativeness are appropriate. Other topics that need further research are appropriate strategies to respond to positive WOM and the effects of combining different webcare strategies in the same response. Also, very little research is devoted to investigating how webcare is understood across cultures and languages, which is fundamental considering previous findings that cultural differences affect how eWOM is produced and perceived (e.g., Fong & Burton, 2008). Although a common practice adopted by firms, there is no research on the effects of hiring external professionals to provide webcare on business outcomes. As mentioned previously, timely webcare yields positive outcomes; however, more research is needed to understand what is considered to be a timely response and to what extent a late response produces negative effects. Finally, we propose that future research addresses the mechanisms (mediators and moderators) that explain and qualify the effects of the different webcare strategies on business outcomes.

Finally, and considering the many inconsistencies and under researched topics uncovered in chapter 4, chapter 5 builds upon past research on the effect of these strategies on attitudes and intentions by investigating how webcare in general and specific webcare strategies in particular influence (future) hotel bookings. We thus aim to answer the following research questions:

*How does webcare affect hotel bookings?*

*How do specific webcare strategies affect hotel bookings?*

In this chapter, we confirm some of the conclusions of previous research discussed in chapter 4, and extend them by testing how webcare strategies influence actual hotel bookings. We find that responding improves the number of future hotel bookings, in line with most previous studies on webcare (e.g., Chen et al., 2019). In terms of specific webcare strategies, the results show that the one with the most positive effect on bookings is changing the conversation to a private conversation channel by inviting the reviewer to further discuss the situation via email or phone. Through this strategy, hotels show bystanders that they care about dissatisfied reviewers and can avoid firestorms (e.g., Herhausen et al., 2019). When managers want to give a public response, the best strategy is to offer compensation, as also found previously (e.g., Rose & Blodgett, 2016). Being defensive also yields positive business results, which is in line with findings from previous research that found that not being complacent can be good for firms (Johnen & Schnittka, 2019; Scholz & Smith, 2019). This result seems to defy the established idea that accommodative webcare, exclusively, is the only type that can result in positive outcomes (e.g., Casado-Díaz et al., 2020).

Whilst on chapter 4 we posited, based on previous literature, that having staff members providing webcare was the suggested strategy to follow, we found in our study that having the manager sign the response with their title and name is the strategy that leads to the most positive business outcomes. This strategy results in more bookings since it signals that the manager (representing the hotel) takes responsibility for the situation (Tathagata & Amar, 2018). Having staff members sign the webcare does not have a significant effect on future bookings.

Our findings also show that certain webcare strategies should be avoided since they hurt future bookings. Apologizing signals review readers that the firm assumes guilt for reviewers' accusations (Lee & Song, 2010; Weiner, 2010). Unlike compensation, which increases the perceptions of justice (Liu, Jayawardhena, Dibb, et al., 2019) and minimizes the cost of choosing a bad hotel (Lamb et al., 2020), offering an apology to dissatisfied reviewers does not reassure bystanders. Asking for more information also reduces future hotel bookings. By publicly asking for more information, the hotel is undermining the feeling of closure that bystanders might have by seeing a negative complaint being solved.

The results of the study presented in chapter 5 also seem to contradict previous research in terms of how using a conversational tone affects consumers' responses. As discussed in chapter 4, we expected that a more conversational (and friendly) tone would lead to more positive outcomes than using a more professional (and distant) tone. However, inviting customers for another visit and using non-verbal cues (e.g., emoji), two elements that determine the tone of voice used in the webcare message, exert adverse effects on future hotel bookings. These findings, together with a non-significant effect found for personalizing the response with the name of the reviewer, seem to point to a possible negative influence of using a conversational tone over a professional tone. The reason for this discrepancy between what we found in our study and what emerges from previous research could be related to the context in which previous studies achieved their findings since most studies rely on papers on services research and complaint handling that, for the most part, consider offline interactions (e.g., Davidow, 2000; Davidow, 2003) or interactions with customers that do not occur in a public sphere being, therefore, more focused on pleasing the dissatisfied customer than on the opinion of the 'crowd'. As mentioned in chapter 5, the impact of webcare on bystanders exerts considerable influence on business, specifically hotel bookings in the case of our study, and these review and webcare readers seem to prefer a less conversational tone, conveying professionalism.

Finally, in this study, we found that expressing gratitude does not have a significant effect on future bookings. Since this is such a commonly used strategy, thanking the reviewer for their comment does not foster positive outcomes. Bystanders might see this strategy as meaningless.

## Contributions to theory

The findings of this dissertation extend prior research in the field of eWOM, online reviews and webcare by studying how review characteristics and managerial responses affect consumer responses and business performance.

The first main contribution of this dissertation is that it looks into the relative importance of review cues for perceived usefulness and credibility that each have been studied separately but never simultaneously. As such, the study in chapter 2 offers a more comprehensive analysis than previous research. Looking at the cues whose role was not clear in previous literature, we found that consumers perceive one-sided reviews as more useful and credible than two-sided reviews, which sheds new light on the role of sidedness in online reviews. The fact that review volume contributes positively to usefulness but negatively to credibility is also an important contribution since consumers could perceive a higher number of online reviews as less credible because they suspect that they may be getting fake reviews. Contrary to previous research (Kolomiiets et al., 2016), we find that peripheral cues (for instance, 'writing quality') are more important for low involvement individuals than for highly involved ones. This finding confirms what could be expected based on the ELM (Richard E. Petty & John T. Cacioppo, 1986). Another contribution of this dissertation is that it helps to understand the role of the number of arguments in a review. Previous research showed contradictory findings for this review characteristic but, according to our results, the number of arguments appears to be processed centrally, being more important for highly involved individuals than for lowly involved ones. Our findings also contribute to how further studies should look into these effects and explore them further.

This dissertation also provides insights into the relevance and the applicability of well-established theories, such as the bandwagon effect, negativity bias, the truth effect, and the repetition-variation hypothesis. First, we challenge the bandwagon effect, by showing that consumers do not always follow the majority opinion (Sundar et al., 2008). In chapter 3, we found that review readers show a positive staying intention at a hotel even when they read a predominantly negative set of online reviews. Second, contrary to what is commonly theorized in eWOM studies, the findings from chapter 3 show that the negativity bias (Rozin & Royzman, 2001; Wu, 2013) does not completely explain how review readers assess a mix of positive and negative information. They take other review characteristics into account, such as the importance of the arguments used in the reviews. Third, the positive effect of providing different arguments (versus repeating arguments) in a set of reviews indicates that information richness and completeness are also important determinants of consumers' responses. This finding supports the repetition-variation hypothesis (Cacioppo & Petty, 1989) rather than the truth effect of repetition (Dechêne et al., 2010) and strengthens the importance of having greater information diagnosticity of reviews (Herr et al., 1991). Indeed, our results indicate that the truth effect does not apply in the context of online reviews, as review readers prefer more diversified information than the same argument in all positive reviews. By exploring the applicability of these well-established psychological mechanisms in the context of online reviews, we shed light on how future studies should consider them when theorizing on the response to online review information.

In terms of its contribution to the knowledge on webcare and how it affects reviewers and bystanders, this dissertation offers a broad look into a vast panoply of webcare strategies, identifies research gaps, and proposes areas for further study. Chapter 4 indeed develops a comprehensive overview of the vast previous literature on webcare strategies that is published in dozens of journals, and is frequently contradictory and inconsistent. It offers a framework that contributes to future research by providing an overview of all research on this topic, discussing possible explanations for inconsistencies, and identifying under-researched topics that need to be further explored. Chapter 5 adds to previous research in advancing the theoretical framework for webcare developed in chapter 4 by simultaneously assessing the effects of several commonly used webcare strategies (managerial responses to online reviews) on actual business outcomes (hotel bookings). Based on the social learning theory (Bandura & McClelland, 1977) and on the effects that previous research found for how webcare affects bystanders, we show that business performance is affected by how review readers perceive different webcare strategies. This chapter disentangles contradictory findings (for instance, for the effects of having people with different roles in the organization replying to the reviews) and focuses on under-researched strategies (for example, the effect of asking for more information or showing gratitude). Second, the findings presented in chapter 5 help demystify the idea that accommodative webcare, exclusively, fosters positive brand and business outcomes.

This dissertation also has important methodological contributions. In chapter 2 we use a conjoint analysis, allowing the simultaneous comparison of several review cues to find their the relative importance in the presence or absence of each other. This is a relatively uncommon setting in studies in eWOM, that typically use experiments and are limited in the number of constructs that can be included in their design in order to be feasible. Besides, this method simulates a realistic context where consumers make (implicit or explicit) trade-off between cues in a review.

Chapter 5 also presents relevant methodological advancements for the studies on webcare. Most previous webcare studies are based on experimental designs with limited external validity and/or explore the effect of webcare on (future) hotel ratings or booking intention and/or are limited to one hotel only. Chapter 5, instead, uses a machine learning approach to explore the effect of specific webcare strategies on the actual bookings of seven hotels while controlling for other factors that can affect business in the hospitality industry, such as bookings' seasonality. In chapter 5, we develop automated machine learning tools and methods for coding webcare responses. Several machine learning text classifiers (support vector machines, boosted trees, random forests, naïve Bayes and BERT [bidirectional encoder representations from transformers]) are thoroughly tested for the different variables in the study and their performance evaluated. We find that BERT outperforms the other tested machine learning algorithms, consistent with previous studies using text classifiers for similar tasks (e.g.,González-Carvajal & Garrido-Merchán, 2020). This study, therefore, helps to develop automated webcare, for instance, chatbots, which are fundamental tools in the future of webcare management (Li et al., 2021; Liebrecht & van Hooijdonk, 2019; Navío-Marco et al., 2018).

## Contributions to practice

This dissertation also offers several contributions for practitioners. It helps them understand better how people process eWOM and how it should be managed to dissipate negative outcomes of NWOM and boost positive brand outcomes. This advice is extremely relevant to managers and marketers, as the volume of eWOM, and online reviews in particular, keeps increasing over the years and managing it becomes a challenge, especially because eWOM has a strong influence on consumer behavior.

Specifically, this dissertation provides insights for marketers by showing which review cues are important to evaluate perceived usefulness and credibility, two important 'gatekeepers' to the further decision-making process. In light of the findings from chapter 2, reviewers should be requested to write about their experiences rather than merely provide a star rating, preferably mentioning several strong arguments in a well-written way. Managers could incentivize strong arguments by rewarding reviews with higher-rated usefulness or by suggesting important attributes or aspects that the review could mention. It might be interesting for practitioners to consider different website layouts depending on consumers' involvement considering the differences between high and low involvement individuals found in chapter 2. Personalization of web layout and content is increasingly feasible through artificial intelligence, and previous research showed that it influences how reviews are perceived (Aerts et al., 2017). Involvement could be deduced from, for example, previous searches for products in the same category or likes or interests on social media. For more highly involved individuals, the central arguments of a review should be easily accessible and could be highlighted using bold font. It would be more useful for more lowly involved individuals to see an overall assessment of the review, such as the rated review usefulness.

Chapter 3 of this dissertation also offers practical contributions by showing that businesses can be positively evaluated even when most reviews are negative. Although receiving little to no negative reviews about your product or service is often considered an ideal scenario, our research shows that brand managers should not necessarily fear negative reviews, as long as the positive reviews are diverse and about important product or service characteristics. Online review managers should thus encourage diversity in online reviews, for instance, by asking reviewers to comment on aspects neglected in previous reviews. Besides, practitioners should try to diversify the arguments of the reviews picked to be displayed on the brand's website.

The framework proposed in chapter 4 provides guidance to managers seeking to respond to eWOM as organizations must know whether and how to invest their efforts in webcare. Although not many consistencies can be drawn from previous research, we provide guidance for practitioners based on previous research on which webcare strategies lead to positive outcomes, which ones to avoid, in which contexts. In summary, responding is advisable, especially if the reviews are negative, and this should be done promptly. The webcare should be tailored and, in most circumstances, accommodative. Our results are crucial to develop automated response mechanisms, such as chatbots (Dao & Theotokis, 2021; Li et al., 2021; Liebrecht & van Hooijdonk, 2019). While companies with smaller review volumes might manually tackle online reviews, others might need to apply advanced artificial intelligence algorithms to face the volume of reviews they receive. In either case, knowing the best way to respond to each review is crucial to overcome the nefarious effects of NWOM and boost the positive impact of PWOM.

In terms of managerial contributions, our findings from chapter 5 help hotel managers optimize their webcare strategy for better business results. Our findings show that the webcare strategies used by hotels matter above and beyond the other components that have shown to be related to hotel business performance (specific hotel characteristics, previous volume of bookings, seasonality, etc.). More specifically, hotels should embrace the strategies that show a positive effect on future bookings: directing reviewers to a private channel, being defensive, offering compensation, and having managers signing the response. Webcare strategies that should be avoided are apologies, asking for more information, inviting customers for another visit, and non-verbal cues, since our findings show that they lower future hotel bookings. Strategies that do not appear to have any effect on future bookings are expressing gratitude, personalizing, and having staff members signing webcare; therefore, webcare efforts should not be directed at employing those.

Our thesis also offers practical contributions to review platforms. According to our findings from chapter 2 and 3 that diversity of arguments leads to positive attitudes and intentions, review platforms could provide people with a template or criteria for reviewers to comment on or rate products and services to promote diversity of arguments. Besides, considering the relative importance of rated review usefulness found in chapter 2, reviews could be sorted based on helpfulness votes. Some practices, such as explicitly asking reviewers to write both positive and negative arguments, could be impairing the review credibility since we found in chapter 2 that review readers may prefer one-sided messages. As mentioned before as a recommendation for managers, review platforms could also implement mechanisms that reward reviews rated as more useful or by suggesting the reviewer to mention in important attributes or aspects in the review. Writing quality, an important review cue, could be ensured by providing automatic grammar and spelling controls. A few peripheral cues, such as the number of reviews and the presence of a star rating, are relatively unimportant and should thus not necessarily be included in the case of single short reviews. Both these elements could become more useful, though, when people are exposed to a larger set of reviews and/or longer reviews.

The findings from chapter 3 that multiple attributes are more influential than always repeating the same attribute also provide insights to review platforms. Review platforms could use artificial intelligence to generate a list of attributes that are not being mentioned commonly or recently and present the reviewer with this list as a suggestion of aspects to comment upon.

Finally, based on which strategies have shown to lead to positive or negative results in chapters 4 and 5, review platforms can develop systems similar to those employed for how the reviews are displayed (i.e., organized according to the rated review helpfulness). For instance, it can both help businesses and bystanders looking for information if the platforms prioritize reviews with responses that lead to higher bookings. This would be rewarding for hotels that reply adequately.

## Limitations and suggestions for further research

The limitations of the present dissertation have already been discussed in each of the empirical chapters. Although more specific issues will not be reiterated here, a discussion of the more important limitations and ways to address them in future research follows next.

The first limitation has to do with the method used in the study reported in chapter 2. In conjoint analysis, the relative importance of attributes is determined by the selection of attributes and their levels. Therefore, despite the extensive list of cues included in the study, our choices were determinant for the outcome. Further research should test the relative importance of cues (attributes) in both the presence or absence of other cues, since this can have an impact on the relative importance of cues. Besides, we used an orthogonal design that does not account for interaction effects; further studies may look at how the different review attributes interact. For instance, it might be possible that (bad) writing quality impairs perceived argument strength, as the competence of the reviewer is put into question (Clare et al., 2018). Furthermore, other product categories and types should be tested and compared to test the robustness of our findings or differences between product and service types. Future research could also incorporate other cues not related to review content, such as sender characteristics (expertise, and homophily, credibility).

There is also an influence of our choices on the study's outcome in the experiment mentioned in chapter 3. Further studies should opt for other scenarios and contexts (e.g., for other services or products), contributing to the findings' generalizability or adding other variables such as personality traits of the respondents or other review cues such as star or usefulness ratings. Previous studies suggest that multiple sources presenting multiple arguments enhances persuasion over having either (multiple sources or multiple arguments) alone (Harkins & Petty, 1981). Future research could disentangle the effects of these two aspects of ‘variability’ in reviews. More research should be devoted to understanding the effect of source credibility on how review readers interpret negative information. An information adoption mechanism where credibility moderates the effects of message repetition on intentions (e.g., Ernst et al., 2017) can also be applied to online reviews and should, therefore, be further studied. Future research could also focus on how the tipping point we found (4 out of 12 versus 3 out of 12 positive reviews) evolves in larger or smaller review sets and study other valence ratios. Our research could also be expanded to positively valenced review sets, as well as look at the effect of attribute repetition in negative reviews. Exploring varying ratios of positive and negative reviews in predominantly positive review sets to test the 'positivity effect' (Shoham et al., 2017) would allow finding nuances on when a positive review set might harm intentions. Other moderators should also be studied to improve our understanding of the nuances to the bandwagon effect and negativity bias. For instance, in chapter 2 we find that argument strength and writing quality were crucial for review readers when assessing a review's credibility and helpfulness. Therefore, further research could look into these characteristics to extend our findings from chapter 3 and provide a more thorough understanding of the negativity bias and bandwagon effect, being these commonly used theories on the studies in eWOM.

The literature review presented in chapter 4 also has some limitations, the biggest being that for purposes of parsimony and feasibility, the scope is narrowed to managerial responses to eWOM or online reviews, leaving aside the extensive literature on (offline) complaint handling. Also, it does not include research published in languages other than English or prior to 2000, compromising a more geographically diverse or historical perspective on the topic of webcare. The suggestions for further studies in chapter 4 were discussed earlier in this chapter.

Chapter 5, building upon chapter 4, solves some of the gaps raised in previous literature. However, our findings in this study leave room for further research due to some inconsistencies between what is expected based on previous research and the results from our analysis. For instance, the positive effect of having managers signing the webcare on future hotel bookings and the negative effects of some of the conversational elements that determine the tone of voice (invite for another visit and use non-verbal cues) seem to contradict what was previously found. Further research should apply our study to other contexts and look into the reasons why this occurs. As previously mentioned, these differences might be probably due to what should be expected for the reviewers and for the bystanders, that read the review and the webcare but are not personally involved in the situation. For instance, while reviewers might appreciate a more friendly tone since they had a previous interaction with the hotel, bystanders might appreciate a more professional tone.

Further research should also look into the effects on future business of combining the webcare strategies covered in chapter 5, following recommendations from previous research (Dens et al., 2015; van Hooijdonk & Liebrecht, 2021). Finally, considering that hotels received a very high volume of reviews and that managing them requires resources that are not easy to allocate, especially for smaller businesses, we suggest that further research looks into which online reviews should be answered. Our findings from chapters 2 and 3 show that well-written reviews with multiple and important arguments have the biggest influence on consumers' attitudes and intentions. Therefore, future research should look into finding other characteristics that allow managers to know which reviews to prioritize as well as the best webcare strategies to use when responding to these reviews in order to achieve the best business results.

# 7. References

Aerts, G., Smits, T., & Verlegh, P. W. (2017). The platform shapes the message: How website design affects abstraction and valence of online consumer reviews. *Decision Support Systems*, *104*, 104-112.

Allard, T., Dunn, L. H., & White, K. (2020). Negative Reviews, Positive Impact: Consumer Empathetic Responding to Unfair Word of Mouth. *Journal of Marketing*, 0022242920924389.

Allsop, D. T., Bassett, B. R., & Hoskins, J. A. (2007). Word-of-mouth research: principles and applications. *Journal of advertising research*, *47*(4), 398-411.

Anderson, C. K., & Han, S. (2016). *Hotel performance impact of socially engaging with consumers* (Cornell Hospitality Report, Issue 16).

Askalidis, G., Kim, S. J., & Malthouse, E. C. (2017). Understanding and overcoming biases in online review systems. *Decision Support Systems*, *97*, 23-30. https://doi.org/10.1016/j.dss.2017.03.002

Babić Rosario, A., Sotgiu, F., De Valck, K., & Bijmolt, T. H. (2016). The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors. *Journal of Marketing Research*, *53*(3), 297-318.

Bach, S. B., & Kim, S. (2012). Online consumer complaint behaviors: The dynamics of service failures, consumers' word of mouth, and organization-consumer relationships. *International Journal of Strategic Communication*, *6*(1), 59-76.

Baek, H., Ahn, J., & Choi, Y. (2012). Helpfulness of Online Consumer Reviews: Readers' Objectives and Review Cues. *International Journal of Electronic Commerce*, *17*(2), 99-126. https://doi.org/10.2753/JEC1086-4415170204

Baek, H., Lee, S., Oh, S., & Ahn, J. (2015). Normative social influence and online review helpfulness: Polynomial modeling and response surface analysis. *Journal of Electronic Commerce Research*, *16*(4), 290.

Baek, S.-H., Ham, S., & Yang, I.-S. (2006). A cross-cultural comparison of fast food restaurant selection criteria between Korean and Filipino college students. *International Journal of Hospitality Management*, *25*(4), 683-698.

Bandura, A., & McClelland, D. C. (1977). *Social learning theory* (Vol. 1). Englewood cliffs Prentice Hall.

Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Perception & psychophysics*, *8*(5), 279-286.

Bhandari, M., & Rodgers, S. (2018). What does the brand say? Effects of brand feedback to negative eWOM on brand trust and purchase intentions. *International Journal of Advertising*, *37*(1), 125-141.

Blattberg, R. C., Briesch, R., & Fox, E. J. (1995). How promotions work. *Marketing Science*, *14*(3_supplement), G122-G132.

Brooks Jr, R. C. (1957). “Word-of-mouth” advertising in selling new products. *Journal of Marketing, 22*(2), 154-161.

Brunner, C. B., Ullrich, S., & De Oliveira, M. J. (2019). The most optimal way to deal with negative consumer review: Can positive brand and customer responses rebuild product purchase intentions? *Internet Research, 29*(1), 104-122.

Cacioppo, J. T., & Petty, R. E. (1979). Effects of message repetition and position on cognitive response, recall, and persuasion. *Journal of personality and social psychology, 37*(1), 97.

Cacioppo, J. T., & Petty, R. E. (1989). Effects of message Repetition on Argument Processing, Recall, and Persuasion. *Basic and Applied Social Psychology, 10*(1), 3-12. https://doi.org/10.1207/s15324834basp1001_2

Calder, B. J., Insko, C. A., & Yandell, B. (1974). The Relation of Cognitive and Memorial Processes to Persuasion in a Simulated Jury Trial. *Journal of Applied Social Psychology, 4*(1), 62-93. https://doi.org/doi:10.1111/j.1559-1816.1974.tb02600.x

Cao, Q., Duan, W., & Gan, Q. (2011). Exploring determinants of voting for the “helpfulness” of online user reviews: A text mining approach. *Decision Support Systems, 50*, 511-521. https://doi.org/10.1016/j.dss.2010.11.009

Capoccia, C. (2018). *Online Reviews Are The Best Thing That Ever Happened To Small Businesses*. Forbes. Retrieved 11/03/2021 from https://www.forbes.com/sites/forbestechcouncil/2018/04/11/online-reviews-are-the-best-thing-that-ever-happened-to-small-businesses/?sh=53dcf650740a

Carstensen, L. L., & DeLiema, M. (2018). The positivity effect: a negativity bias in youth fades with age. *Current Opinion in Behavioral Sciences, 19*, 7-12. https://doi.org/10.1016/j.cobeha.2017.07.009

Casado-Díaz, A. B., Andreu, L., Beckmann, S. C., & Miller, C. (2020). Negative online reviews and webcare strategies in social media: effects on hotel attitude and booking intentions. *Current Issues in Tourism, 23*(4), 418-422.

Cenni, I., & Goethals, P. (2020). Responding to negative hotel reviews: A cross-linguistic perspective on online rapport-management. *Discourse, Context & Media, 37*, 100430.

Chang, C. (2009). Repetition variation strategies for narrative advertising. *Journal of advertising, 38*(3), 51-66.

Chang, H. H., Tsai, Y.-C., Wong, K. H., Wang, J. W., & Cho, F. J. (2015). The effects of response strategies and severity of failure on consumer attribution with regard to negative word-of-mouth. *Decision Support Systems, 71*, 48-61.

Chen, W., Gu, B., Ye, Q., & Zhu, K. X. (2019). Measuring and managing the externality of managerial responses to online customer reviews. *Information Systems Research, 30*(1), 81-96.

Cheung, C., Lee, M. K. O., & Rabjohn, N. (2008). The impact of electronic word-of-mouth: The adoption of online opinions in online customer communities. *Internet Research, 18*, 229-247. https://doi.org/10.1108/10662240810883290

Cheung, C. M.-Y., Sia, C.-L., & Kuan, K. K. Y. (2012). Is this review believable? A study of factors affecting the credibility of online consumer reviews from an ELM perspective. *Journal of the Association for Information Systems, 13*, 618.

Cheung, C. M. K., & Lee, M. K. O. (2008). Online consumer reviews: Does negative electronic word-of-mouth hurt more? *AMCIS 2008 Proceedings*, 143.

Cheung, M. Y., Luo, C., Sia, C. L., & Chen, H. (2009). Credibility of Electronic Word-of-Mouth: Informational and Normative Determinants of On-line Consumer Recommendations. *International Journal of Electronic Commerce, 13*(4), 9-38. https://doi.org/10.2753/JEC1086-4415130402

Chevalier, J. A., Dover, Y., & Mayzlin, D. (2018). Channels of Impact: User reviews when quality is dynamic and managers respond. *Marketing Science, 37*(5), 688-709.

Chevalier, J. A., & Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research, 43*, 345-354. https://doi.org/10.1509/jmkr.43.3.345

Chong, A. Y. L., Khong, K. W., Ma, T., McCabe, S., & Wang, Y. (2018). Analyzing key influences of tourists' acceptance of online reviews in travel decisions. *Internet Research, 28*(3), 564-586.

Chua, A. Y. K., & Banerjee, S. (2015). Understanding review helpfulness as a function of reviewer reputation, review rating, and review depth. *Journal of the Association for Information Science and Technology, 66*, 354-362. https://doi.org/10.1002/asi.23180

Clare, C. J., Wright, G., Sandiford, P., & Caceres, A. P. (2018). Why should I believe this? Deciphering the qualities of a credible online customer review. *Journal of Marketing Communications, 24*, 823-842. https://doi.org/10.1080/13527266.2016.1138138

Colicev, A., Kumar, A., & O'Connor, P. (2019). Modeling the relationship between firm and user generated content and the stages of the marketing funnel. *International Journal of Research in Marketing, 36*(1), 100-116.

Colliander, J., Dahlén, M., & Modig, E. (2015). Twitter for two: Investigating the effects of dialogue with customers in social media. *International Journal of Advertising, 34*(2), 181-194.

Collinger, T., Malthouse, E., & Maslowska, E. (2011). *The Consumer Engagement Engine: A Contemporary Framework*. https://spiegel.medill.northwestern.edu/_pdf/NU_Consumer%20Engagement%20Engine.pdf

Cowley, E. (2005). Views from consumers next in line: the fundamental attribution error in a service setting. *Journal of the Academy of Marketing Science, 33*(2), 139-152.

Craciun, G., & Moore, K. (2019). Credibility of negative online product reviews: Reviewer gender, reputation and emotion effects. *Computers in Human Behavior, 97*, 104-115.

Crijns, H., Cauberghe, V., Hudders, L., & Claeys, A.-S. (2017). How to deal with online consumer comments during a crisis? The impact of personalized organizational responses on organizational reputation. *Computers in Human Behavior, 75*, 619-631.

Dao, H. M., & Theotokis, A. (2021). Self-Service Technology Recovery: The Effect of Recovery Initiation and Locus of Responsibility. *Journal of interactive marketing, 54*, 25-39.

Davidow, M. (2000). The bottom line impact of organizational responses to customer complaints. *Journal of Hospitality & Tourism Research, 24*(4), 473-490.

Davidow, M. (2003). Organizational responses to customer complaints: What works and what doesn't. *Journal of service research, 5*(3), 225-250.

De Keyzer, F., Dens, N., & De Pelsmacker, P. (2017). Don't be so emotional! How tone of voice and service type affect the relationship between message valence and consumer responses to WOM in social media. *Online Information Review, 41*, 905-920. https://doi.org/10.1108/OIR-08-2016-0219

de Matos, C. A., & Rossi, C. A. V. (2008). Word-of-mouth communications in marketing: a meta-analytic review of the antecedents and moderators. *Journal of the Academy of Marketing Science, 36*, 578-596. https://doi.org/10.1007/s11747-008-0121-1

De Meulenaer, S., Dens, N., & De Pelsmacker, P. (2015). Which cues cause consumers to perceive brands as more global? A conjoint analysis. *International Marketing Review, 32*, 606-626. https://doi.org/10.1108/IMR-04-2014-0144

De Pelsmacker, P., Dens, N., & Kolomiiets, A. (2018). The impact of text valence, star rating and rated usefulness in online reviews. *International Journal of Advertising, 37*, 340-359. https://doi.org/10.1080/02650487.2018.1424792

De Pelsmacker, P., Geuens, M., & Van den Bergh, J. (2021). *Marketing communications: A European perspective* (7th edition ed.). Pearson education.

De Pelsmacker, P., Van Tilburg, S., & Holthof, C. (2018). Digital marketing strategies, online reviews and hotel performance. *International Journal of Hospitality Management, 72*, 47-55.

Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The Truth About the Truth: A Meta-Analytic Review of the Truth Effect. *Personality and Social Psychology Review, 14*(2), 238-257. https://doi.org/10.1177/1088868309352251

Deighton, J., & Kornfeld, L. (2009). Interactivity's unanticipated consequences for marketers and marketing. *Journal of interactive marketing, 23*(1), 4-10.

Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management science, 49*(10), 1407-1424.

Demmers, J., Van Dolen, W. M., & Weltevreden, J. W. (2018). Handling consumer messages on social networking sites: customer service or privacy infringement? *International Journal of Electronic Commerce, 22*(1), 8-35.

Dens, N., & De Pelsmacker, P. (2010). Consumer response to different advertising appeals for new products: The moderating influence of branding strategy and product category involvement. *Journal of Brand Management, 18*(1), 50-65. https://doi.org/10.1057/bm.2010.22

Dens, N., De Pelsmacker, P., & Purnawirawan, N. (2015). "We(b)care" How review set balance moderates the appropriate response strategy to negative online reviews. *Journal of Service Management, 26*, 486-515. https://doi.org/10.1108/JOSM-03-2014-0082

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Dixon, L., Li, J., Sorensen, J., Thain, N., & Vasserman, L. (2018). Measuring and mitigating unintended bias in text classification. Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society,

Doh, S.-J., & Hwang, J.-S. (2009). How Consumers Evaluate eWOM (Electronic Word-of-Mouth) Messages. *CyberPsychology & Behavior, 12*(2), 193-197. https://doi.org/10.1089/cpb.2008.0109

Drollinger, T., Comer, L. B., & Warrington, P. T. (2006). Development and validation of the active empathetic listening scale. *Psychology & Marketing, 23*(2), 161-180.

Edwards, A., & de Kool, D. (2015). Webcare in Public Services: Deliver Better with Less? In *Social Media for Government Services* (pp. 151-166). Springer.

Einwiller, S. A., & Steilen, S. (2015). Handling complaints on social network sites–An analysis of complaints and complaint responses on Facebook and Twitter pages of large US companies. *Public Relations Review, 41*(2), 195-204.

Ernst, N., Kühne, R., & Wirth, W. (2017). Effects of message repetition and negativity on credibility judgments and political attitudes. *International Journal of Communication, 11*, 21.

Esmark Jones, C. L., Stevens, J. L., Breazeale, M., & Spaid, B. I. (2018). Tell it like it is: The effects of differing responses to negative online reviews. *Psychology & Marketing, 35*(12), 891-901.

Evans, A. M., Stavrova, O., & Rosenbusch, H. (2020). Expressions of doubt and trust in online user reviews. *Computers in Human Behavior, 114*, 106556.

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. https://doi.org/https://doi.org/10.1016/j.jbusres.2014.11.006

Filieri, R., Hofacker, C. F., & Alguezaui, S. (2018). What makes information in online consumer reviews diagnostic over time? The role of review relevancy, factuality, currency, source credibility and ranking score. *Computers in Human Behavior, 80*, 122-131.

Filieri, R., McLeay, F., Tsui, B., & Lin, Z. (2018). Consumer perceptions of information helpfulness and determinants of purchase intention in online consumer reviews of services. *Information & Management, 55*(8), 956-970.

Floyd, K., Freling, R., Alhoqail, S., Cho, H. Y., & Freling, T. (2014). How online product reviews affect retail sales: A meta-analysis. *Journal of Retailing, 90*(2), 217-232.

Fong, J., & Burton, S. (2008). A cross-cultural comparison of electronic word-of-mouth and country-of-origin effects. *Journal of Business Research, 61*(3), 233-242.

Gelbrich, K., & Roschk, H. (2011). A meta-analysis of organizational complaint handling and customer responses. *Journal of service research, 14*(1), 24-43.

Ghosh, T. (2017). Managing negative reviews: the persuasive role of webcare characteristics. *Journal of Internet Commerce, 16*(2), 148-173.

González-Carvajal, S., & Garrido-Merchán, E. C. (2020). Comparing BERT against traditional machine learning text classification. *arXiv preprint arXiv:2005.13012*.

Gottschalk, S. A., & Mafael, A. (2017). Cutting Through the Online Review Jungle — Investigating Selective eWOM Processing. *Journal of interactive marketing, 37*, 89-104. https://doi.org/10.1016/j.intmar.2016.06.001

Grégoire, Y., Salle, A., & Tripp, T. M. (2015). Managing social media crises with your customers: The good, the bad, and the ugly. *Business Horizons, 58*(2), 173-182.

Grewal, L., & Stephen, A. T. (2019). In mobile we trust: The effects of mobile versus nonmobile reviews on consumer purchase intentions. *Journal of Marketing Research, 56*(5), 791-808.

Gu, B., & Ye, Q. (2014). First step in social media: Measuring the influence of online management responses on customer satisfaction. *Production and Operations Management, 23*(4), 570-582.

Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tourism Management, 59*, 467-483.

Ha, J., & Jang, S. S. (2009). Perceived justice in service recovery and behavioral intentions: The role of relationship quality. *International Journal of Hospitality Management, 28*(3), 319-327.

Hair, J. J., Anderson, R. E., Tatham, R. L., & Black, W. C. (1992). *Multivariate Data Analysis*. Prentice Hall.

Hair, M. L., & Bond, S. D. (2018). Attribute Dismissal and Valence Effects in Preferential Decision Processing. *Journal of Behavioral Decision Making, 31*(1), 164-178. https://doi.org/10.1002/bdm.2054

Harkins, S. G., & Petty, R. E. (1981). The multiple source effect in persuasion: The effects of distraction. *Personality and Social Psycology Bulletin, 7*(4), 627-635.

Hartmann, J., Huppertz, J., Schamp, C., & Heitmann, M. (2019). Comparing automated text classification methods. *International Journal of Research in Marketing, 36*(1), 20-38.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.

He, S. X., & Bond, S. D. (2015). Why is the crowd divided? Attribution for dispersion in online word of mouth. *Journal of Consumer research, 41*(6), 1509-1527.

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (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. https://doi.org/10.1002/dir.10073

Hennig-Thurau, T., Hofacker, C. F., & Bloching, B. (2013). Marketing the pinball way: Understanding how social media change the generation of value for consumers and companies. *Journal of interactive marketing, 27*(4), 237-241.

Hennig-Thurau, T., Malthouse, E. C., Friege, C., Gensler, S., Lobschat, L., Rangaswamy, A., & Skiera, B. (2010). The impact of new media on customer relationships. *Journal of service research, 13*(3), 311-330.

Herhausen, D., Ludwig, S., Grewal, D., Wulf, J., & Schoegel, M. (2019). Detecting, preventing, and mitigating online firestorms in brand communities. *Journal of Marketing, 83*(3), 1-21.

Herr, P. M., Kardes, F. R., & Kim, J. (1991). Effects of Word-of-Mouth and Product-Attribute Information on Persuasion: An Accessibility-Diagnosticity Perspective. *Journal of Consumer research, 17*, 454. https://doi.org/10.1086/208570

Homburg, C., Ehm, L., & Artz, M. (2015). Measuring and managing consumer sentiment in an online community environment. *Journal of Marketing Research, 52*(5), 629-641.

Homburg, C., & Fürst, A. (2007). See no evil, hear no evil, speak no evil: a study of defensive organizational behavior towards customer complaints. *Journal of the Academy of Marketing Science, 35*(4), 523-536.

Hong, H., Xu, D., Wang, G. A., & Fan, W. (2017). Understanding the determinants of online review helpfulness: A meta-analytic investigation. *Decision Support Systems, 102*, 1-11.

Honisch, S. V., & Manchón, L. M. (2019). The Effects of Paracrisis Origin and Response Strategy on Facebook Audience's Perceived Organisational Reputation and Behavioural Intentions. *Corporate Reputation Review*, 1-12.

Huang, R., & Ha, S. (2020). The effects of warmth-oriented and competence-oriented service recovery messages on observers on online platforms. *Journal of Business Research, 121*, 616-627.

Ismagilova, E., Dwivedi, Y. K., Slade, E., & Williams, M. D. (2017). Electronic Word of Mouth (eWOM) in the Marketing Context. 138. https://doi.org/10.1007/978-3-319-52459-7

Janssens, W., De Pelsmacker, P., Wijnen, K., & Van Kenhove, P. (2008). *Marketing research with SPSS*. Pearson Education.

Javornik, A., Filieri, R., & Gumann, R. (2020). “Don't Forget that Others Are Watching, Too!” The Effect of Conversational Human Voice and Reply Length on Observers' Perceptions of Complaint Handling in Social Media. *Journal of interactive marketing, 50*, 100-119.

Jeong, H.-J., & Koo, D.-M. (2015). Combined effects of valence and attributes of e-WOM on consumer judgment for message and product. *Internet Research*, *25*(1), 2-29. https://doi.org/10.1108/IntR-09-2013-0199

Jiang, L., O'Hern, M., & Hanson, S. (2019). Who's got my back? Comparing consumers' reactions to peer-provided and firm-provided customer support. *Psychology & Marketing*.

Jiménez, F. R., & Mendoza, N. A. (2013). Too popular to ignore: The influence of online reviews on purchase intentions of search and experience products. *Journal of interactive marketing*, *27*(3), 226-235.

Johnen, M., & Schnittka, O. (2019). When pushing back is good: the effectiveness of brand responses to social media complaints. *Journal of the Academy of Marketing Science*, 1-21.

Keller, E. (2007). Unleashing the power of word of mouth: Creating brand advocacy to drive growth. *Journal of advertising research*, *47*(4), 448-452.

Kemp, E., Porter, M., Fuller, N. R., & Min, K. S. (2020). I heard it through the grapevine: managing and engaging customers on the web. *Journal of Marketing Theory and Practice*, 1-16.

Kim, S. J., Maslowska, E., & Malthouse, E. C. (2018). Understanding the effects of different review features on purchase probability. *International Journal of Advertising*, *37*, 29-53. https://doi.org/10.1080/02650487.2017.1340928

Kim, S. J., Wang, R. J.-H., Maslowska, E., & Malthouse, E. C. (2016). “Understanding a fury in your words”: The effects of posting and viewing electronic negative word-of-mouth on purchase behaviors. *Computers in Human Behavior*, *54*, 511-521. https://doi.org/10.1016/j.chb.2015.08.015

Kim, W. G., Lim, H., & Brymer, R. A. (2015). The effectiveness of managing social media on hotel performance. *International Journal of Hospitality Management*, *44*, 165-171.

King, R. A., Racherla, P., & Bush, V. D. (2014). What We Know and Don't Know About Online Word-of-Mouth: A Review and Synthesis of the Literature. *Journal of interactive marketing, 28*, 167-183. https://doi.org/10.1016/j.intmar.2014.02.001

Kniesel, H., Waiguny, M. K., & Diehl, S. (2016). Effects of online review response strategies on attitudes toward the hotel. In *Advances in Advertising Research (Vol. VI)* (pp. 85-98). Springer.

Kolomiiets, A., Dens, N., & De Pelsmacker, P. (2016). The Wrap Effect in Online Review Sets Revisited: How Perceived Usefulness Mediates the Effect on Intention Formation. *Journal of Electronic Commerce Research, 17*(4), 280-288.

Kozinets, R. V., De Valck, K., Wojnicki, A. C., & Wilner, S. J. (2010). Networked narratives: Understanding word-of-mouth marketing in online communities. *Journal of Marketing, 74*(2), 71-89.

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage publications.

Ku, H.-H., Shang, R.-A., & Fu, Y.-F. (2021). Social learning effects of complaint handling on social media: Self-construal as a moderator. *Journal of Retailing and Consumer Services, 59*.

Kwok, L., & Xie, K. L. (2016). Factors contributing to the helpfulness of online hotel reviews: does manager response play a role? *International Journal of Contemporary Hospitality Management, 28*(10), 2156-2177.

Lamb, Y., Cai, W., & McKenna, B. (2020). Exploring the complexity of the individualistic culture through social exchange in online reviews. *International Journal of Information Management, 54*, 102198.

Lang, A. (2000). The Limited Capacity Model of Mediated Message Processing. *Journal of communication, 50*, 46-70. https://doi.org/10.1111/j.1460-2466.2000.tb02833.x

Lappeman, J., Patel, M., & Appalraju, R. (2018). Firestorm Response: Managing Brand Reputation during an nWOM Firestorm by Responding to Online Complaints Individually or as a Cluster. *Communicatio, 44*(2), 67-87.

Lavidge, R. J., & Steiner, G. A. (1961). A model for predictive measurements of advertising effectiveness. *Journal of Marketing, 25*(6), 59-62.

Lee, C. H., & Cranage, D. A. (2014). Toward understanding consumer processing of negative online word-of-mouth communication: the roles of opinion consensus and organizational response strategies. *Journal of Hospitality & Tourism Research, 38*(3), 330-360.

Lee, J., Park, D.-H., & Han, I. (2008). The effect of negative online consumer reviews on product attitude: An information processing view. *Electronic Commerce Research and Applications, 7*, 341-352. https://doi.org/10.1016/j.elerap.2007.05.004

Lee, M., Rodgers, S., & Kim, M. (2009). Effects of Valence and Extremity of eWOM on Attitude toward the Brand and Website. *Journal of Current Issues & Research in Advertising, 31*(2), 1-11. https://doi.org/10.1080/10641734.2009.10505262

Lee, S., Ha, T., Lee, D., & Kim, J. H. (2018). Understanding the majority opinion formation process in online environments: An exploratory approach to Facebook. *Information Processing & Management, 54*(6), 1115-1128.

Lee, Y. L., & Song, S. (2010). An empirical investigation of electronic word-of-mouth: Informational motive and corporate response strategy. *Computers in Human Behavior, 26*(5), 1073-1080.

Levy, D. S. (1995). Modern marketing research techniques and the property professional. *Property Management, 13*(3), 33-40.

Li, C., Cui, G., & Peng, L. (2018). Tailoring management response to negative reviews: The effectiveness of accommodative versus defensive responses. *Computers in Human Behavior, 84*, 272-284.

Li, H., Wang, C. R., Meng, F., & Zhang, Z. (2019). Making restaurant reviews useful and/or enjoyable? The impacts of temporal, explanatory, and sensory cues. *International Journal of Hospitality Management, 83*, 257-265.

Li, K., Chen, Y., & Zhang, L. (2020). Exploring the influence of online reviews and motivating factors on sales: A meta-analytic study and the moderating role of product category. *Journal of Retailing and Consumer Services, 55*, 102107.

Li, L., Lee, K. Y., Emokpae, E., & Yang, S.-B. (2021). What makes you continuously use chatbot services? Evidence from chinese online travel agencies. *Electronic Markets*, 1-25.

Liebrecht, C., Tsaousi, C., & van Hooijdonk, C. (2021). Linguistic elements of conversational human voice in online brand communication: Manipulations and perceptions. *Journal of Business Research, 132*, 124-135.

Liebrecht, C., & van Hooijdonk, C. (2019). Creating Humanlike Chatbots: What Chatbot Developers Could Learn From Webcare Employees In Adopting A Conversational Human Voice. International Workshop on Chatbot Research and Design, Amsterdam.

Liu, H., Jayawardhena, C., Dibb, S., & Ranaweera, C. (2019). Examining the trade-off between compensation and promptness in eWOM-triggered service recovery: A restorative justice perspective. *Tourism Management, 75*, 381-392.

Liu, H., Jayawardhena, C., Osburg, V.-S., & Babu, M. M. (2019). Do online reviews still matter post-purchase? *Internet Research, 30*(1), 109-139.

Liu, W., & Ji, R. (2019). Do Hotel Responses Matter?: A Comprehensive Perspective on Investigating Online Reviews. *Information Resources Management Journal (IRMJ), 32*(3), 1-20.

Liu, X., Schuckert, M., & Law, R. (2015). Can response management benefit hotels? Evidence from Hong Kong hotels. *Journal of Travel & Tourism Marketing, 32*(8), 1069-1080.

Lopes, A. I., Dens, N., De Pelsmacker, P., & De Keyzer, F. (2020). Which cues influence the perceived usefulness and credibility of an online review? A conjoint analysis. *Online Information Review, 45*(1), 1-20. https://doi.org/10.1108/OIR-09-2019-0287

Lui, T.-W., Bartosiak, M., Piccoli, G., & Sadhya, V. (2018). Online review response strategy and its effects on competitive performance. *Tourism Management, 67*, 180-190.

Luo, L., Duan, S., Shang, S., & Pan, Y. (2021). What makes a helpful online review? Empirical evidence on the effects of review and reviewer characteristics. *Online Information Review*.

Ma, L., Sun, B., & Kekre, S. (2015). The Squeaky Wheel Gets the Grease—An empirical analysis of customer voice and firm intervention on Twitter. *Marketing Science, 34*(5), 627-645.

Ma, Y., Xiang, Z., Du, Q., & Fan, W. (2018). Effects of user-provided photos on hotel review helpfulness: An analytical approach with deep leaning. *International Journal of Hospitality Management, 71*, 120-131.

Mafael, A., Gottschalk, S. A., & Kreis, H. (2016). Examining Biased Assimilation of Brand-related Online Reviews. *Journal of interactive marketing, 36*, 91-106. https://doi.org/10.1016/j.intmar.2016.06.002

Majors, I. (2020). *Trust in Advertising is Dying. where do we go from here?* Because XM. Retrieved February 10 from https://www.becausexm.com/blog/trust-in-advertising-is-dying.-where-do-we-go-from-here

Maslowska, E., Malthouse, E. C., & Bernritter, S. F. (2017). Too good to be true: the role of online reviews' features in probability to buy. *International Journal of Advertising, 36*, 142-163. https://doi.org/10.1080/02650487.2016.1195622

Maslowska, E., Malthouse, E. C., & Collinger, T. (2016). The customer engagement ecosystem. *Journal of Marketing Management, 32*(5-6), 469-501.

Maslowska, E., Malthouse, E. C., & Viswanathan, V. (2017). Do customer reviews drive purchase decisions? The moderating roles of review exposure and price. *Decision Support Systems, 98*, 1-9. https://doi.org/10.1016/j.dss.2017.03.010

Mate, M. J., Trupp, A., & Pratt, S. (2019). Managing negative online accommodation reviews: Evidence from the Cook Islands. *Journal of Travel & Tourism Marketing, 36*(5), 627-644.

Mathwick, C., & Mosteller, J. (2017). Online reviewer engagement: A typology based on reviewer motivations. *Journal of service research, 20*(2), 204-218.

Mayweg-Paus, E., & Jucks, R. (2018). Conflicting evidence or conflicting opinions? Two-sided expert discussions contribute to experts' trustworthiness. *Journal of Language and Social Psychology, 37*(2), 203-223.

McCullough, J. L., & Ostrom, T. M. (1974). Repetition of highly similar messages and attitude change. *Journal of Applied Psychology*, *59*(3), 395-397. https://doi.org/10.1037/h0036658

Metzger, M. J., Hartsell, E. H., & Flanagin, A. J. (2020). Cognitive dissonance or credibility? A comparison of two theoretical explanations for selective exposure to partisan news. *Communication Research, 47*(1), 3-28.

Min, H., Lim, Y., & Magnini, V. P. (2015). Factors affecting customer satisfaction in responses to negative online hotel reviews: The impact of empathy, paraphrasing, and speed. *Cornell Hospitality Quarterly*, *56*(2), 223-231.

Moran, G., & Muzellec, L. (2017). eWOM credibility on social networking sites: A framework. *Journal of Marketing Communications, 23*, 149-161. https://doi.org/10.1080/13527266.2014.969756

Morwitz, V. G., Steckel, J. H., & Gupta, A. (2007). When do purchase intentions predict sales? *International Journal of Forecasting, 23*(3), 347-364.

Murphy, R. (2020). *Local Consumer Review Survey 2020*. https://www.brightlocal.com/research/local-consumer-review-survey/

Nan, G., Yang, J., & Dou, R. (2017). Do only review characteristics affect consumer's online behavior? A study of relationship between reviews. *Journal of Electronic Commerce Research*, *18*(4).

Navío-Marco, J., Ruiz-Gómez, L. M., & Sevilla-Sevilla, C. (2018). Progress in information technology and tourism management: 30 years on and 20 years after the internet-Revisiting Buhalis & Law's landmark study about eTourism. *Tourism Management, 69*, 460-470.

Netemeyer, R. G., Maxham, J. G., & Pullig, C. (2005). Conflicts in the Work–Family Interface: Links to Job Stress, Customer Service Employee Performance, and Customer Purchase Intent. *Journal of Marketing, 69*(2), 130-143. https://doi.org/10.1509/jmkg.69.2.130.60758

Nghiêm-Phú, B. (2018). Comment And Comment Response Strategies-An Analysis Of Gay Hotel Guests' Comments And Managers' Responses. *Tourism and hospitality management, 24*(1), 133-149.

Park, C., Wang, Y., Yao, Y., & Kang, Y. R. (2011). Factors influencing eWOM effects: Using experience, credibility, and susceptibility. *International Journal of Social Science and Humanity, 1*(1), 74.

Park, D.-H., & Kim, S. (2008). The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews. *Electronic Commerce Research and Applications, 7*, 399-410. https://doi.org/10.1016/j.elerap.2007.12.001

Park, D.-H., & Lee, J. (2008). eWOM overload and its effect on consumer behavioral intention depending on consumer involvement. *Electronic Commerce Research and Applications, 7*, 386-398. https://doi.org/10.1016/j.elerap.2007.11.004

Park, D.-H., Lee, J., & Han, I. (2007). The Effect of On-Line Consumer Reviews on Consumer Purchasing Intention: The Moderating Role of Involvement. *International Journal of Electronic Commerce, 11*(4), 125-148. https://doi.org/10.2753/JEC1086-4415110405

Park, J., Yi, Y., & Kang, D. (2019). The Effects of One-Sided vs. Two-Sided Review Valence on Electronic Word of Mouth (e-WOM): The Moderating Role of Sponsorship Presence. *ASIA MARKETING JOURNAL, 21*(2), 1-19.

Park, S.-Y., & Allen, J. P. (2013). Responding to online reviews: Problem solving and engagement in hotels. *Cornell Hospitality Quarterly, 54*(1), 64-73.

Pentina, I., Bailey, A. A., & Zhang, L. (2018). Exploring effects of source similarity, message valence, and receiver regulatory focus on yelp review persuasiveness and purchase intentions. *Journal of Marketing Communications, 24*(2), 125-145.

Petty, R. E., & Cacioppo, J. T. (1984). The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of personality and social psychology, 46*(1), 69.

Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of Persuasion. *Advances in Experimental Social Psychology, 19*, 123-205.

Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of Persuasion. *Communication and persuasion*, 123-205.

Phillips, P., Barnes, S., Zigan, K., & Schegg, R. (2017). Understanding the impact of online reviews on hotel performance: an empirical analysis. *Journal of Travel Research, 56*(2), 235-249.

Podium. (2017). *Consumers Get "Buy" With a Little Help From Their Friends*. Podium. Retrieved January 10 2021 from https://www.podium.com/resources/podium-state-of-online-reviews/

Proserpio, D., & Zervas, G. (2017). Online reputation management: Estimating the impact of management responses on consumer reviews. *Marketing Science, 36*(5), 645-665.

Purnawirawan, N., De Pelsmacker, P., & Dens, N. (2012). Balance and Sequence in Online Reviews: How Perceived Usefulness Affects Attitudes and Intentions. *Journal of interactive marketing, 26*(4), 244-255. https://doi.org/10.1016/j.intmar.2012.04.002

Purnawirawan, N., Dens, N., & De Pelsmacker, P. (2012). Balance and Sequence in Online Reviews: The Wrap Effect. *International Journal of Electronic Commerce, 17*(2), 71-98. https://doi.org/10.2753/JEC1086-4415170203

Purnawirawan, N., Dens, N., & De Pelsmacker, P. (2014). Expert reviewers beware! The effects of review set balance, review source and review content on consumer responses to online reviews. *Journal of Electronic Commerce Research, 15*(3), 162.

Purnawirawan, N., Eisend, M., De Pelsmacker, P., & Dens, N. (2015). A Meta-analytic Investigation of the Role of Valence in Online Reviews. *Journal of interactive marketing, 31*, 17-27. https://doi.org/10.1016/j.intmar.2015.05.001

Purnawirawan, N. A. (2013). *Consumer Responses to Positive and Negative Online Reviews* [PhD, University of Antwerp]. Antwerp.

Raju, A. (2019). Can reviewer reputation and webcare content affect perceived fairness? *Journal of Research in Interactive Marketing*.

Resnick, P., & Zeckhauser, R. (2002). Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system. *The Economics of the Internet and E-commerce, 11*(2), 23-25.

Rhee, H. T., Yang, S.-B., & Kim, K. (2016). Exploring the comparative salience of restaurant attributes: A conjoint analysis approach. *International Journal of Information Management, 36*(6), 1360-1370.

Riffe, D., Lacy, S., Fico, F., & Watson, B. (2019). *Analyzing media messages: Using quantitative content analysis in research*. Routledge.

Ro, H., & Wong, J. (2012). Customer opportunistic complaints management: A critical incident approach. *International Journal of Hospitality Management, 31*(2), 419-427.

Roggeveen, A. L., & Johar, G. V. (2002). Perceived Source Variability Versus Familiarity: Testing Competing Explanations for the Truth Effect. *Journal of consumer psychology, 12*(2), 81-91. https://doi.org/10.1207/S15327663JCP1202_02

Roozen, I., & Raedts, M. (2018). The effects of online customer reviews and managerial responses on travelers' decision-making processes. *Journal of Hospitality Marketing & Management, 27*(8), 973-996.

Rosario, A. B., de Valck, K., & Sotgiu, F. (2020). Conceptualizing the electronic word-of-mouth process: What we know and need to know about eWOM creation, exposure, and evaluation. *Journal of the Academy of Marketing Science*, 1-27.

Rose, M., & Blodgett, J. G. (2016). Should hotels respond to negative online reviews? *Cornell Hospitality Quarterly, 57*(4), 396-410.

Rozin, P., & Royzman, E. B. (2001). Negativity Bias, Negativity Dominance, and Contagion. *Personality and Social Psychology Review, 5*(4), 296-320. https://doi.org/10.1207/S15327957PSPR0504_2

Rucker, D. D., Petty, R. E., & Briñol, P. (2008). What's in a frame anyway?: A meta-cognitive analysis of the impact of one versus two sided message framing on attitude certainty. *Journal of consumer psychology*, *18*(2), 137-149.

Santini, F., Ladeira, W. J., Pinto, D. C., Herter, M. M., Sampaio, C. H., & Babin, B. J. (2020). Customer engagement in social media: a framework and meta-analysis. *Journal of the Academy of Marketing Science, 48*, 1-18.

Schaefers, T., & Schamari, J. (2016). Service recovery via social media: The social influence effects of virtual presence. *Journal of service research, 19*(2), 192-208.

Schamari, J., & Schaefers, T. (2015). Leaving the home turf: how brands can use webcare on consumer-generated platforms to increase positive consumer engagement. *Journal of interactive marketing, 30*, 20-33.

Schindler, R. M., & Bickart, B. (2005). Published word of mouth: Referable, consumer-generated information on the Internet. In C. P. Haugtvedt, K. Machleit, & R. F. Yalch (Eds.), *Online consumer psychology: Understanding and influencing consumer behavior in the virtual world* (Vol. 32, pp. 35-61). Lawrence Erlbaum Associates Hillsdale, NJ.

Schindler, R. M., & Bickart, B. (2012). Perceived helpfulness of online consumer reviews: The role of message content and style. *Journal of Consumer Behaviour, 11*, 234-243. https://doi.org/10.1002/cb.1372

Schlosser, A. E. (2011). Can including pros and cons increase the helpfulness and persuasiveness of online reviews? The interactive effects of ratings and arguments. *Journal of consumer psychology, 21*, 226-239. https://doi.org/10.1016/j.jcps.2011.04.002

Schmidt, S., & Eisend, M. (2015). Advertising repetition: A meta-analysis on effective frequency in advertising. *Journal of advertising, 44*(4), 415-428.

Scholz, J., & Smith, A. N. (2019). Branding in the age of social media firestorms: how to create brand value by fighting back online. *Journal of Marketing Management, 35*(11-12), 1100-1134.

Sebastianelli, R., & Tamimi, N. (2018). E-tailer website attributes and trust: understanding the role of online reviews. *Online Information Review, 42*(4), 506-519.

Sheng, J. (2019). Being Active in Online Communications: Firm Responsiveness and Customer Engagement Behaviour. *Journal of interactive marketing, 46*, 40-51.

Sheng, J., Amankwah-Amoah, J., Wang, X., & Khan, Z. (2019). Managerial Responses to Online Reviews: A Text Analytics Approach. *British Journal of Management, 30*(2), 315-327.

Sheng, J., Wang, X., & Amankwah-Amoah, J. (2021). The Value of Firm Engagement: How Do Ratings Benefit from Managerial Responses? *Decision Support Systems*.

Shoham, M., Moldovan, S., & Steinhart, Y. (2017). Positively useless: irrelevant negative information enhances positive impressions. *Journal of consumer psychology, 27*(2), 147-159.

Siering, M., Muntermann, J., & Rajagopalan, B. (2018). Explaining and predicting online review helpfulness: The role of content and reviewer-related signals. *Decision Support Systems, 108*, 1-12.

Singh, J. P., Irani, S., Rana, N. P., Dwivedi, Y. K., Saumya, S., & Roy, P. K. (2017). Predicting the “helpfulness” of online consumer reviews. *Journal of Business Research, 70*, 346-355.

Soh, H., Reid, L. N., & King, K. W. (2009). Measuring trust in advertising. *Journal of advertising, 38*, 83-104.

Sotiriadis, M. D., & Van Zyl, C. (2013). Electronic word-of-mouth and online reviews in tourism services: the use of twitter by tourists. *Electronic Commerce Research, 13*(1), 103-124.

Sparks, B. A., & Bradley, G. L. (2017). A "Triple A" typology of responding to negative consumer-generated online reviews. *Journal of Hospitality & Tourism Research, 41*(6), 719-745.

Sparks, B. A., So, K. K. F., & Bradley, G. L. (2016). Responding to negative online reviews: The effects of hotel responses on customer inferences of trust and concern. *Tourism Management, 53*, 74-85.

Sreejesh, S., Anusree, M., & Abhilash, P. (2019). Can online service recovery interventions benignly alter customers' negative review evaluations? Evidence from the hotel industry. *Journal of Hospitality Marketing & Management, 28*(6), 711-742.

Srivastava, V., & Kalro, A. D. (2019). Enhancing the helpfulness of online consumer reviews: The role of latent (content) factors. *Journal of interactive marketing, 48*, 33-50.

Stankevich, A. (2017). Explaining the consumer decision-making process: Critical literature review. *Journal of International Business Research and Marketing, 2*(6).

Stevens, J. L., Spaid, B. I., Breazeale, M., & Jones, C. L. E. (2018). Timeliness, transparency, and trust: A framework for managing online customer complaints. *Business Horizons, 61*(3), 375-384.

Sundar, S. S., Oeldorf-Hirsch, A., & Xu, Q. (2008). *The bandwagon effect of collaborative filtering technology*, ACM.

Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *12*(2), 257-285. https://doi.org/doi:10.1207/s15516709cog1202_4

Tang, F., Fu, L., Yao, B., & Xu, W. (2019). Aspect based fine-grained sentiment analysis for online reviews. *Information Sciences, 488*, 190-204.

Tata, S. V., Prashar, S., & Gupta, S. (2020). An examination of the role of review valence and review source in varying consumption contexts on purchase decision. *Journal of Retailing and Consumer Services, 52*.

Tathagata, G., & Amar, R. G. (2018). Gulping the poison: How webcare attributes reduce damages to brands caused by negative reviews. *Journal of Internet Commerce, 17*(3), 216-254.

Tathagata, G., & Mandal, S. (2020). Webcare quality: conceptualization, scale development and validation. *Journal of Marketing Management*, 1-35.

Thomas, M.-J., Wirtz, B. W., & Weyerer, J. C. (2019). Determinants of Online Review Credibility and its Impact on Consumer Purchase Intention. *Journal of Electronic Commerce Research*, *20*(1), 1-20.

Treviño, T., & Castaño, R. (2013). How should managers respond? Exploring the effects of different responses to negative online reviews. *International Journal of Leisure and Tourism Marketing*, *3*(3), 237-251.

Tsao, W.-C., Hsieh, M.-T., Shih, L.-W., & Lin, T. M. (2015). Compliance with eWOM: The influence of hotel reviews on booking intention from the perspective of consumer conformity. *International Journal of Hospitality Management*, *46*, 99-111.

Ullrich, S., & Brunner, C. B. (2015). Negative online consumer reviews: effects of different responses. *Journal of Product & Brand Management*, *24*(1), 66-77.

Valentini, S., Orsingher, C., & Polyakova, A. (2020). Customers' emotions in service failure and recovery: a meta-analysis. *Marketing Letters*, 1-18.

van Hooijdonk, C., & Liebrecht, C. (2021). Sorry but no sorry: The use and effects of apologies in airline webcare responses to NeWOM messages of flight passengers. *Discourse, Context & Media*, *40*, 100442.

Van Noort, G., & Willemsen, L. M. (2012). Online damage control: The effects of proactive versus reactive webcare interventions in consumer-generated and brand-generated platforms. *Journal of interactive marketing*, *26*(3), 131-140.

Van Noort, G., Willemsen, L. M., Kerkhof, P., & Verhoeven, J. W. (2015). Webcare as an integrative tool for customer care, reputation management, and online marketing: A literature review. In *Integrated communications in the postmodern era* (pp. 77-99). Springer.

Vargo, S. L., & Lusch, R. F. (2004). Evolving to a New Dominant Logic for Marketing. *Journal of Marketing*, *68*, 1-17. https://doi.org/10.1509/jmkg.68.1.1.24036

Ventre, I., & Kolbe, D. (2020). The Impact of Perceived Usefulness of Online Reviews, Trust and Perceived Risk on Online Purchase Intention in Emerging Markets: A Mexican Perspective. *Journal of International Consumer Marketing*, *32*(4), 287-297.

Verma, S., & Yadav, N. (2021). Past, Present, and Future of Electronic Word of Mouth (EWOM). *Journal of interactive marketing*, *53*, 111-128.

Vermeer, S. A., Araujo, T., Bernritter, S. F., & van Noort, G. (2019). Seeing the wood for the trees: How machine learning can help firms in identifying relevant electronic word-of-mouth in social media. *International Journal of Research in Marketing*.

Vermeulen, I. E., & Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism Management*, *30*, 123-127. https://doi.org/10.1016/j.tourman.2008.04.008

Wang, P., & Li, H. (2019). Understanding the antecedents and consequences of the perceived usefulness of travel review websites. *International Journal of Contemporary Hospitality Management*, *31*(3), 1086-1103.

Wang, S., Cunningham, N. R., & Eastin, M. S. (2015). The Impact of eWOM Message Characteristics on the Perceived Effectiveness of Online Consumer Reviews. *Journal of Interactive Advertising*, *15*, 151-159. https://doi.org/10.1080/15252019.2015.1091755

Wang, X., Peng, L., & Xu, F. (2018). Do incentives in SWOM communication matter? A positive emotion perspective. *Journal of Electronic Commerce Research, 19*(2), 135-153.

Wang, Y., & Chaudhry, A. (2018). When and how managers' responses to online reviews affect subsequent reviews. *Journal of Marketing Research*, *55*(2), 163-177.

Wang, Y., Wang, J., & Yao, T. (2019). What makes a helpful online review? A meta-analysis of review characteristics. *Electronic Commerce Research, 19*(2), 257-284.

Wang, Z., Mao, H., Li, Y. J., & Liu, F. (2017). Smile big or not? Effects of smile intensity on perceptions of warmth and competence. *Journal of Consumer research, 43*(5), 787-805.

Warren, C., Barsky, A., & McGraw, A. P. (2018). Humor, comedy, and consumer behavior. *Journal of Consumer research, 45*(3), 529-552.

Wei, W., Miao, L., & Huang, Z. J. (2013). Customer engagement behaviors and hotel responses. *International Journal of Hospitality Management*, *33*, 316-330.

Weiner, B. (2000). Attributional thoughts about consumer behavior. *Journal of Consumer research*, *27*(3), 382-387.

Weiner, B. (2010). The development of an attribution-based theory of motivation: A history of ideas. *Educational psychologist*, *45*(1), 28-36.

Weitzl, W., & Hutzinger, C. (2017). The effects of marketer-and advocate-initiated online service recovery responses on silent bystanders. *Journal of Business Research, 80*, 164-175.

Weitzl, W., Hutzinger, C., & Einwiller, S. (2018). An empirical study on how webcare mitigates complainants' failure attributions and negative word-of-mouth. *Computers in Human Behavior, 89*, 316-327.

Weitzl, W. J. (2019). Webcare's effect on constructive and vindictive complainants. *Journal of Product & Brand Management*.

Weitzl, W. J., & Einwiller, S. A. (2020). Profiling (un-) committed online complainants: Their characteristics and post-webcare reactions. *Journal of Business Research, 117*, 740-753.

Weitzl, W. J., & Hutzinger, C. (2019). Rise and fall of complainants' desires: The role of pre-failure brand commitment and online service recovery satisfaction. *Computers in Human Behavior, 97*, 116-129.

Wen, J., Lin, Z., Liu, X., Xiao, S. H., & Li, Y. (2020). The Interaction Effects of Online Reviews, Brand, and Price on Consumer Hotel Booking Decision Making. *Journal of Travel Research*, 0047287520912330.

Willemsen, L. M., Neijens, P. C., Bronner, F., & de Ridder, J. A. (2011a). "Highly Recommended!" The Content Characteristics and Perceived Usefulness of Online Consumer Reviews. *Journal of Computer-Mediated Communication, 17*(1), 19-38. https://doi.org/10.1111/j.1083-6101.2011.01551.x

Willemsen, L. M., Neijens, P. C., Bronner, F., & de Ridder, J. A. (2011b). "Highly Recommended!" The Content Characteristics and Perceived Usefulness of Online Consumer Reviews. *Journal of Computer-Mediated Communication, 17*(1), 19-38. https://doi.org/10.1111/j.1083-6101.2011.01551.x

Williams, M., & Buttle, F. (2011). The eight pillars of WOM management: Lessons from a multiple case study. *Australasian Marketing Journal (AMJ), 19*(2), 85-92.

WiT. (2019). *Google is now Top Site for hotel reviews, overtaking Booking.com*. World in Travel. Retrieved 11/03/2021 from https://www.webintravel.com/google-is-now-top-site-for-hotel-reviews-overtaking-booking-com/#:~:text=Booking.com%20only%20saw%20a,in%20the%20online%20travel%20sphere.

Wu, J., Wu, T., & Schlegelmilch, B. B. (2020). Seize the Day: How Online Retailers Should Respond to Positive Reviews. *Journal of interactive marketing*, *52*, 52-60.

Wu, P. F. (2013). In Search of Negativity Bias: An Empirical Study of Perceived Helpfulness of Online Reviews. *Psychology & Marketing*, *30*(11), 971-984. https://doi.org/10.1002/mar.20660

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.

Xia, L. (2013). Effects of companies' responses to consumer criticism in social media. *International Journal of Electronic Commerce*, *17*(4), 73-100.

Xia, L., & Bechwati, N. N. (2008). Word of mouse: the role of cognitive personalization in online consumer reviews. *Journal of Interactive Advertising*, *9*(1), 3-13.

Xie, K. L., So, K. K. F., & Wang, W. (2017). Joint effects of management responses and online reviews on hotel financial performance: A data-analytics approach. *International Journal of Hospitality Management, 62*, 101-110.

Xie, K. L., Zhang, Z., & Zhang, Z. (2014). The business value of online consumer reviews and management response to hotel performance. *International Journal of Hospitality Management*, *43*, 1-12.

Xie, K. L., Zhang, Z., Zhang, Z., Singh, A., & Lee, S. K. (2016). Effects of managerial response on consumer eWOM and hotel performance: Evidence from TripAdvisor. *International Journal of Contemporary Hospitality Management*, *28*(9), 2013-2034.

Xun, J., & Guo, B. (2017). Twitter as customer's eWOM: an empirical study on their impact on firm financial performance. *Internet Research, 27*(5), 1014-1038.

Yang, S.-B., Shin, S.-H., Joun, Y., & Koo, C. (2017). Exploring the comparative importance of online hotel reviews' heuristic attributes in review helpfulness: a conjoint analysis approach. *Journal of Travel & Tourism Marketing*, *34*(7), 963-985.

Yang, Y., Park, S., & Hu, X. (2018). Electronic word of mouth and hotel performance: A meta-analysis. *Tourism Management*, *67*, 248-260. https://doi.org/10.1016/j.tourman.2018.01.015

Ye, Q., Law, R., & Gu, B. (2009). The impact of online user reviews on hotel room sales. *International Journal of Hospitality Management*, *28*, 180-182. https://doi.org/10.1016/j.ijhm.2008.06.011

Zablocki, A., Schlegelmilch, B., & Houston, M. J. (2018). How valence, volume and variance of online reviews influence brand attitudes. *AMS Review*, *9*(1-2), 61-77. https://doi.org/10.1007/s13162-018-0123-1

Zhang, J. Q., Craciun, G., & Shin, D. (2010). When does electronic word-of-mouth matter? A study of consumer product reviews. *Journal of Business Research*, *63*, 1336-1341. https://doi.org/10.1016/j.jbusres.2009.12.011

Zhang, K. Z. K., Zhao, S. J., Cheung, C. M. K., & Lee, M. K. O. (2014). Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model. *Decision Support Systems*, *67*, 78-89. https://doi.org/10.1016/j.dss.2014.08.005

Zhang, L., Gao, Y., & Zheng, X. (2019). Let's Talk About This in Public: Consumer Expectations for Online Review Response. *Cornell Hospitality Quarterly*, 1938965519864864.

Zhang, X. J., Ko, M., & Carpenter, D. (2016). Development of a scale to measure skepticism toward electronic word-of-mouth. *Computers in Human Behavior*, *56*, 198-208. https://doi.org/10.1016/j.chb.2015.11.042

Zhang, Y., & Vásquez, C. (2014). Hotels' responses to online reviews: Managing consumer dissatisfaction. *Discourse, Context & Media*, *6*, 54-64.

# 8. Appendix

# Appendix 1: Stimuli and Questionnaire for Chapter 2

### Intro

Beste, Ik ben een student Toegepaste Economische Wetenschappen - Bedrijfskunde en ik ben geïnteresseerd in hoe mensen online reviews beoordelen.

Zou u mij willen helpen bij mijn onderzoek? Dit zou ongeveer 15 minuten van uw tijd in beslag nemen. Er zal vertrouwelijk met uw gegevens worden omgegaan en de resultaten worden geheel anoniem verwerkt.

### Intructions

Hier zijn nog enkele instructies voor u met de vragenlijst kan starten:

Stel u voor dat u een gps gaat kopen, voor deze aankoop gaat u online reviews raadplegen. U gaat verschillende online reviews te zien krijgen over eenzelfde gps die u moet beoordelen. Het is hierbij belangrijk dat u eerst de productomschrijving goed leest en de reviews los van elkaar beoordeeld.

### Product information

Tomo

€128,00

**Tomo Spirit 7500 – West-Europa**

5 inch (13 cm) | Gratis kaartupdates | 3 maanden gratis flitsupdates

Productbeschrijving

Ben je op zoek naar een eenvoudig navigatiesysteem? Dan is deze gps wat voor jou. Dit model is voorzien van rijbaanbegeleiding zodat je precies weet welke baan je moet aanhouden op snelwegen en in de stad. Je ontvangt levenslange kaartupdates voor West-Europa. Daarnaast geeft deze gps je gesproken navigatie-instructies zodat je de ogen op weg kan houden.

Productspecificaties

| | | | |
|---|---|---|---|
| Intern geheugen | 8 GB | Oplaadtijd | 120 uur |
| Schermgrootte | 5 inch (13 cm) inch | Inclusief USB-kabel | Ja |
| Schermresolutie | 480 x 272 pixels | Verkeersinformatie | Ja |
| Levensduur batterij | Tot 1 uur | Fabrieksgarantie | 2 jaar |

## Cards

### Card 1

Review (2 reviews)

+ goede prijs/kwaliteit
+ snelle routebepaling
+ stemherkenning werkt goed
+ scherm heeft een hoge resoluttie

- geen outomatische updates
- niet comptatibel met mijn mobiele telefoon

prijs/kwaliteit goed, voor prijs goed navigatiesysteem dat je lijdt naar bestemming.ook snelle routebepaling; stem goed herkennen + hoge resoluttie scherm. Nadeel aan Gps geen outomatische updates en niet comptatibel met gsm.

Vond je dit een nuttige review? 235 7

### Card 2

Review (2 reviews)

+ leuke accessoires bij te verkrijgen
+ mooi design
+ uitbereiding beschikbaar voor Amerika
+ fast food restaurants standaard op de kaart

- het paarse pijltje zou geel mogen zijn
- past niet in oude gps houder

Deze gps beschikt over een mooi design. Verder zijn er leuke accessoires bij dit model te verkrijgen. Ook is er een uitbereiding beschikbaar voor Amerika en alle fast food restaurants staan standaard op de kaart. Een nadeel aan dit toestel is dat het paarse pijltje op het scherm geel zou mogen zijn naar mijn persoonlijke mening. Daarnaast past dit model niet in mijn oude gps houder.

Vond je dit een nuttige review? 7 235

## Card 3

Review (2 reviews)

+ leuke acessoires bij te verkrijgen

+ mooi design

gps heeft mooi design en leuke acesoires bij model te krijgen.

Vond je dit een nuttige review? 235 7

## Card 4

Review (274 reviews)

+ leuke accessoires bij te verkrijgen - het paarse pijltje zou geel mogen zijn

+ mooi design

Deze gps heeft een mooi design en er zijn leuke accessoires bij dit model te verkrijgen. Een nadeel aan dit toestel is dat het paarse pijltje op het scherm geel zou mogen zijn naar mijn persoonlijke mening.

Vond je dit een nuttige review? 235 7

## Card 5

Review (2 reviews)

+ goede prijs/kwaliteit

+ snelle routebepaling

De prijs/kwaliteit van deze gps zit goed, voor deze prijs heb je een goed navigatiesysteem dat je leidt naar de plaats van bestemming.

Vond je deze review nuttig? 7 235

## Card 6

Review (274 reviews)

+ goede prijs/kwaliteit - geen outomatische updates

+ snelle routebepaling

De prijs/kwaliteit goed, voor prijs goed navigatiesysteem dat je lijdtt naar bestemming. Nadeel aan Gps geen outomatische updates

Vond je deze review nuttig? 7 235

## Card 7

Review (274 reviews)

+ goede prijs/kwaliteit

+ snelle routebepaling

+ stemherkenning werkt goed

+ scherm heeft een hoge resolutie

Deze gps is eenvoudig te bedienen en bijgevolg gebruiksvriendelijk. De prijs/kwaliteit zit goed, voor deze prijs heb je een goed navigatiesysteem dat je leidt naar de plaats van bestemming. Bovendien, heeft dit model een goede stemherkenning en het scherm van de gps heeft een hoge resolutie.

Vond je deze review nuttig? 235 7

<u>Card 8</u>

Review (274 reviews)

+ leuke acessoires bij te verkrijgen
+ mooi design
+ uitbereiding beschikbaar voor Amerika
+ fast foed restourants standaard op de kaart

gps heeft mooi design. zijn leuke acesoires bij model te krijgen; ook uitbereiding voor Amerika en fast foed restourants standaard op kaart

Vond je deze review nuttig? 7 235

## Questions displayed for each Card

Deze review

| | -3 | -2 | -1 | 0 | +1 | +2 | +3 | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| vond ik niet nuttig | ○ | ○ | ○ | ○ | ○ | ○ | ○ | vond ik nuttig |
| heeft mij niet geholpen om mijn mening ten opzichte van de gps te bepalen | ○ | ○ | ○ | ○ | ○ | ○ | ○ | heeft mij geholpen om mijn mening ten opzichte van de gps te bepalen |
| heeft mij niet geholpen om mijn aankoopbeslissing te maken | ○ | ○ | ○ | ○ | ○ | ○ | ○ | heeft mij geholpen om mijn aankoopbeslissing te maken |

Mijn oordeel over deze review is de volgende:

| | -3 | -2 | -1 | 0 | +1 | +2 | +3 | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| ik vind deze review slecht | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ik vind deze review goed |
| mijn houding ten opzichte van de review is zeer negatief | ○ | ○ | ○ | ○ | ○ | ○ | ○ | mijn houding ten opzichte van de review is zeer positief |
| deze review beantwoordt niet aan mijn verwachtingen | ○ | ○ | ○ | ○ | ○ | ○ | ○ | deze review beantwoordt aan mijn verwachtingen |

Ik denk dat de meeste mensen na het lezen van deze review

| | -3 | -2 | -1 | 0 | +1 | +2 | +3 | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| de gps slecht vinden | ○ | ○ | ○ | ○ | ○ | ○ | ○ | de gps goed vinden |
| een negatief oordeel over de gps houden | ○ | ○ | ○ | ○ | ○ | ○ | ○ | een positief oordeel over de gps houden |
| de gps zullen afraden bij hun familie of vrienden | ○ | ○ | ○ | ○ | ○ | ○ | ○ | de gps zullen aanraden bij hun familie of vrienden |

Hoe geloofwaardig is deze review?

| | -3 | -2 | -1 | 0 | +1 | +2 | +3 | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| heel ongeloofwaardig | ○ | ○ | ○ | ○ | ○ | ○ | ○ | heel geloofwaardig |

Zou je overwegen om deze gps te kopen op basis van bovenstaande online review?

| | -3 | -2 | -1 | 0 | +1 | +2 | +3 | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Ik zou de gps zeker niet kopen | ○ | ○ | ○ | ○ | ○ | ○ | ○ | Ik zou de gps zeker kopen |

**Measuring Involvement**

Een gps...

| | -3 | -2 | -1 | 0 | +1 | +2 | +3 | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| is onbelangrijk voor mij | ○ | ○ | ○ | ○ | ○ | ○ | ○ | belangrijk voor mij |
| betekent niets voor mij | ○ | ○ | ○ | ○ | ○ | ○ | ○ | betekent veel voor mij |
| maakt niet uit voor mij | ○ | ○ | ○ | ○ | ○ | ○ | ○ | maakt veel uit voor mij |

**Demographics**

Wat is uw geslacht?

- Man (1)
- Vrouw (2)

Wat is uw leeftijd?

__________________________________________________

Wat is uw hoogst genoten opleiding?

- Geen (1)
- Lager onderwijs (2)
- Middelbaar onderwijs (3)
- Hoger onderwijs (4)
- Universitair onderwijs (5)

# Appendix 2: Stimuli and Questionnaire for Chapter 3

**Intro**

Dear Madam, dear Sir,

This study is part of a PhD research project conducted by the University of Antwerp, Belgium. It aims to assess how people make decisions based on online reviews.

You can stop your participation and close the survey at any time, but please keep in mind that you should complete the whole survey in order to be compensated. We would also like to note that the survey contains several "attention checks", questions that probe if participants are paying attention while completing the survey.

All your responses are anonymous and will only be processed in aggregated form. Your responses will only be used for the purposes of the current research and will not be shared with third parties. All data collection and data processing complies with the Privacy Policy of the University of Antwerp, which you can consult here. This research has been approved by the University Ethics committee. Data will be preserved for 5 years.

If you have any questions or remarks concerning the questionnaire, please contact Ana Lopes via ana.lopes@uantwerpen.be.

By proceeding you confirm that you understand the information above and agree to take part in this study.

Thanks in advance for your cooperation.

Ana Lopes

prof.dr. Nathalie Dens

prof.dr. Patrick De Pelsmacker

**Instructions**

Before proceeding to the questions, please read the following definition:

An all-inclusive resort is a holiday resort that charges a set price for a total service that includes lodging, drinks (both alcoholic and non-alcoholic), food (meals and snacks), sports activities, for instance a gym, and entertainment. Depending on the hotel some additional services, for example safety deposit box or towel service, may not be included.

**Introductory questions**

How many times have you visited an all-inclusive-resort?

- Never
- 1 time
- 2-3 times
- 4-5 times
- More than 5 times

(If ‘How many times have you visited an all-inclusive-resort?’ was never:)

Do you know someone that has spent their holidays in an all-inclusive resort?

- Yes
- No

(If ‘*How many times have you visited an all-inclusive-resort?’* was never:)

Have you ever considered going to an all-inclusive resort?

- Yes
- No

How likely is it that you will visit an all-inclusive resort in the next year?

- Very unlikely
- Somewhat unlikely
- Neither likely nor unlikely
- Somewhat likely
- Very likely

**Stimuli** (randomly assigned to participants; participants saw the one of the sets of reviews that was assigned to them)

*Intro for stimuli*

You are planning your next vacation and after discussing with your partner, a family member or a friend, you have decided that you want to travel to a sunny destination. The destination is already fixed, you choose to go to an all-inclusive resort, and now you are looking for one. You find online an all-inclusive resort that is available for your dates. Please read the following reviews written by different travelers to help you decide if you are going to book it.

Stimulus 1:

| |
|---|
| It was a shame that the minibar didn’t offer much choice. |
| I was disappointed that the hotel does not offer any motorized watersports. |
| I wanted to rent a boat and had to find a rental company in the town… It would have been easier to book it through the hotel. |
| The sunbeds were very uncomfortable. I couldn’t lie on them for a long time. |
| The hotel garden was very small, I really felt like I could use some more green. |
| The hotel lobby was small, I felt crammed when we were checking in. |
| It was too bad that the hotel didn't have a wellness area. |
| The gym offers little variety in equipment. More choices would have been better. |
| The best spots by the pool were always taken. |
| There was no one at the hotel of our age. |
| The room was incredibly clean… It was really amazing to have such a tidy place to stay. |
| It was great that we could access the Wi-Fi in the room, really nice. |

Stimulus 2:

| |
|---|
| It was a shame that the minibar didn’t offer much choice. |
| I was disappointed that the hotel does not offer any motorized watersports. |
| I wanted to rent a boat and had to find a rental company in the town… It would have been easier to book it through the hotel. |
| The sunbeds were very uncomfortable. I couldn’t lie on them for a long time. |
| The hotel garden was very small, I really felt like I could use some more green. |
| The hotel lobby was small, I felt crammed when we were checking in. |
| It was too bad that the hotel didn't have a wellness area. |
| The gym offers little variety in equipment. More choices would have been better. |
| The best spots by the pool were always taken. |
| There was no one at the hotel of our age. |
| The food was always freshly made, amazing. |
| It was great that the buffet always had fresh food available. |

Stimulus 3:

| |
|---|
| It was a shame that the minibar didn’t offer much choice. |
| I was disappointed that the hotel does not offer any motorized watersports. |
| I wanted to rent a boat and had to find a rental company in the town… It would have been easier to book it through the hotel. |
| The sunbeds were very uncomfortable. I couldn’t lie on them for a long time. |
| The hotel lobby was small, I felt crammed when we were checking in. |
| It was too bad that the hotel didn't have a wellness area. |
| The gym offers little variety in equipment. More choices would have been better. |
| The best spots by the pool were always taken. |
| There was no one at the hotel of our age. |
| The bed was very comfortable. It was great to rest. |
| The room was incredibly clean… It was really amazing to have such a tidy place to stay. |
| It was great that we could access the Wi-Fi in the room, really nice. |

Stimulus 4:

| |
|---|
| I was disappointed that the hotel does not offer any motorized watersports. |
| I wanted to rent a boat and had to find a rental company in the town… It would have been easier to book it through the hotel. |
| The sunbeds were very uncomfortable. I couldn’t lie on them for a long time. |
| The hotel lobby was small, I felt crammed when we were checking in. |
| It was too bad that the hotel didn't have a wellness area. |
| The gym offers little variety in equipment. More choices would have been better. |
| The best spots by the pool were always taken. |
| There was no one at the hotel of our age. |
| The food was always freshly made, amazing. |
| I really enjoyed the food at the hotel, always fresh! |
| It was great that the buffet always had fresh food available. |

Stimulus 5:

| |
|---|
| It was a shame that the minibar didn’t offer much choice. |
| I was disappointed that the hotel does not offer any motorized watersports. |
| The sunbeds were very uncomfortable. I couldn’t lie on them for a long time. |
| The hotel lobby was small, I felt crammed when we were checking in. |
| It was too bad that the hotel didn't have a wellness area. |
| The gym offers little variety in equipment. More choices would have been better. |
| The best spots by the pool were always taken. |
| There was no one at the hotel of our age. |
| The food at the hotel restaurant was delicious, I loved it. |
| The bed was very comfortable. It was great to rest. |
| The room was incredibly clean… It was really amazing to have such a tidy place to stay. |
| It was great that we could access the Wi-Fi in the room, really nice. |

Stimulus 6:

| |
|---|
| It was a shame that the minibar didn’t offer much choice. |
| I was disappointed that the hotel does not offer any motorized watersports. |
| The sunbeds were very uncomfortable. I couldn’t lie on them for a long time. |
| The hotel lobby was small, I felt crammed when we were checking in. |
| It was too bad that the hotel didn't have a wellness area. |
| The gym offers little variety in equipment. More choices would have been better. |
| The best spots by the pool were always taken. |
| There was no one at the hotel of our age. |
| The food was always freshly made, amazing. |
| I loved the freshness of the food! |
| It was great that the buffet always had fresh food available. |
| I really enjoyed the food at the hotel, always fresh! |

Stimulus 7:

| |
|---|
| It was a shame that the minibar didn't offer much choice. |
| I was disappointed that the hotel does not offer any motorized watersports. |
| The sunbeds were very uncomfortable. I couldn't lie on them for a long time. |
| The hotel lobby was small, I felt crammed when we were checking in. |
| The gym offers little variety in equipment. More choices would have been better. |
| The best spots by the pool were always taken. |
| There was no one at the hotel of our age. |
| The food at the hotel restaurant was delicious, I loved it. |
| The bed was very comfortable. It was great to rest. |
| The room was incredibly clean… It was really amazing to have such a tidy place to stay. |
| The food was always freshly made, amazing. |
| It was great that we could access the Wi-Fi in the room, really nice. |

Stimulus 8:

| |
|---|
| It was a shame that the minibar didn't offer much choice. |
| I was disappointed that the hotel does not offer any motorized watersports. |
| The sunbeds were very uncomfortable. I couldn't lie on them for a long time. |
| The hotel lobby was small, I felt crammed when we were checking in. |
| The gym offers little variety in equipment. More choices would have been better. |
| The best spots by the pool were always taken. |
| There was no one at the hotel of our age. |
| The food was always freshly made, amazing. |
| I loved the freshness of the food! |
| It was great that the buffet always had fresh food available. |
| I really enjoyed the food at the hotel, always fresh! |
| The best fresh food. Definitely a plus. |

**Questions**

Based on the reviews that you read, to what extent do you agree or disagree with the following statements?

| | Totally disagree | Somewhat disagree | Neither agree nor disagree | Somewhat agree | Totally agree |
|---|---|---|---|---|---|
| It is very likely that I will stay at this resort | ○ | ○ | ○ | ○ | ○ |
| I will certainly try the mentioned resort | ○ | ○ | ○ | ○ | ○ |
| There is a great chance that I will choose the mentioned resort | ○ | ○ | ○ | ○ | ○ |
| If you are reading this, mark somewhat agree | ○ | ○ | ○ | ○ | ○ |

**Demographics**

Please indicate your year of birth.

______________________________________________

Please select gender.

- Male
- Female
- Unspecified

Please indicate your level of education.

- Primary school
- Middle (junior high) school
- High school
- Bachelor's degree (undergraduate)
- Master's degree (graduate) or higher

# Appendix 3: Summary table of papers included in Chapter 4

Summary of Framework-Related Literature.

* = Bystander perspective

| Reference | Method | Independent variable(s) | Dependent variable(s) | Moderators | Mediators |
|---|---|---|---|---|---|
| Anderson and Han (2016) | Secondary data analysis | Response vs no response<br>Response to negative reviews<br>Response to positive reviews | Hotel revenue | | |
| Bach and Kim (2012) | Case study | Response vs no response<br>Accommodative vs defensive | Unspecified | | |
| Bhandari and Rodgers (2018) | Experimental | Response vs no response | *Purchase intention | Problem attribution in the review | Brand trust |
| Brunner et al. (2019) | Experimental | Managerial response vs consumer response | *Purchase intention | Brand equity | |
| Casado-Díaz et al. (2020) | Experimental | Defensive<br>Accommodative<br>No response | *Hotel attitude<br>*Booking intentions | Social Media type | |
| Cenni and Goethals (2020) | Content analysis | Thank<br>Apologize/ Express regret<br>Take responsibility<br>Refer to corrective actions<br>Offer explanations<br>Dismiss | Not applicable | Language/Culture | |

| Reference | Method | Independent variable(s) | Dependent variable(s) | Moderators | Mediators |
|---|---|---|---|---|---|
| | | Invite further contact<br>Solicit future visit | | | |
| Chang et al. (2015). | Experimental | Accommodative vs defensive | NWOM intention | | Attribution of locus and controllability<br>Reputation |
| Chen et al. (2019) | Secondary data analysis | Response vs no response (overall, and to positive and negative) | Future review volume<br>Future review valence | Response length | |
| Chevalier et al. (2018) | Secondary data analysis | Response vs no response (overall and to negative reviews) | Reviewer motivation to post | | |
| Colliander et al. (2015) | Experimental | Response vs no response | *Brand attitude<br>*Purchase intention | | |
| Crijns et al. (2017) | Experimental | Personalized vs corporate | *Organizational reputation | Review valence | Conversational Human Voice<br>Skepticism |
| Demmers et al. (2018) | Experimental | Response vs. no response | Satisfaction<br>Repurchase intentions | Message addressee<br>Message valence | Perceived usefulness<br>Perceived privacy violation |
| Dens et al. (2015) | Experimental | Refutation<br>Apology<br>Apology + prospective explanation<br>Apology + compensation<br>Apology + prospective explanation + compensation | *Reader's attitude<br>*Patronage intention<br>*PWOM intention | Review set balance | Perceived trust |

| Reference | Method | Independent variable(s) | Dependent variable(s) | Moderators | Mediators |
|---|---|---|---|---|---|
| Einwiller and Steilen (2015) | Secondary data analysis | Inquiring further information (attentiveness)<br>Gratitude (attentiveness)<br>Regret (attentiveness)<br>Corrective action<br>Explanation (credibility)<br>Active transfer<br>Passive transfer<br>Apology<br>Understanding (attentiveness) | Complaint satisfaction | | |
| Esmark Jones et al. (2018) | Experimental | Response vs no response<br>Managerial response vs consumer response | *Purchase intention<br>*Attitude toward the company | | |
| Gelbrich and Roschk (2011) | Meta-analysis | Compensation<br>Favorable employee behavior<br>Organizational procedures | Customer behavioral intentions: loyalty and positive WOM | | |
| Ghosh (2017) | Experimental | Explanation (accommodative) vs no explanation<br>Timely response vs late response | Loyalty | Review helpfulness | Consumer forgiveness |
| Tathagata and Amar (2018) | Experimental | Explanation<br>Signed response (owner, manager, etc.) vs response by team<br>Sidedness (accept some complaints and reject others) | *Brand attitude<br>*Purchase intention<br>*Satisfaction with webcare | Consumer forgiveness | Severity of failure |

| Reference | Method | Independent variable(s) | Dependent variable(s) | Moderators | Mediators |
|---|---|---|---|---|---|
| Gu and Ye (2014) | Secondary data analysis | Response vs No response (to negative review) | *Customer satisfaction | | |
| Ha and Jang (2009) | Questionnaire | Level of service recovery efforts (high/low) | Behavioral intentions | Relationship quality | Perceived justice |
| Herhausen et al. (2019) | Secondary data analysis | Compensation<br>Apology<br>Channel change | Virality | Level of arousal of the complainant | |
| Homburg et al. (2015) | Secondary data analysis and experimental | Response vs<br>No response | *Consumer sentiment | | Type and topic of conversation |
| Honisch and Manchón (2019) | Experimental | Reform<br>Humor<br>Refuse<br>Refute | *Organizational reputation<br>*Behavioral intentions | | |
| Huang and Ha (2020) | Experimental | Warmth-oriented responses<br>Competence-oriented responses | *Positive word-of-mouth intentions | Relationship orientation | Perceived diagnosticity<br>Perceived sincerity<br>Satisfaction with service recovery efforts |
| Javornik et al. (2020) | Experimental | Conversational Human vs Corporate Voice<br>Reply Length (short vs long) | *Observer's Satisfaction with Complaint Handling | | Perceived justice dimensions |
| Johnen and Schnittka (2019) | Experimental | Accommodative vs defensive | *Purchase intention | Perceived sought benefits | Reasoning in the complaint<br>Brand's communication style |

| **Reference** | **Method** | **Independent variable(s)** | **Dependent variable(s)** | **Moderators** | **Mediators** |
|---|---|---|---|---|---|
| | | | | | Hedonic vs. utilitarian benefits |
| Kim et al. (2015) | Secondary data analysis | Response vs<br>No response (negative reviews) | Hotel performance | | |
| Kim et al. (2016) | Secondary data analysis<br>Experimental | Apology vs no Apology | *Behavioral intentions | Reviewer vs bystander | |
| Kniesel et al. (2016) | Experimental | Response vs<br>No response<br>Human voice vs corporate voice<br>Manager vs<br>Staff response | *Likeability of the hotel<br>*Review credibility<br>*Review usefulness | | External attribution |
| Kwok and Xie (2016) | Secondary data analysis | Response vs<br>No response | Review helpfulness | | |
| Lappeman et al. (2018) | Experimental | Personalized vs standard response | *Brand reputation | | |
| Lee and Cranage (2014) | Experimental | Accommodative vs defensive | *Attitude change | Response strategy | External causal attribution |
| Lee and Song (2010) | Content analysis and experimental | Accommodative vs defensive | *Problem attribution to company<br>*Company evaluation | | |
| Li et al. (2018) | Field study<br>Experimental | Accommodative vs Defensive | Hotel sales revenue<br>*Purchase intention | Ordinary negative review vs Product review failure | Attribution of negative review towards brand |
| Liu and Ji (2019) | Text mining | Response length<br>Response voice: disputed voice, professional voice, Empathetic voice | *Perceived helpfulness | | |

| Reference | Method | Independent variable(s) | Dependent variable(s) | Moderators | Mediators |
|---|---|---|---|---|---|
| Liu, Jayawardhena, Dibb, et al. (2019) | Experimental | Response vs no response (negative reviews)<br>Compensation<br>Timeliness | eWOM continuance<br>Attitude towards the hotel | | Failure severity |
| Liu et al. (2015) | Secondary data analysis | Response vs no response vs targeted response | Hotel ratings | | Hotel class |
| Lui et al. (2018) | Secondary data analysis | No response vs Response vs Response only to extreme reviews | Firm's competitive performance | Review rating | |
| Ma et al. (2015) | Secondary data analysis | Response vs no response | *Sentiment index | | |
| Mate et al. (2019) | Content Analysis<br>Interviews | Response dimensions: timeliness, style, structure, source, credibility, attentiveness, values culture<br>Accommodative approach: acknowledge, explanation, apology, corrective actions, corrective statements, customer input, compensation<br>Defensive approach: dismiss, denial, avoidance, criticize, shift the blame | Not applicable | | |
| Min et al. (2015) | Experimental | Timeliness<br>Empathy<br>Paraphrasing | *Satisfaction with response | | |
| Nghiêm-Phú (2018) | Secondary data analysis | For positive comments:<br>Appreciation Agreement<br>Downgrade | Not applicable | | |

| **Reference** | **Method** | **Independent variable(s)** | **Dependent variable(s)** | **Moderators** | **Mediators** |
|---|---|---|---|---|---|
| | | Disagreement Question<br>Challenge<br>Shift credit<br>Informative comment<br>Ignorance<br>Legitimate evasion Request reassurance<br>For negative comments:<br>Gratis<br>Discount Coupon<br>Free upgrade Free ancillary<br>Managerial intervention<br>Replacement<br>Correction Substitution<br>Apology | | | |
| Park and Allen (2013) | Case study | Response frequency | Unspecified | | |
| Proserpio and Zervas (2017) | Secondary data analysis | Response vs no response | Future review ratings<br>Future review volume<br>Future review length | | |
| Raju (2019) | Experimental | Specific webcare vs vague webcare<br>High vs low webcare source credibility | *Perceived fairness | Reviewer reputation | |
| Roozen and Raedts (2018) | Experimental | Personalized vs. General | *Booking intention<br>*Recommendation intention | | |
| Rose and Blodgett (2016) | Experimental | Response vs no response | Company reputation | Apology with assurance vs | |

| Reference | Method | Independent variable(s) | Dependent variable(s) | Moderators | Mediators |
|---|---|---|---|---|---|
| | | | | apology with correction action | |
| Schamari and Schaefers (2015) | Experimental | Response vs no response<br>Personal vs impersonal response | *Consumer engagement | Platform type | |
| Sheng (2019) | Secondary data analysis | Response volume<br>Timeliness<br>Response length | Future review volume | | |
| Sheng et al. (2019) | Text analysis | Response vs no response<br>Timeliness<br>Response length<br>Response sentiment | Future review ratings | | Level of satisfaction<br>Previous experience |
| Sparks and Bradley (2017) | Interviews and content analysis | Acknowledge<br>Account<br>Take Action<br>Content attributes<br>Style characteristics | *Customer perceptions | | |
| Sparks et al. (2016) | Experimental | Response vs No response<br>Timely response vs late response<br>Conversational human voice vs professional voice | *Consumer inferences of trust and concern | | |
| Sreejesh et al. (2019) | Experimental | Perception of failure (high vs low)<br>Review agreement (high vs low) | *Attitude<br>*Patronage intentions | Webcare:<br>Apology + compensation + explanation<br>Apology + explanation<br>Apology + compensation | Justice perceptions |

| Reference | Method | Independent variable(s) | Dependent variable(s) | Moderators | Mediators |
|---|---|---|---|---|---|
| Stevens et al. (2018) | Literature review | Timeliness<br>Transparent response<br>Response that fosters trust | Effective management of complaints | | |
| Treviño and Castaño (2013) | Interviews | Denying<br>Accepting<br>Mention changes | *Purchase intentions | | |
| Ullrich and Brunner (2015) | Experimental | No response<br>Response by the brand<br>Response by other consumers | *Product purchase intention | | |
| Valentini et al. (2020) | Meta-analysis | Compensation with money<br>Compensation without money | Valence of emotions | | |
| Van Noort and Willemsen (2012) | Experimental | No webcare<br>Reactive webcare<br>Proactive webcare | *Brand evaluation | Human voice | Reactive vs proactive webcare<br>Platform type |
| Wang and Chaudhry (2018) | Secondary data analysis | No response<br>Responding to all reviews<br>Responding to negative reviews<br>Responding to positive reviews | *Future review ratings | Review platform<br>Response tailoring | |
| Wei et al. (2013) | Experimental | Personalized vs standard | *Perceived communication quality<br>*Trust towards the response | Review valence | |
| Weitzl (2019) | Survey<br>Experimental | No response<br>Defensive response<br>Accommodative response | *Favorable brand-related outcomes | Complainant-type (constructive vs vindictive) | Webcare reaction |

| Reference | Method | Independent variable(s) | Dependent variable(s) | Moderators | Mediators |
|---|---|---|---|---|---|
| | | | *Unfavorable Brand-related outcomes | | |
| Weitzl and Einwiller (2020) | Survey<br>Experimental | No response<br>Defensive response<br>Accommodative response | *Future NWOM | Complainant type | |
| Weitzl and Hutzinger (2017) | Experimental | Managerial response vs consumer response<br>Accommodative vs no response<br>Defensive vs no response | *Favorable band reactions<br>*Unfavorable brand reactions | | Webcare credibility |
| Weitzl et al. (2018) | Survey<br>Quasi- experimental | No response<br>Defensive response<br>Accommodative response | *Post webcare satisfaction<br>*Post NWOM | Prior failure experiences (few vs multiple)<br>Advocate-initiated webcare (no response vs defensive response) | Failure attributions |
| J. Wu et al. (2020) | Experimental | Passive-constructive response<br>Active-constructive response | Repurchase intention | Response style (official vs friendly) | Positive affect |
| Xia (2013) | Questionnaire and experimental | Vulnerable (accommodative) response<br>Defensive response | *Satisfaction<br>*Purchase intention<br>*PWOM intention | Sincerity<br>Respect<br>Appropriateness | Brand personality |
| Xie et al. (2016) | Secondary data analysis | No response vs Response | Future review ratings<br>Future review volume<br>Hotel sales | | Response ratio<br>Hotel class |

| Reference | Method | Independent variable(s) | Dependent variable(s) | Moderators | Mediators |
|---|---|---|---|---|---|
| Xie et al. (2017) | Secondary data analysis | Timeliness<br>Response length<br>Response by executives or by staff<br>Repeat (or no repeat) of the topics in the review) | Financial performance (revenue, average daily rate (ADR), and occupancy) | | Review rating<br>Review volume |
| Xie et al. (2014) | Secondary data analysis | Response vs<br>No response | Hotel performance | | |
| Zhang and Vásquez (2014) | Genre analytic research | Express gratitude<br>Apologize<br>Invitation for a second visit<br>Opening pleasantries<br>Proof of action<br>Acknowledge feedback<br>Refer to customer reviews<br>Closing pleasantries<br>Avoidance of reoccurring problems<br>Solicit response | Not applicable | | |
| Zhang et al. (2019) | Experimental | Explanation type (explained action or explained reaction)<br>Response channel (public vs private space) | Consumer expectations | Review valence | Perceived usefulness |

# Appendix 4: Operationalization of the webcare variables used in chapter 5

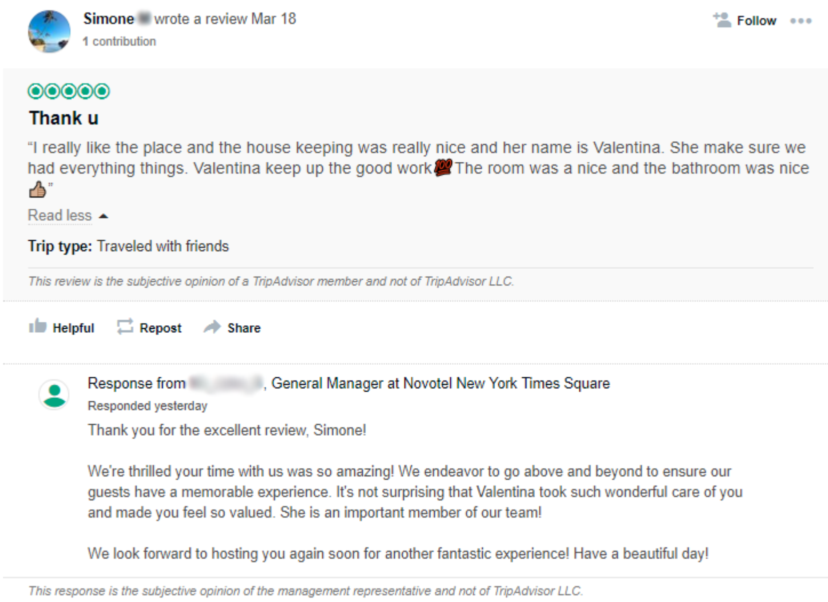
This section provides details on how the variables used in this study were created. The procedures can be divided into two types: (1) variables requiring manual coding (therefore a detailed codebook with examples is needed) in order to develop a training set for machine learning text classification, and (2) variables created automatically based on instructions.

## *Variables requiring manual coding*

We created the codebook below as guidelines for the researcher that coded the sample and by the coder that coded the subsample to calculate the intercoder reliability. This codebook comprises both the definition of the constructs, examples, and instructions for the coding.

### Coding instructions: Webcare strategies

This codebook contains definitions and coding instructions for the different webcare strategies found in relevant academic literature.



The purpose of this codebook is to analyze the text of the managerial responses. We provide an Excel file containing the original reviews and responses as well as the English translation.

- Each response is coded for all variables. This codebook provides information on each variable and general coding guidelines for each category, along with examples.
- The general rules for coders are:
    - Each judgment is mostly binary (0/1). That is, coders choose whether content falls into each category or not.
    - Content can be assigned to multiple topics. For example, a response can contain both "apology" and "explanation". A 0/1 judgment on one issue should not affect a judgment on a different topic labeling on the same level.
    - Although the coder might have prior experience reading or writing reviews for other hotels, or if the hotel is known, this prior information should not be used to attribute the codes. The code should only be based on each response individually.
- Whenever necessary, we provide examples of the variables/constructs.

**a) Who responds to the review**

The person responding is the one who signs the response. Several sub-categories can be selected within the options below. For instance, the reply might include the position, e.g., general manager, and the name of the manager. We create indicator (0-1) variables for whether the response was signed with a name, by someone with a manager title, someone with a staff title, the "hotel" and "department x." For example:

| Not signed: | Signed with own name: | Signed by hotel manager: | Signed by staff member | Signed with hotel name | Signed with department title |
|---|---|---|---|---|---|
| *Dear guest, Thank you for your comment!* | *Dear guest, Thank you for your comment! Sincerely, John Doe* | *Dear guest, Thank you for your comment! Sincerely, General Manager* | *Dear guest, Thank you for your comment! Sincerely, John Doe, Front Desk* | *Dear guest, Thank you for your comment! Sincerely, The Best Suits Hotel* | *Dear guest, Thank you for your comment! Sincerely, Social media team* |

**b) Invitation for another visit**

The invitation for another visit is when hotels invite their guests to return to the hotel, making their guests feel like they are welcome. For example:

| Invites another visit | Does not invite for another visit |
|---|---|
| *Dear,*<br>*Thank you for your comment. I'm glad that you enjoyed your stay. We hope to have you again in our hotel in the future!* | *Dear,*<br>*Thank you for your comment. I'm glad that you enjoyed your stay.* |
| *Dear,*<br>*Thank you for your comment. You are always welcome at our hotel!* | *Dear,*<br>*Thank you for your comment. Will you visit us again soon?* |

**c) Non-verbal cues**

The tone of voice can be described by non-verbal characteristics. The following characteristics can be coded for its presence (1) and absence (0): abbreviations, emoticons, words in upper case, onomatopoeia, sound stretching, and using punctuation to pass a message. Examples:

| Abbreviations | Emoticons | Words in upper case | Onomatopoeia | Sound stretching | Punctuation |
|---|---|---|---|---|---|
| *LOL, DM, PM, OMG* | ☺, ☹ | *SUPER, WOW, AMAZING* | *Haha* | *Veeeery* | *..., !!!, ???, ??!, (use only) ?, !* |

**d) Provide explanation**

When hotels receive a negative review, they might provide an explanation for what happened. Webcare also comprises an explanation when it gives instructions to guests on how they should have proceeded to avoid the reported issue. It also consists of an explanation when the guests compliment a certain aspect of the stay (e.g., 'good croissants in the breakfast buffet'), and the response explores this compliment by explaining the feature (e.g., 'our croissants are baked every morning by our talented guest chef'). As such, the explanation does not need to always be explicit (e.g., 'we had no heating because of an issue in the central system'). It could explain details regarding a specific issue mentioned by the review. Examples are as follows:

| Explanation | No explanation |
|---|---|
| *Dear customer,*<br>*Thank you for your comment. We are sorry to hear that you were not satisfied with your stay. The roadworks are temporarily causing some noise problems, but we expect that it will be finished in a week.* | *Dear guest,*<br>*Thank you for your comment. We are sorry to hear that you were not satisfied with your stay because of the noise.* |
| *Dear,*<br>*Thank you for your review. I am sorry that you couldn't find the remote, but there was a remote available in the room. It was inside the drawer in the bedside table.* | |
| *Dear,*<br>*I am glad to hear that you found our beds comfortable. We use the greatest linen available and the mattress is made for us by order.* | |
| *Dear,*<br>*Our location is indeed very good. We are just 200 meters from the train station and you find the nicest street shop just 50 meters from the hotel door.* | |

**e) Defensiveness**

A response is defensive when it denies or refutes the review. Management may argue that the events described by the dissatisfied customer are untrue. Arguing against the review by saying that it is the first complaint the hotel receives about the issue or that other guests liked what is criticized is a way of refuting what was written. The same can be made by presenting 'facts' that put in jeopardy what is written in the review. Examples are as follows:

| Defensive | Not defensive |
|---|---|
| *Dear customer,*<br>*We are sorry that the noise bothered you but we did not receive any complaints from other guests. Besides, our staff that was working that night says that they did not hear anything.*<br><br>*Dear guest,*<br>*We are sorry that you considered our hotel too far from the city center. Our guests usually don't mind to walk there from the hotel, so I guess this is a matter of opinion.*<br><br>*Dear guest,*<br>*I was surprised that you found our rooms too small because they are is 25m2.* | *Dear guest,*<br>*Thank you for your comment. We are sorry to hear that you were not satisfied with your stay because of the noise.* |

**f) Response Tailoring**

Assess if the topics in the review and mentioned in the response. For instance, if the review mentions the quality of the bed and the response mentions the quality of the bed that is tailoring. If the review mentions the quality of the bed and the response only thanks for the comment, that does not constitute tailoring. It is different from the variable explanation, as for the tailoring no additional information needs to be provided.

## *Variables automatically coded*

This section presents the variables that are coded using automated methods since they do not require human interpretation.

**a) Response vs No response**

For each review, either there is a response or there is not a response. If the response field is empty (no response, 0), no further coding is naturally necessary. We can use a simple R function to detect if the response field is empty and code as follows:

**b) Response timeliness**

Create dummy variable for the number of days between review and response.

**c) Offer apology**

Use the dictionary developed by Herhausen et al. (2019) to find compensation words (*accept, accident*, acknowledg*, admit, agree, apolog*, approv*, assert, conced*, confess*, excus*, forgiv*, guilt, guilty, pardon, recogni*, sorry*). For this category, we can count the total number of times one or more apology words are used in a response.

**d) Offer compensations**

Use the dictionary developed by Herhausen et al. (2019) to find compensation words (*compensat*, offer*, recover*, refund*, reimburs*, repay*, restor*, return**).

**e) Channel change**

Use the dictionary developed by Herhausen et al. (2019) to find change channel words (*call, chat*, contact, correspondence, email, e-message, letter, mail, offline, phone, ring, visit, voic**).

**f) Showing gratitude**

Use a dictionary to find gratitude words within the response (*thank, grateful, appreciat*).*

**g) Inquiring further information**

It is frequent that hotels show attentiveness by inquiring further information in a complaint, for instance. As such, we can look for the presence or question marks (?) automatically.

**h) Personalization**

This category refers to when the response addresses the reviewer by their own name (first name, last name or both names). Operationalized by looking for the word after 'Dear' to see if there is a name that matches the name in the reviewer column.

# Appendix 5: Performance of different text classification methods used in chapter 5

The table below presents the accuracy and AUC scores achieved by each machine learning text classifier for the different variables with a training set of $n$=810 managerial responses. Besides the average, the minimum and maximum values presented for support vector machines (SVM), boosted trees, random forests, and naïve Bayes represent the lowest and highest values achieved in 10-fold cross-validation. For BERT, the learning rate is adjusted based on empirical simulations for each variable (2e-4 for defensiveness and 2e-5 for the other variables). For the other parameters, we used a batch size of 12 with a max length of 256 words and a validation percentage of 20% of the dataset. The number of training epochs is determined by an automatic stopping rule that terminated training when performance plateaued on the training set and saved the best performing model on the validation set. For defensiveness, we used a balanced training set (same number of cases for each label) to increase BERT's performance and prevent unintended bias since that variable is particularly unbalanced. This is a common practice, as can be seen in previous research (e.g., Dixon et al., 2018). The Python code used for SVM, boosted trees, random forests, naïve Bayes and BERT can be found in https://github.com/aisabel1/webcare.

| | | **Accuracy** | | | **AUC** | | |
|---|---|---|---|---|---|---|---|
| **Variable** | **Classifier** | *Min* | *Max* | *Avg* | *Min* | *Max* | *Avg* |
| Tailoring | SVM | .83 | .96 | .89 | .50 | .95 | .77 |
| | Boosted trees | .71 | .98 | .89 | .54 | .99 | .77 |
| | Random forests | .77 | .98 | .90 | .54 | .99 | .79 |
| | Naïve Bayes | .74 | .91 | .83 | .50 | .78 | .63 |
| | BERT | - | - | **.95** | - | - | **.98** |
| Explanation | SVM | .68 | .84 | .76 | .64 | .83 | .74 |
| | Boosted trees | .65 | .84 | .76 | .67 | .82 | .75 |
| | Random forests | .67 | .90 | .75 | .51 | .84 | .71 |
| | Naïve Bayes | .68 | .85 | .75 | .60 | .86 | .72 |
| | BERT | - | - | **.88** | - | - | **.93** |
| Invitation for visit | SVM | .45 | .94 | .73 | .52 | .92 | .68 |
| | Boosted trees | .51 | .94 | .73 | .54 | .92 | .68 |
| | Random forests | .37 | .89 | .70 | .53 | .90 | .68 |
| | Naïve Bayes | .37 | .94 | .70 | .52 | .93 | .65 |
| | BERT | - | - | **.85** | - | - | **.90** |
| Defensive | SVM | .89 | .91 | .90 | .50 | .50 | .50 |
| | Boosted trees | .88 | .93 | .89 | .49 | .56 | .52 |
| | Random forests | .90 | .91 | .91 | .50 | .56 | .51 |
| | Naïve Bayes | .86 | .93 | .90 | .49 | .70 | .57 |
| | BERT* | - | - | **.84** | - | - | **.93** |
| Non Verbal cues | SVM | .96 | .98 | .97 | 0.50 | .49 | .50 |
| | Boosted trees | .91 | 1 | .97 | 0.50 | 1 | .75 |
| | Random forests | .83 | .98 | .96 | 0.50 | .91 | .54 |
| | Naïve Bayes | .96 | .98 | .97 | 0.49 | .50 | .50 |
| | BERT | - | - | **.99** | - | - | **.99** |
| Signed with department | SVM | .78 | 1 | .95 | .58 | 1 | .92 |
| | Boosted trees | .78 | 1 | .95 | .58 | 1 | .92 |
| | Random forests | .78 | 1 | .95 | .58 | 1 | .94 |
| | Naïve Bayes | .63 | .98 | .92 | .58 | 1 | .89 |
| | BERT | - | - | **.98** | - | - | **1** |
| Signed with hotel | SVM | .85 | 1 | .95 | .50 | 1 | .86 |
| | Boosted trees | .84 | 1 | .97 | .53 | 1 | .92 |
| | Random forests | .86 | 1 | .96 | .50 | 1 | .92 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Naïve Bayes | .86 | 1 | .91 | .50 | 1 | .73 |
| | BERT | - | - | **.99** | - | - | **.99** |
| Signed by manager | SVM | .71 | 1 | .92 | .88 | 1 | .94 |
| | Boosted trees | .85 | 1 | .95 | .86 | 1 | .95 |
| | Random forests | .85 | 1 | .94 | .88 | 1 | .96 |
| | Naïve Bayes | .74 | 1 | .88 | .74 | 1 | .88 |
| | BERT | - | - | **.96** | - | - | **.99** |
| Signed with name | SVM | .86 | 1 | .94 | .77 | 1 | .91 |
| | Boosted trees | .89 | 1 | .93 | .76 | 1 | .95 |
| | Random forests | .89 | .99 | .96 | .85 | 1 | .95 |
| | Naïve Bayes | .69 | 1 | .90 | .57 | 1 | .86 |
| | BERT | - | - | **.98** | - | - | **.99** |
| Signed by staff | SVM | .78 | 1 | .95 | .65 | 1 | .91 |
| | Boosted trees | .78 | 1 | .96 | .65 | 1 | .93 |
| | Random forests | .79 | 1 | .95 | .65 | 1 | .94 |
| | Naïve Bayes | .63 | 1 | .93 | .62 | 1 | .89 |
| | BERT | - | - | **.98** | - | - | **.99** |

*Using a balanced dataset (with the same amount of cases for each label)

Looking at the table, it is clear that BERT outperforms the other text classification methods, both for simple identification tasks (e.g., use of the name in the signature) as well as for more abstract classifications (e.g., defensiveness). These results for BERT are in line with previous research that shows the superiority of BERT over other text classifiers (González-Carvajal & Garrido-Merchán, 2020). The results from SVM, boosted trees, random forests, and naïve Bayes are never too discrepant, contrary to what is found in previous literature (Hartmann et al., 2019).

<u>An experience with reinforced boosted trees</u>

Before we implemented BERT text classification, we experimented with reinforced boosted trees (RBT) since they seemed to yield slightly better results than the other classification methods for some of the variables. Therefore, we proceeded with fine-tuning the parameters for the variables that displayed less than optimal performance with the other classifiers (tailoring, explanation, invitation for visit, defensiveness, and non-verbal cues). The interaction depths of 1, 2, 3 and 4 and the learning rates of .01, .05, .1, .2 were tested for each variable.

The table below compares the baseline parameters (reported in the previous table) with the improved parameters.

| Variable | Parameters boosted trees | AUC Average 10 splits |
|---|---|---|
| Tailoring | Baseline: learning rate 0,1/ interaction depth 3 | .77 |
| | Improved: learning rate 0,05/ interaction depth 3 | **.79** |
| Explanation | Baseline: learning rate 0,1/ interaction depth 3 | .75 |
| | Improved: learning rate 0,1/ interaction depth 2 | **.76** |
| Invitation for visit | Baseline: learning rate 0,1/ interaction depth 3 | .68 |
| | Improved: learning rate 0,05/ interaction depth 4 | **.71** |
| Defensive | Baseline: learning rate 0,1/ interaction depth 3 | .52 |
| | Improved: learning rate 0,2/ interaction depth 3 | **.54** |
| Non Verbal cues | Baseline: learning rate 0,1/ interaction depth 3 | .75 |
| | Improved: learning rate 0,05/ interaction depth 4 | **.78** |

The table shows that AUC for all the variables improves when the learning rate parameters and interaction depth were fine-tuned. The only variable with better AUC using another classifier (that is not BERT) is defensiveness, where naïve Bayes performed better (.57). Still, BERT outperforms RFT for all variables.

# Appendix 6: Correlation matrix, factor analysis, and variable selection performed for chapter 5

Factor analysis of signature variables

| Variable | RC1 | RC2 | RC3 |
|---|---|---|---|
| sigdepart | .98 | | |
| sigstaff | .98 | | |
| signame | .45 | .87 | |
| sigmanager | | .96 | |
| sighotel | | | .99 |

| Variable | RC1 | RC2 | RC3 |
|---|---|---|---|
| SS loadings | 2.18 | 1.70 | 1.01 |
| Proportion Var | .435 | .34 | .202 |
| Cumulative Var | .435 | .775 | .977 |

Correlation matrix webcare variables

| | **nextbook** | **book** | **respond** | **tailor** | **defensive** | **invitevisit** | **explain** | **nonverb** | **apology** | **compensate** | **chachange** | **gratitude** | **info** | **personal** | **sigNameMgr** | **sigDepStaf** | **sighotel** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **nextbook** | 1 | **.82** | .15 | .36 | .4 | .22 | .33 | -.1 | .4 | .27 | .13 | .23 | .37 | .42 | .44 | .35 | .01 |
| **book** | .82 | 1 | .18 | .36 | .37 | .25 | .31 | -.06 | .38 | .24 | .12 | .26 | .31 | .38 | .32 | .28 | .12 |
| **respond** | .15 | .18 | 1 | .7 | .51 | .63 | .57 | .17 | .45 | .36 | .27 | .56 | .26 | .6 | .52 | .18 | .25 |
| **tailor** | .36 | .36 | .7 | 1 | **.83** | **.88** | **.89** | .15 | .74 | .63 | .44 | .79 | .33 | **.88** | .76 | .42 | .17 |
| **defensive** | .4 | .37 | .51 | .83 | 1 | .69 | **.87** | .07 | .74 | .63 | .41 | .64 | .38 | .75 | .67 | .41 | .14 |
| **invitevisit** | .22 | .25 | .63 | .88 | .69 | 1 | **.81** | .21 | .61 | .6 | .42 | **.81** | 0 | .79 | .62 | .46 | .13 |
| **explain** | .33 | .31 | .57 | .89 | .87 | .81 | 1 | .13 | .73 | .66 | .43 | .73 | .26 | .78 | .7 | .49 | .07 |
| **nonverb** | -.1 | -.06 | .17 | .15 | .07 | .21 | .13 | 1 | 0 | .09 | .02 | .21 | .06 | .07 | .02 | 0 | .13 |
| **apology** | .4 | .38 | .45 | .74 | .74 | .61 | .73 | 0 | 1 | .62 | .35 | .62 | .31 | .71 | .6 | .38 | .09 |
| **compensate** | .27 | .24 | .36 | .63 | .63 | .6 | .66 | .09 | .62 | 1 | .31 | .55 | .16 | .55 | .5 | .4 | .02 |
| **chachange** | .13 | .12 | .27 | .44 | .41 | .42 | .43 | .02 | .35 | .31 | 1 | .34 | .12 | .38 | .35 | .21 | .06 |
| **gratitude** | .23 | .26 | .56 | .79 | .64 | .81 | .73 | .21 | .62 | .55 | .34 | 1 | .1 | .71 | .52 | .44 | .18 |
| **info** | .37 | .31 | .26 | .33 | .38 | 0 | .26 | .06 | .31 | .16 | .12 | .1 | 1 | .34 | .39 | -.06 | .15 |
| **personal** | .42 | .38 | .6 | .88 | .75 | .79 | .78 | .07 | .71 | .55 | .38 | .71 | .34 | 1 | .79 | .43 | .05 |
| **sigNameMgr** | .44 | .32 | .52 | .76 | .67 | .62 | .7 | .02 | .6 | .5 | .35 | .52 | .39 | .79 | 1 | .36 | -.25 |
| **sigDepStaf** | .35 | .28 | .18 | .42 | .41 | .46 | .49 | 0 | .38 | .4 | .21 | .44 | -.06 | .43 | .36 | 1 | -.1 |
| **sighotel** | .01 | .12 | .25 | .17 | .14 | .13 | .07 | .13 | .09 | .02 | .06 | .18 | .15 | .05 | -.25 | -.1 | 1 |

Factor analysis of webcare strategies (except signature variables)

| | **RC1** | **RC2** | **RC6** | **RC4** | **RC5** | **RC3** |
|---|---|---|---|---|---|---|
| tailor | .87 | | | | | |
| defensive | .65 | | .50 | | | |
| invitevisit | .90 | | | | | |
| explain | .76 | | .40 | | | |
| gratitude | .86 | | | | | |
| personal | .86 | | | | | |
| information | | .98 | | | | |
| apology | .52 | | .75 | | | |
| chachange | | | | .96 | | |
| compensate | .42 | | | | .87 | |
| nonverbal | | | | | | .99 |

| | **RC1** | **RC2** | **RC6** | **RC4** | **RC5** | **RC3** |
|---|---|---|---|---|---|---|
| SS loadings | 4.55 | 1.20 | 1.20 | 1.10 | 1.10 | 1.03 |
| Proportion Var | .41 | .11 | .11 | .10 | .10 | .09 |
| Cumulative Var | .41 | .52 | .63 | .73 | .83 | .93 |

Panel Poisson regression model containing all webcare variables

| **Coefficients:** | | | | | **VIFs** |
|---|---|---|---|---|---|
| | **Estimate** | **Std. Error** | **z value** | **Pr(>\|z\|)** | |
| (Intercept) | 2.15 | 0.06 | 37.36 | < 2e-16 *** | |
| Spline(time) 5 df | | | | | 2.73 |
| factor(hotel_id)2 | -0.09 | 0.01 | -7.75 | 9.17e-15 *** | 67.96 |
| factor(hotel_id)3 | -1.21 | 0.03 | -37.83 | < 2e-16 *** | |
| factor(hotel_id)4 | -0.86 | 0.02 | -36.56 | < 2e-16 *** | |
| factor(hotel_id)5 | -0.68 | 0.02 | -34.68 | < 2e-16 *** | |
| factor(hotel_id)6 | -0.03 | 0.01 | -1.91 | .06 . | |
| factor(hotel_id)7 | -0.65 | 0.03 | -25.75 | < 2e-16 *** | |
| book | 0.61 | 0.01 | 78.16 | < 2e-16 *** | 4.84 |
| respond | 0.08 | 0.01 | 6.39 | 1.62e-10 *** | 3.46 |
| tailor | 0.04 | 0.01 | 2.70 | .007 ** | 23.74 |
| defensive | 0.03 | 0.01 | 3.48 | .0005 *** | 6.89 |
| invitevisit | -0.02 | 0.01 | -2.71 | .007 ** | 9.55 |
| explain | -0.02 | 0.01 | -2.096 | .04 * | 11.096 |
| nonverbal | -0.05 | 0.01 | -3.64 | .0003 *** | 1.10 |
| apology | -0.04 | 0.01 | -5.67 | 1.47e-08 *** | 3.51 |
| compensate | 0.02 | 0.01 | 1.96 | .05 | 2.15 |
| chachange | 0.04 | 0.01 | 3.87 | .0001 *** | 1.36 |
| gratitude | 0.0001 | 0.01 | 0.01 | .99 | 4.39 |
| information | -0.04 | 0.01 | -3.62 | .0003 *** | 4.48 |
| personal | 0.004 | 0.01 | 0.38 | .70 | 10.60 |
| sigDepStaf | -0.003 | 0.004 | -0.71 | .48 | 2.72 |
| sighotel | -0.06 | 0.01 | -5.49 | 4.04e-08 *** | 4.05 |
| sigNameMgr | -0.02 | 0.01 | -2.11 | .035 * | 17.92 |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05

Null deviance: 120,407 on 1,143 degrees of freedom

Residual deviance: 9,955 on 1,116 degrees of freedom

# 9. Dutch Summary

# Online reviews en hoe ze te managen. De effecten van online word-of-mouth en webcare op consumentenreacties en bedrijfsresultaten

Er is veel onderzoek verricht naar het onderwerp eWOM (electronic Word of Mouth), dat kan worden gedefinieerd als "door de consument gegenereerde, consumptiegerelateerde communicatie die gebruik maakt van digitale hulpmiddelen en voornamelijk gericht is tot andere consumenten" (Rosario et al., 2020, p. 427). Online reviews, d.w.z. online productevaluaties door gebruikers of experts, worden algemeen beschouwd als een vorm van eWOM (bv., Zhang et al., 2010). Daarom worden in deze thesis de termen eWOM en online reviews door elkaar gebruikt.

Het doel van het proefschrift is tweeledig: het bestuderen van de invloed van verschillende eWOM-kenmerken op de reacties van consumenten; en de effecten van webcare - d.w.z. het reageren op online recensies - op de reacties van consumenten en de bedrijfsprestaties. De eerste hoofddoelstelling van deze thesis is te achterhalen welke online review cues belangrijk zijn om de geloofwaardigheid en nuttigheid van een review te bepalen en hoe de samenstelling en inhoud van een set van reviews de reacties van consumenten beïnvloeden. De eerste twee empirische hoofdstukken van het proefschrift gaan in op deze doelstelling.

In hoofdstuk 2 analyseren we het relatieve belang van argumentsterkte, argumentzijdigheid, schrijfkwaliteit, aantal argumenten, beoordeelde bruikbaarheid van de review, samenvattende reviewrating en aantal reviews bij het bepalen van de waargenomen nuttigheid en geloofwaardigheid van een online review. Daarnaast gebruiken we inzichten uit het Elaboration Likelihood Model (ELM) (Richard E Petty & John T Cacioppo, 1986) om het effect van de betrokkenheid van consumenten bij de productcategorie op het relatieve belang van de cues te onderzoeken. Een conjunctanalyse (N= 287) is gebruikt om het relatieve belang van de zeven eerder genoemde attributen te bestuderen. Een gebalanceerd orthogonaal ontwerp genereerde acht kaarten die overeenkomen met individuele beoordelingen. Respondenten scoorden alle acht kaarten in willekeurige volgorde voor waargenomen nuttigheid en geloofwaardigheid. Over het algemeen is de sterkte van de argumenten het belangrijkst, terwijl de samenvattende beoordeling en het aantal beoordelingen het minst belangrijk zijn voor de waargenomen nuttigheid en geloofwaardigheid van de beoordeling. Het aantal argumenten is belangrijker voor mensen die meer betrokken zijn bij het product, terwijl de schrijfkwaliteit en de beoordeling van het nut relatief belangrijker zijn voor de groep met een lage betrokkenheid. Deze studie biedt een uitgebreide test van hoe consumenten online reviews waarnemen, omdat het, voor zover wij weten, de eerste is die tegelijkertijd een grote set van cues onderzoekt met behulp van conjunctanalyse. Deze methode maakt de impliciete waardering (utility) van de individuele cues mogelijk, en onthult het relatieve belang van de cues, in een setting die dicht in de buurt komt van een real-life context. Daarnaast worden inzichten van het Elaboration Likelihood Model (ELM) gebruikt om te begrijpen hoe het relatieve belang van cues verschilt afhankelijk van de mate van betrokkenheid van reviewlezers bij de productcategorie.

Review set valentie (de mate van negativiteit of positiviteit van een set online reviews) bepaalt sterk de reacties van review lezers (Purnawirawan et al., 2015). Eerder onderzoek heeft voornamelijk gekeken naar alleen het aantal positieve en negatieve reviews om de valentie van een review set te bepalen. Daarom heeft hoofdstuk 3 als doel om te bestuderen hoe het verhogen van het aantal belangrijke positieve recensies de hotelverblijfsintentie van lezers beïnvloedt, waarbij we het 'omslagpunt' onderzoeken waarop belangrijke positieve recensies het negatieve effect van een groter aantal minder belangrijke negatieve recensies compenseren. We onderzoeken verder of de reacties van lezers positiever zijn wanneer alle positieve recensies betrekking hebben op hetzelfde productkenmerk of op verschillende kenmerken. We presenteren een 4 (review set valentie) x 2 (attribuut herhaling vs. verschillende attributen voor de positieve reviews) online experiment (N=408). De resultaten tonen aan dat een positievere review set alleen leidt tot een hogere intentie om in een hotel te verblijven wanneer de positieve reviews verschillende attributen bespreken (en niet hetzelfde attribuut herhalen). Het 'omslagpunt' waarop positieve recensies de negatieve compenseren is vier positieve recensies over verschillende kenmerken in een set van 12. Deze studie nuanceert het bandwagon effect, negativiteitsbias en waarheidseffect door aan te tonen dat overwegend negatieve recensiesets positief kunnen worden beoordeeld mits ze voldoende gevarieerde positieve reviews over belangrijke productkenmereken bevatten.

Met het toenemende volume van eWOM is het cruciaal voor organisaties om te weten hoe zij hun inspanningen in webcare moeten investeren om positieve bedrijfsresultaten te behalen (Schamari & Schaefers, 2015; Williams & Buttle, 2011). Organisaties worstelen echter vaak met de vraag welke webcare strategieën ze moeten inzetten om de beste resultaten te behalen (Van Noort et al., 2015). Daarom is de tweede hoofddoelstelling van dit proefschrift het opbouwen van een raamwerk voor webcare en het onderzoeken van de effecten van verschillende webcare strategieën op de bedrijfsreultaten. De hoofdstukken 4 en 5 van dit proefschrift beogen aan deze doelstelling te voldoen door de resultaten van een literatuuronderzoek te presenteren (hoofdstuk 4) en door te onderzoeken hoe specifieke webcare strategieën van invloed zijn op hotelboekingen (hoofdstuk 5).

Hoofdstuk 4 analyseert de gepubliceerde literatuur over webcare en managementreacties op online reviews om een raamwerk te bieden dat eenduidige conclusies uit vorig onderzoek wil identificeren, mogelijke verklaringen voor inconsistenties die verder onderzoek vereisen wil bespreken, en de 'te weinig onderzochte gebieden betreffende de managementreacties op online reviews wil identificeren. Dit raamwerk beantwoordt verschillende praktische en theoretische vragen over eWOM (electronic Word-of-Mouth). Moeten managers reageren op eWOM of niet? Als ze reageren, op welk soort eWOM moeten ze dan reageren en welke strategieën moeten ze gebruiken: wie moet reageren, wanneer, op welke platformen, in welke stijl? Hoe moeten ze specifiek reageren op negatieve recensies? Toekomstig onderzoek moet de vele tegenstrijdige effecten ontwarren (bv. wanneer defensieve webcare te gebruiken) en onderbelichte onderwerpen behandelen (bv. webcare strategieën voor Positive WOM specifiek of de onderliggende mechanismen die de effecten van verschillende webcare strategieën verklaren).

Gebaseerd op het raamwerk ontwikkeld in hoofdstuk 4, bouwt hoofdstuk 5 voort op de vaak schaarse of inconsistente bevindingen met betrekking tot de effecten van specifieke webcare strategieën op de bedrijfsprestaties om te testen of en hoe verschillende webcare strategieën hotelboekingen beïnvloeden. Na het testen van verschillende machine learning classifiers, levert BERT (Bidirectional Encoder Representations from Transformers) de beste prestatie voor het classificeren van webcare variabelen. De strategieën die een positief effect hebben op boekingen zijn het doorverwijzen van reviewers naar een privékanaal, defensief zijn, compensatie bieden en managers de reactie laten ondertekenen. Webcare strategieën die vermeden moeten worden zijn excuses, alleen om meer informatie vragen, klanten uitnodigen voor een volgend bezoek, en het toevoegen van informele non-verbale signalen. Strategieën die geen invloed lijken te hebben op toekomstige boekingen zijn het uiten van dankbaarheid, personaliseren, en het laten ondertekenen van webcare door personeelsleden (in plaats van managers). Deze bevindingen helpen hotelmanagers om hun webcare strategie te optimaliseren voor betere bedrijfsresultaten en om geautomatiseerde webcare te ontwikkelen.

Deze dissertatie voorziet in een leidraad voor verder onderzoek en voor praktijkmensen door te onderzoeken hoe mensen eWOM-informatie verwerken en hoe antwoorden op online reviews moeten worden gemanaged om negatieve effecten van eWOM te vermijden en positieve merkeffecten te stimuleren. Dit advies is uiterst relevant voor managers en marketeers, aangezien het volume van eWOM, en online reviews in het bijzonder, blijft toenemen en het beheren ervan een uitdaging wordt, vooral omdat eWOM een sterke invloed heeft op het gedrag van consumenten.